ABSTRACT

Title of Dissertation: EFFECTS OF AGE ON CONTEXT BENEFIT FOR UNDERSTANDING COCHLEAR-IMPLANT PROCESSED SPEECH

Anna R. Tinnemore, Doctor of Philosophy,

2024

Dissertation directed by: Professor Sandra Gordon-Salant, Hearing and Speech Sciences
Professor Matthew J. Goupell, Hearing and Speech Sciences

The number of people over 65 years old in the United States is rapidly growing as the generation known as “Baby Boomers” reaches this milestone. Currently, at least 16 million of these older adults struggle to communicate effectively because of disabling hearing loss. An increasing number of older adults with hearing loss are electing to receive a cochlear implant (CI) to partially restore their ability to

communicate effectively. CIs provide access to speech information, albeit in a highly degraded form. This degradation can frequently make individual words unclear. While predictive sentence contexts can often be used to resolve individual unclear words, there are many factors that either enhance or diminish the benefit of sentence contexts. This dissertation presents three complementary studies designed to address some of these factors, specifically: (1) the location of the unclear word in the context sentence, (2) how much background noise is present, and (3) individual factors such as age and hearing loss.

The first study assessed the effect of context for adult listeners with acoustic hearing when a target word is presented in different levels of background noise at the beginning or end of sentences that vary in predictive context. Both context sentences and target words were spectrally degraded as a simulation of sound processed by a CI. The second study evaluated how listeners with CIs use context under the same conditions of background noise, sentence position, and predictive contexts as the group with acoustic hearing. The third study used eye-tracking methodology to infer information about the real-time processing of degraded speech across ages in a group of people who had acoustic hearing and a group of people who used CIs. Results from these studies indicate that target words at the beginning of the context sentence are more likely to be interpreted to be consistent with the following context sentence than target words at the end of the context sentences. In addition, the age of the listener interacted with some of the other experimental variables to predict phoneme categorization performance and response times in both listener groups. In the study of

real-time language processing, there were no significant differences in the gaze trajectories between listeners with CIs and listeners with acoustic hearing.

Together, these studies confirm that older listeners can use context in a manner similar to younger listeners, although at a slower speed. These studies expand the field's knowledge of the importance of an unclear word's location within a sentence and draw attention to the strategies employed by individual listeners to use context. The results of these experiments provide vital data needed to assess the current usage of context in the aging population with CIs and to develop age-specific auditory rehabilitation efforts for improved communication.

EFFECTS OF AGE ON CONTEXT BENEFIT FOR UNDERSTANDING COCHLEAR-IMPLANT PROCESSED SPEECH

by

Anna Rose Tinnemore

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Sandra Gordon-Salant, Co-Chair
Professor Matthew J. Goupell, Co-Chair
Associate Professor Yi Ting Huang
Dr. Stefanie Kuchinsky
Dean's Representative: Professor Catherine Carr

Acknowledgements

Thank you to my mentors, Sandy Gordon-Salant and Matt Goupell, who have been extremely generous with their time and feedback to help shape the researcher I have become. Thank you to the members of my committee: Stefanie Kuchinsky, Yi Ting Huang, and Catherine Carr, who have always supported me in so many ways. I am indebted to Monita Chatterjee, who supervised my early research attempts and has championed and supported me since we first met. Thank you!

I would like to thank my parents for instilling a love of learning, encouraging my curiosity, and supporting me with hugs and a willingness to listen. I would also like to thank my found family – the friends I've made along the way. To those who supported me through my first research experience, my AuD, and the decision to apply for a PhD: Jason Smith, Laura Gaeta, Sarah Bone Schroyer and David Schroyer. I am so grateful for you. To my fellow students in HESP and NACS at the University of Maryland, including, but not limited to: Rebecca Beiber, Maureen Shader, Brittany Jaekel, Zoe Ovans, my NACS cohort, and Julianne Garbarino. Thank you for your support, even if it was just to listen! To my writing/support group: Ren Salig, Rachel Thompson, Mine Muezzinoglu and others – who else in the world would hear "T-BONE" and think "dinosaur"? You have truly helped keep me sane! To the students I've had the privilege of working with: Becca Higgins Kelly, Erin Doyle, Vanessa Reyes, and many others. I'm so proud of you. Thank you for your help as I developed

and ran these projects. To my parkrun friends, including, but not limited to: Neil and Jules, Diana Gough, Lisa Wilson, Külli Crespin, Lori Dominick, Mike and Bonnie McClellan, Andrea Zukowski and Colin Phillips, and Rebecca and Joe White. My life is richer for having you in it! And finally, to those who friends with whom I formed a COVID pod and without whom I might not have succeeded in this venture: Allison Johnson, Jake Cox, Mary Barrett, and Matthew Brown. I've learned so much from each of you! You mean the world to me. Thank you.

This work was supported, in part, by the National Institute on Deafness and Other Communication Disorders (NIDCD) from an institutional training grant (T32DC000046: Co-PIs: C.E. Carr and M.J. Goupell) and by the Ann G. Wylie Dissertation Fellowship.

Table of Contents

This Table of Contents is automatically generated by MS Word, linked to the Heading formats used within the Chapter text.

List of Tables

List of Figures

List of Acronyms

Acoustic Hearing (AH)

Cochlear Implant (CI)

Decibel (dB)

Hertz (Hz)

Mixed Effects Model (MEM)

Milliseconds (ms)

Neighborhood Activation Model (NAM)

Normal Hearing (NH)

Reverse Hierarchy Theory (RHT)

Root-Mean-Square (RMS)

Signal-to-Noise Ratio (SNR)

Visual World Paradigm (VWP)

Voice Onset Time (VOT)

1 Introduction

Of the rapidly growing population over 65 years old in the United States, at least 16 million people struggle to communicate effectively because of disabling hearing loss (NIDCD, 2016). Older adults with acoustic hearing often perform more poorly than younger adults with acoustic hearing on a wide range of hearing tasks because of peripheral (e.g., age-related hearing loss), central processing (e.g., reduced temporal processing abilities), or cognitive (e.g., slowed processing speed) deficits (e.g., CHABA, 1988; Gordon-Salant et al., 2010; Helfer et al., 2020). On some speech understanding tasks, however, older listeners with acoustic hearing demonstrate no performance deficits compared to younger listeners, occasionally even outperforming the younger listening group. Despite age-related reductions in audibility, supra-threshold processing, and cognition, some older listeners appear to be able to use other strategies to boost their performance to match or exceed that of younger listeners (e.g., O'Neill et al., 2021). One such strategy is a reliance on sentence context to compensate for a loss of access to acoustic information (Boothroyd & Nittrouer, 1988; Nittrouer & Boothroyd, 1990; Winn, 2016). According to the Reverse Hierarchy Theory of Sensory Processing (RHT; Ahissar et al., 2009; Ahissar & Hochstein, 2004; Nahum et al., 2008; Nelken & Ahissar, 2006), speech may be perceived first at a high level, with little to no conscious access to fine-grained acoustic details. Words and phrases are perceived as units. If a listener were asked about a single sound within a word, they would need to concentrate, using effort and cognitive resources to try to remember the small details

of the sounds that made up the word they just heard. Strategies that would save a listener from needing to expend that type of concentration and effort in order to understand speech are very helpful. Knowledge of the language, as defined using vocabulary measures, is roughly equivalent between younger and older adults (Verhaeghen, 2003), while the use of sentence context to predict unclear words can be more extensive in older adults than in younger adults (e.g., Sheldon et al., 2008).

An increasing number of the over 16 million older adults with hearing loss in the United States are electing to receive a cochlear implant (CI) to partially restore some of their ability to understand speech. CIs have been shown to be safe and effective for older adults (e.g., Castiglione et al., 2015; Dillon et al., 2013; Forli et al., 2019; Yang & Cosetti, 2016). In general, the speech understanding performance of older adults with CIs is poorer than that of younger adults with CIs (Blamey et al., 2013; Murr et al., 2021). Speech sounds that have been processed through a CI are highly degraded (e.g., Macherey & Carlyon, 2014). It is likely that older adults with CIs may rely less on the degraded acoustic information received through their CIs than on the information provided by sentence contexts, compared to younger adults with CIs (Amichetti et al., 2018; Moberly et al., 2021; Moberly & Reed, 2019; O'Neill et al., 2021).

This dissertation explores differences in the use of sentence context versus the acoustic information of the target word alone in listeners with CIs across a range of ages using behavioral and eye-tracking methods. An enhanced use of sentence context with age may explain the general success of the prosthesis in older listeners. It may

also explain why younger and older CI users sometimes exhibit no differences in their understanding of target words in sentences (O'Neill et al., 2021) despite the established age-related differences in neural survival and other peripheral and central factors that support speech understanding. The next sections will provide a foundation for understanding how CIs degrade speech, some of the difficulties faced by older listeners when trying to understand degraded speech, the use of context in speech understanding, the age-related changes that may affect a listener's ability to use context, and the interactions between the effects of age and context on speech understanding in listeners who use one or more CIs. The overarching premise of this dissertation is that adults with CIs, especially those who are older, use context to support speech understanding. Given the large amount of variability in speech understanding outcomes in older adults with CIs, the ability to use context successfully may underlie the high levels of speech understanding some listeners with CIs can achieve. The age of the listener may impact the extent to which an individual intuitively relies on sentence context. Ultimately, these findings could lead to the development of age-specific rehabilitation therapies to improve speech understanding in older adult listeners with CIs.

1.1 Cochlear Implants

When hearing loss is severe to profound, an individual may elect to get a CI. These devices are auditory prostheses that are surgically implanted into the inner ear, or cochlea. Worldwide, roughly one million people have been fitted with a CI, ranging from infants who were born deaf to adults who lost some or all of their hearing later in

life (Zeng, 2022). A CI consists of an external sound processor and transmitter, and an internal receiver/stimulator and electrode array. The external components capture and process acoustic sound before transmitting electrical signals through the skin to the internal receiver/stimulator. The internal device sends electrical pulses to the electrodes that have been implanted into the cochlea (as seen in Figure 1-1). These electrical pulses, emitted from the electrodes, stimulate the spiral ganglion cells of the auditory nerve. Auditory information is transmitted to the brain via the auditory nerve.

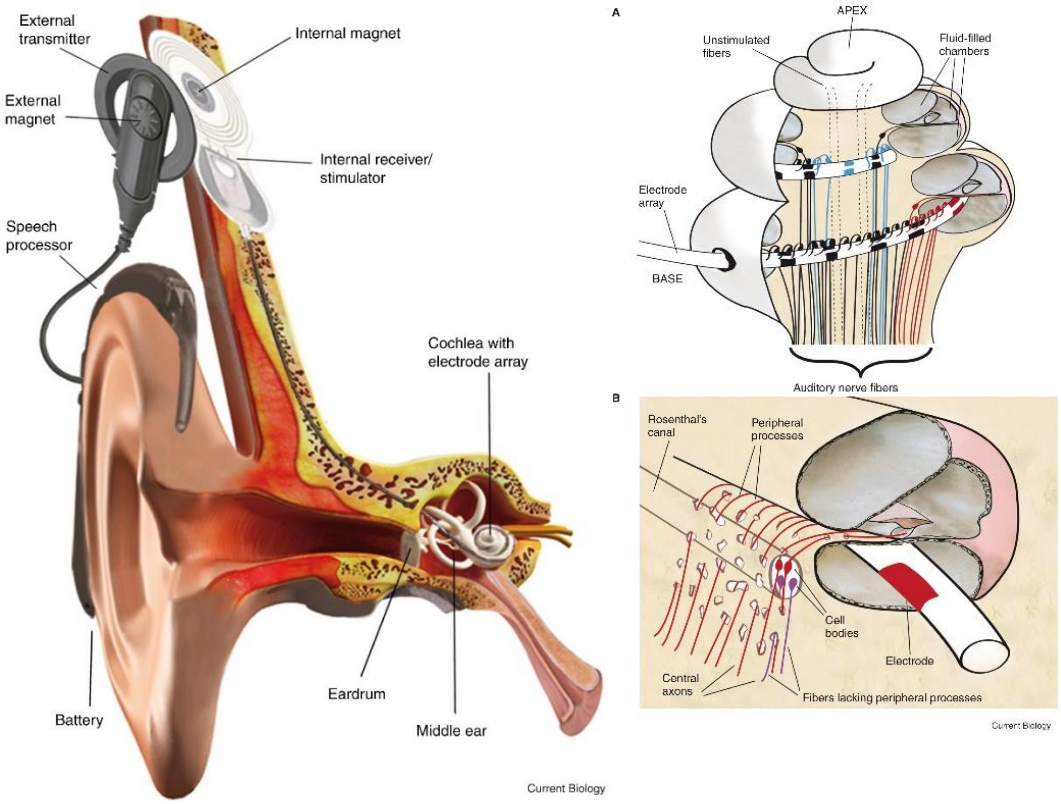

Figure 1-1. Representation of a cochlear implant. The left panel shows the external speech processor and transmitter and the internal receiver/stimulator and electrode array inserted into the cochlea. Panel A shows a cut-away view of the cochlea with the electrode contacts and auditory nerve fibers. Panel B shows a close-up view of an electrode contact within the cochlea and the spiral ganglion cells that are electrically stimulated [from Macherey and Carlyon (2014)].

The transformation of the acoustic waveform into electrical pulses results in the loss of temporal fine structure features while maintaining the temporal envelope of the acoustic waveform (see Figure 1-2, third row). This illustration from Loizou (1999) shows the processing that takes place to convert the acoustic waveform to a series of electrical pulses. The signal is filtered into frequency bands, the temporal envelope of each band is extracted, and electrical pulses are generated that follow the temporal envelope of each band. The resulting signal is a degraded version of the acoustic waveform that the listener's brain interprets as sound (Wouters et al., 2015).

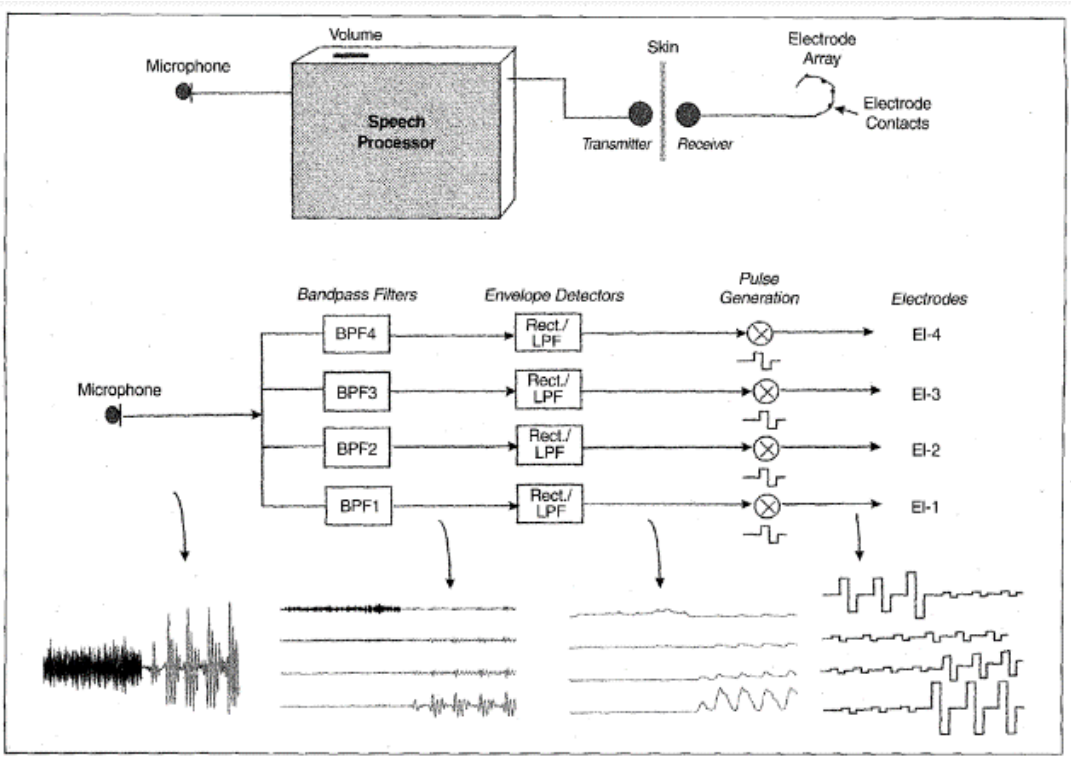

Figure 1-2. Diagram showing a four-channel cochlear implant. The components of a cochlear implant are represented in the top row. The middle row shows the input from the microphone being divided into four bandpass filters (BPFs), the envelopes of those bands being detected with rectified lowpass filters (LPF) before pulses are generated, modified by the envelopes, and sent to the electrodes. The lower row illustrates the transformation of an acoustic waveform into four electrical pulse trains [from Loizou (1999)].

Most CI users have greatly improved speech understanding in quiet listening situations after implantation (e.g., Boisvert et al., 2020; Lenarz et al., 2012; McRackan et al., 2018; Rasmussen et al., 2022). Children who are implanted early in life are able to develop spoken language with great success (e.g., Svirsky et al., 2004; Tye-Murray et al., 1995). Over 80% of adults with hearing loss that occurred after they learned to speak (i.e., postlingual hearing loss) improved their speech recognition ability in quiet by 15 percent or more compared to pre-implantation measurements (Boisvert et al., 2020). Individual outcomes are highly variable (e.g., Blamey et al., 2013). Older CI users (those implanted in middle-age or later) show improved speech understanding, but average performance increases are not as large as for younger CI users, especially when listening to speech in noise (e.g., Dillon et al., 2013; Forli et al., 2019; Haensel et al., 2005; Murr et al., 2021).

1.2 Age-Related Changes to Hearing

It is difficult to disentangle what processes or abilities influence age-related changes to speech recognition. One theory is that there is a common cause (such as the integrity of brain structures and functions) for both sensory and cognitive deficits seen with aging (e.g., Baltes & Lindenberger, 1997; Lin, 2011). The changes seen in cognitive functioning might be driven, for example, by age-related changes in sensory functioning to such an extent that the correlation between age and cognitive processing is quite small when sensory abilities are taken into account (e.g., Humes et al., 2013). Another theory, the processing speed theory of cognitive aging proposed by Salthouse

(1996), describes a general slowing of cognitive processing at all levels, which is then reflected as deficits in cognitive measures and sensory perception.

Processing speed deficits are widely accepted as at least one source of the difference between older and younger listeners' performance on language processing tasks (Wingfield & Tun, 2001). In background noise, in the presence of other people talking, or in other degraded sound conditions, speech understanding becomes more difficult for everyone. Older listeners, however, have disproportionately more difficulty than younger listeners – even with audiometrically normal or near-normal hearing (e.g., Dubno et al., 1984; Helfer & Wilber, 1990).

The Working Group On Speech Understanding and Aging identified three general categories of changes that contribute to age-related declines in speech understanding. Those categories are peripheral hearing ability, central auditory processing, and cognition (CHABA, 1988). While each of these components individually contribute to age-related changes to speech understanding, they also interact with each other.

1.2.1 Peripheral Physiological Changes

Hearing loss often accompanies aging (Cruickshanks et al., 2010; Humes & Dubno, 2010). The inner ear is especially susceptible to age-related anatomical and physiological changes. These changes can include: damaged hair cells and stereocilia, less efficient blood transfer, a reduction in the endo-cochlear potential, and damage to synapses and nerve fibers (e.g., Makary et al., 2011; Schmiedt, 2010; Schuknecht &

Gacek, 1993). Together, these changes result in less frequent auditory nerve firing overall, as well as less synchronous firing. The hearing loss caused by these changes affects the ability to detect quiet sounds (e.g., Kirikae, 1969), as well as performance on suprathreshold tasks such as speech recognition in noise (e.g., Dubno et al., 1984; Helfer & Wilber, 1990; Takahashi & Bacon, 1992). Significant loss of spiral ganglion neurons in the inner ear as a consequence of aging (Makary et al., 2011; Schuknecht & Gacek, 1993; Willott, 1991) can cause deficits in suprathreshold hearing tasks (Kujawa & Liberman, 2009, 2015; e.g., Otte et al., 1978). In general, older people with hearing loss perform even more poorly than might be expected from simple loss of audibility, perhaps because of the loss of spiral ganglion cells. Other factors, such as changes to the central auditory nervous system, may explain some of the age-related hearing deficits that are not explained by peripheral hearing loss.

1.2.2 Central Physiological and Processing Changes

In the central auditory nervous system, changes induced by the biological process of aging can contribute to auditory processing deficits (Canlon et al., 2010). In the cochlear nucleus, there are fewer neurons overall in older mice as well as deficits in inhibitory glycinergic receptors that allow for precise neural synchrony, and significant decreases in the neurotransmitter GABA which is used for inhibition (Canlon et al., 2010; Willott, 1996). In humans, these physiological changes result in a decrease in the precise synchrony of neuronal firing throughout the midbrain (Anderson et al., 2012; Parthasarathy, Herrmann, et al., 2019; Presacco et al., 2016) and into the

auditory cortex (Goossens et al., 2016; Parthasarathy, Bartlett, et al., 2019; Tremblay et al., 2002) as measured with electrophysiology. One of the behavioral impacts of these age-related physiological changes is an auditory temporal processing deficit (Fitzgibbons & Gordon-Salant, 1996; Gordon-Salant et al., 2008; Gordon-Salant et al., 2006; Gordon-Salant & Fitzgibbons, 1993, 1999) that can affect speech perception in many ways. One specific area of speech perception that is profoundly affected by temporal processing abilities is that of distinguishing certain speech sounds. Several distinguishing cues between speech sounds, or phonemes, are temporal in nature. For example, voice onset time (VOT) cues differentiate voiced from voiceless initial plosives and vowel duration cues differentiate voiced from voiceless final plosives. Deficits in auditory temporal processing can result in a reduced ability to distinguish contrasting phonemes. This can result in words being misheard, leading to communication difficulties.

1.2.3 Age-Related Changes to Cognition

As an individual ages, they may experience declines in many cognitive abilities (e.g., Baltes & Lindenberger, 1997; Humes et al., 2013; Salthouse, 1996; Wingfield & Tun, 2001). Deficits can occur in specific cognitive domains that affect speech understanding such as processing speed (Rush et al., 2006; Salthouse, 1996; Wingfield, 1996), working memory (Hasher & Zacks, 1988; Rönnberg et al., 2010; Rudner et al., 2011; Stine & Wingfield, 1987), and inhibition (Hasher et al., 1991; Sommers & Danielson, 1999; Stenbäck et al., 2021). Processing speed refers to the time needed for

a stimulus to be detected and a decision or response to be generated (Salthouse, 2000). Working memory is commonly defined as the ability to manipulate information in short-term memory storage (e.g., Stine & Wingfield, 1987). This mental manipulation could be sorting or otherwise transforming the original information to be remembered. Working memory is the cognitive ability that has been shown to correlate the most with degraded speech understanding (Just & Carpenter, 1992; Rönnberg et al., 2010; Rudner et al., 2011; Schubotz et al., 2021). Individuals with poor working memory are likely to perform poorly when asked to understand degraded speech, even if they have audiometrically normal hearing thresholds.

Another area of cognition that changes with aging is the ability to inhibit irrelevant information (Hasher et al., 1991). There is evidence of a decline in the neurotransmitters and receptors necessary for neuronal inhibition as mentioned in the previous section. In addition, the cortical representation of speech is exaggerated in older adults compared to younger adults, and this enhanced cortical response seems to be negatively correlated with cognitive tests of attention and inhibition (Presacco et al., 2016, 2019). The impact of decreased neural inhibition is that older adults have more difficulty ignoring irrelevant information in auditory and visual domains than younger adults (e.g., Alain & Woods, 1999; Guerreiro et al., 2013; Mattys & Scharenborg, 2014). In models of language processing, such as the TRACE model (McClelland & Elman, 1986) or the Neighborhood Activation Model (Luce & Pisoni, 1998), decreased inhibition in older adults may also result in the activation of more potential lexical matches to the incoming sound than may be activated in younger adults – making the

ultimate word identification choice slower and perhaps more difficult (e.g., Rush et al., 2006; Sommers & Danielson, 1999; Stenbäck et al., 2021).

1.2.4 Summary of age-related changes to hearing

As the Working Group On Speech Understanding and Aging noted, each of these three areas (peripheral hearing, central auditory processing, and cognition) can contribute to age-related declines in speech understanding (CHABA, 1988). The reality is that changes in each of these areas interact within an individual to create a highly personalized hearing experience that may change as an individual grows older. Understanding the effects of age-related changes in central auditory processing and cognition in listeners who use CIs may help explain the highly variable performances of those implanted as adults. The effects of age on ideal device settings, speech perception, and communication success remain fertile ground for future research.

1.3 Context

In contrast to the common declines in sensory and cognitive abilities associated with increasing age, many of the abilities useful for benefiting from context to aid speech understanding may be well-preserved in older listeners. These include knowledge of the language, vocabulary, and world knowledge (Burke & Peters, 1986; Verhaeghen, 2003). Several studies have shown that older listeners use context more than younger listeners with normal hearing or may receive more benefit from highly predictable contexts than younger listeners (e.g., Pichora-Fuller et al., 1995; Sommers

& Danielson, 1999; Speranza et al., 2000). Other studies have not found the same enhanced benefit of context for older listeners (e.g., Dagerman et al., 2006; Dubno et al., 2000). The use of context might offset the effects of slowed processing of sensory information or declines in working memory, and allow older listeners to perform at the same level of accuracy as younger listeners on certain speech recognition tasks.

Context has many definitions across many research disciplines (for reviews, see: Abowd et al., 1999; Bazire & Brézillon, 2005; Becker, 1980). For this dissertation, context is defined as the words and phrases surrounding the target word in a sentence. Context can exist in the words preceding or following a target word. While in the real world, context does not necessarily stop at sentence boundaries, the proposed studies will present sentences in isolation (i.e., not within a discourse). The effects of context will necessarily be limited to the single sentence containing the target word. The effects of context are typically measured by comparing speech recognition scores or response times for target words presented in isolation versus in the presence of a context sentence. In addition, context benefit is typically assessed for the final word in a sentence (e.g., Amichetti et al., 2018; Connine, 1987; Kalikow et al., 1977; Patro & Mendel, 2016; Wagner et al., 2016; Winn, 2016). Contexts can be predictive, neutral, or incongruent. A predictive, or congruent, context exists when the target word can be predicted logically and probabilistically, given the surrounding words or sounds. A neutral context exists when the surrounding words or sounds provide no clues to the identity of the target word, although the sentence is syntactically and semantically

correct. An incongruent context exists when the surrounding words or sounds cause the listener to predict a word or sound other than the target. See Table 1-I for examples.

Table 1-I. Example context sentences for the target word "doctor."

Context Type	Sentence
Predictive	He took the sick child to see the **doctor**.
Neutral	I heard about the **doctor**.
Incongruent	The fierce storm blew a branch off the **doctor**.

Context provides the listener with cues to effectively use the rules of language and shared world knowledge to fill in a word they did not fully hear or to predict a future word (e.g., Moss & Marslen-Wilson, 1993; Wingfield et al., 1994). When a word is not heard clearly, whether it was masked by competing noise or simply not said, a listener initially has no acoustic cues to help identify the word. They must rely instead on "top-down processing" to fill in the missing word and extract the talker's intended meaning. Top-down processing is defined as the use of cognitive skills, resources, and effort to facilitate the decryption of sensory input (Marslen-Wilson & Tyler, 1980; Morton, 1969). Reverse Hierarchical Theory (RHT) would postulate that top-down processing includes the use of effort and cognitive resources to access the limited low-level acoustic cues that were initially processed subconsciously (Nahum et al., 2008; Nelken & Ahissar, 2006). For example, if one heard the incomplete sentence, "Leaves turn beautiful colors in the…", the rules of the English language indicate that the next word must be a noun and it must be a location, object, or time. World knowledge

reveals that leaves change color as daily temperatures decrease. Collectively, this information will likely lead to the prediction of "fall" as the final word.

Obtaining a benefit from context is usually accomplished by combining "bottom-up information" from peripheral inputs with top-down processing. Bottom-up information is defined as basic sensory information from peripheral sensory organs. As the neural representation of the signal ascends toward the cortex, minor transformations and processing occur through the various central auditory nervous system nuclei. Top-down processing, as defined previously, involves cognitive skills, resources, and effort to facilitate the decryption of sensory input (Marslen-Wilson & Tyler, 1980; Morton, 1969). Top-down processing, logically necessary for context benefit, is most useful when the sensory input is degraded (Sohoglu et al., 2012), such as when speech is heard through a CI processor (Loebach et al., 2010; O'Neill et al., 2021; Patro & Mendel, 2020) or in the presence of background noise (Bronkhorst et al., 2002; Dubno et al., 2000) or both (Jaekel et al., 2021). Context cues trigger top-down processes that allow the listener to fill in missing bottom-up information for less effortful understanding of the intended message (Sohoglu et al., 2012). Knowledge of the topic of conversation narrows the range of possible intended messages and provides considerable context benefit (McClelland & Elman, 1986). The addition of a sentence context can improve the threshold for recognizing a word in background noise by 6 dB signal-to-noise ratio (SNR) in young normal-hearing listeners (Miller et al., 1951).

Most research into the benefit of context is focused on the ability to predict a single word at the end of a sentence (e.g., Wingfield, 1996). The use of a carrier phrase

that strongly predicts a target word has been shown to facilitate recognition of a final word better than phrases that are syntactically compatible but not semantically predictive of that final word, or phrases that are merely incoherent chunks of speech (e.g., Kalikow et al., 1977).

Very few studies have examined the benefit of subsequent context for predicting a target word. One such study found that subsequent context did affect the identification of a sentence-initial word, but only when the predictive context occurred within three syllables after the target word (Connine et al., 1991). Another study provided context for a sentence-final word and then found that they could disrupt the benefit of that context for listeners with CIs by playing another, disruptive sound after the end of the target word (Winn & Moore, 2018). Yet another study showed that while older listeners with AH benefitted as much as younger listeners with AH when the context was presented before the target word, the older listeners did not benefit as much from context presented after the target word as younger listeners (Wingfield et al., 1994). These studies show that subsequent words or noises can affect previously heard target words. However, the use of subsequent context and the magnitude of the context effect by different listener groups are not yet known. It remains to be seen, therefore, how listening to a degraded signal, such as through a CI, might affect listeners' ability to use subsequent context and what strategies might be most effective for older listeners to take advantage of subsequent context cues – whether they use a CI or not.

1.3.1 Factors that Affect the Use of Context by Older Adults

Context can be used by older adults with and without hearing loss to predict degraded words and thereby improve speech understanding (Federmeier & Kutas, 2005; Nittrouer & Boothroyd, 1990) and reduce listening effort (Goy et al., 2013; Hunter & Humes, 2022). As noted above, some studies have shown that older adults receive a greater benefit from context compared to younger adults (e.g., Goy et al., 2013; Madden, 1988; Sheldon et al., 2008; Sommers & Danielson, 1999), while others have shown an equal benefit of context between older and younger adults (e.g., O'Neill et al., 2021; Stine & Wingfield, 1987; Stine-Morrow et al., 1999).

One factor affecting the benefit of context by older adults is signal degradation. Background noise and reduced spectral content are two methods by which a signal can be degraded. The amount of signal degradation can affect the potential benefits of context, with the greatest context benefit at moderate levels of degradation and less benefit when there is either no degradation or high levels of degradation (Bhandari et al., 2021; Dubno et al., 2000; Goy et al., 2013; Moberly et al., 2021; Obleser & Kotz, 2010; Sohoglu et al., 2012).

A second factor affecting the benefit of context is the listener's ability to use cognitive skills, such as working memory and processing speed, to augment their sensory perception. Specific cognitive abilities have been shown to be predictive of speech-in-noise performance in listeners with AH – both those with typical hearing and those with hearing impairment (Akeroyd, 2008; Dryden et al., 2017). In addition, adults

using CIs have been shown to use top-down processes, such as cognitive abilities and context, to improve speech perception (Gianakas & Winn, 2019; Moberly & Reed, 2019; Patro & Mendel, 2018, 2020). In one recent study, knowledge of vocabulary and working memory — cognitive abilities generally thought to be useful for top-down processing — did not correlate with speech understanding performance in listeners with CIs (Bosen et al., 2021). The specific measures tested in Bosen et al. (2021) were Forward Digit Span, a working memory task involving mostly storage of information, and word familiarity measures of vocabulary. Measures of working memory that require manipulation of information, such as the Reading Span or Listening Span, on the other hand, seem to be highly predictive of speech understanding performance in adult listeners with CIs (e.g., Kaandorp et al., 2017), as well as adult listeners with AH (e.g., Gordon-Salant & Cole, 2016). Thus, performance on some tests of working memory or vocabulary may be more sensitive than others for identifying the cognitive domains and abilities that older listeners with CIs use to support speech understanding.

1.3.2 Measuring the Use of Context

Throughout the years, several methods have been employed to determine how context might affect an individual's perception. One of the simplest methods measures the difference in word recognition accuracy for words presented with and without context (e.g., Amichetti et al., 2018; Sheldon et al., 2008; Stine-Morrow et al., 1999; Wingfield et al., 1994). A similar method measures phoneme categorization for ambiguous words presented in one context or another (e.g., Borsky et al., 1998;

Connine et al., 1991; Mattys & Scharenborg, 2014; Nittrouer & Boothroyd, 1990). A third method compares response times to stimuli presented with and without context (e.g., Connine, 1987; Goy et al., 2013). Yet another method involves measuring evoked potentials to stimuli presented in congruent compared to incongruent contexts (e.g., Federmeier & Kutas, 2005; Hannemann et al., 2007; Hunter, 2020; Signoret et al., 2020). A final method is analysis of listeners' eye movements as they are looking at pictures representing words that might be the target word in the sentence they are currently hearing, a method known as the Visual World Paradigm (VWP; for a detailed review, see Huettig et al., 2011). The VWP allows for insight into the real-time processing of speech as it unfolds throughout the duration of the word or sentence. Behavioral differences in categorization response times have been measured to describe the benefit of context (e.g., Connine, 1987), but this is an indirect measure of the effect of context on the categorical decision-making process. The VWP provides a more direct measurement of the decision-making process of real-time speech understanding in degraded listening conditions by tracking the time a listener's eyes remain fixed on the pictures representing candidates for the target word. This measurement could provide insight into the mechanisms and strategies used by listeners of different ages and different hearing abilities.

Some researchers propose that uncertainty about the first phoneme in a word may delay a person's recognition of that word. Listeners with CIs, knowing that they might mishear an onset phoneme, might purposefully delay committing to a lexical choice to make it easier to revise if needed (Farris-Trimble et al., 2014). In a similar

study, McMurray et al. (2017) tested pre-lingually deafened listeners with CIs and age-matched peers listening to severely degraded speech. They found that both groups significantly delayed their decision about the target word to allow for more auditory information to be accumulated. They proposed that individuals listening to degraded speech might employ a "wait-and-see" strategy compared to those listening to clear speech. Indeed, further work seems to support this theory (McMurray et al., 2019; F. Smith & McMurray, 2018). It is worthwhile to note that this research has thus far mostly been limited to individual words. Recent studies have begun to evaluate the impact of listening to a target word in a context sentence (F. X. Smith & McMurray, 2022).

The "wait-and-see" strategy proposed by McMurray and colleagues has implications for the real-time processing of running speech for older listeners with CIs and the methods we use to measure speech processing. One method used to measure top-down processing and the effect of context requires listeners to understand speech that is interrupted at regular intervals with bursts of noise. There is currently no evidence that adult listeners with CIs can restore interrupted speech using top-down processing (Jaekel, 2020; Jaekel et al., 2018, 2021). When context effects are tested by examining the impact of context on a sentence-final target word, older listeners with CIs may obtain a greater benefit from context than younger listeners with CIs (Amichetti et al., 2018). This finding is consistent with age-related context benefits that have been reported in listeners with AH (Sommers & Danielson, 1999).

This dissertation presents data to address the questions of whether: (1) age-related context benefits will be affected by the sentence position of the target word; (2) age-related context benefits will be predicted by cognitive measures of processing speed or inhibition; (3) older listeners with CIs will show similar context benefit as older listeners with AH while listening to degraded speech; (4) the "wait-and-see" strategy will be used in sentences containing predictive context; and (5) the use of the "wait-and-see" strategy will be dependent on the listener's age, the listener's cognitive abilities, or the signal degradation (background noise, CI processing, etc.). The answers to these questions will provide a foundation for developing age-specific therapeutic strategies to maximize the benefits of context for adult listeners with CIs.

1.4 Specific Aims

This dissertation research is at the intersection of speech perception, age-related changes in hearing and cognitive abilities, and CI use. The central hypothesis is that increased age will be associated with greater change in the identification of a word in a sentence because of the semantic context conveyed by the sentence. The central hypothesis will be tested in the following Specific Aims.

Aim 1: Quantify the extent to which increasing age predicts the context effect for a listener with AH.

Older adults employ listening strategies that are shaped by their knowledge of the language, their ability to process language (e.g., using cognitive abilities like working memory), and the state of their auditory periphery and central auditory nervous

system (Pichora-Fuller, 2008). It has been shown that older adults with AH use context more than younger adults (Sheldon et al., 2008). Given these findings, it is hypothesized that older listeners with AH will show greater sentence context effect sizes than younger listeners with AH. Context effects will be measured by comparing the behavioral categorization of target words from an acoustic continuum embedded in sentences that provide no context cues (acoustic cues only) to categorization of the same target words in sentences that provide predictive context to support the identification of the target word (context and acoustic cues). Previous studies have found that subsequent context does not affect performance as much as preceding context, especially for older listeners (Connine et al., 1991; Wingfield et al., 1994). It is therefore hypothesized that there will be an interaction between age and location of the target word. Older listeners are hypothesized to demonstrate a similar context effect compared to younger listeners when the target word is at the end of the sentence, but less of a context effect compared to younger listeners when the target word is at the beginning of the sentence.

In addition, cognitive measures of working memory and inhibition are hypothesized to correlate with less reliance on context and more accurate categorization of target words, especially for older listeners. When an unclear word is at the beginning of a sentence, a listener with better working memory ability may retain more acoustic cues to the identity of that word than a listener with poorer working memory. A listener with a strong memory of the target word or an ability to suppress irrelevant information would likely be less swayed by a context sentence that was incongruent with the target

word than a listener with little to no memory of the target word or who was unable to suppress irrelevant information.

Aim 2: Quantify the extent to which increasing age predicts the context effect for a listener with a CI.

Older adults with CIs use context to an equal or greater extent than younger adults with CIs (Amichetti et al., 2018). Older adults with CIs, like older adults with AH, likely maintain their knowledge of the language and experience age-related declines in the cognitive abilities that are used to process language (e.g., working memory) (Pichora-Fuller, 2008). Similarly, age-related declines are as likely to occur in the central auditory nervous systems of older adults with CIs as in those of older adults with AH (e.g., Canlon et al., 2010; Gordon-Salant et al., 2020; Shader, Gordon-Salant, et al., 2020). Given these findings, it is hypothesized that older listeners with CIs will show greater context effects than younger listeners with CIs. Context effects will be measured by comparing the behavioral categorization of target words from an acoustic continuum at one end or the other of sentences that provide no context cues (acoustic cues only) to categorization of the same target words in sentences that provide predictive context to support the identification of the target word (context and acoustic cues). With no previous studies examining subsequent context effects for listeners with CIs, the hypotheses about the effects of target location must be based on the findings in listeners with AH. Thus, it is hypothesized that there will be an interaction between age and location of the target word. Older listeners with CIs are hypothesized to perform similarly to younger listeners with CIs when the target word is at the end of

the sentence but will demonstrate smaller effects of context than younger listeners when the target word is at the beginning of the sentence. As with listeners with AH, when an unclear word is at the beginning of a sentence, a listener with better working memory or inhibition ability may retain more acoustic cues to the identity of that word and be less swayed by the context of the sentence than a listener with poorer working memory or inhibition ability. Cognitive measures of working memory and inhibition are hypothesized to correlate with less reliance on context and more accurate categorization of target words in both sentence locations, especially for older listeners.

Aim 3: Determine the extent to which real-time speech processing differs as a function of age and signal degradation.

Older listeners are expected to have more difficulty overriding contextual predictions that conflict with degraded acoustic information (Harel-Arbeli et al., 2021), and listeners with CIs have been shown to use a "wait-and-see" strategy during sentence comprehension (McMurray et al., 2017). Given these findings, it is hypothesized that older listeners with AH and all listeners who use CIs will differ in their real-time speech processing in the presence of predictive context sentences compared to younger listeners with AH. This difference in processing may take the form of delays in looking exclusively at the target words that are ultimately chosen as the response. The hypothesized differences in speech processing may be the result of differences in cognitive abilities (e.g., age-related working memory differences) or differences in strategies (e.g., wait-and-see). Real-time language processing will be

measured objectively in adult listeners who use CIs or who have AH and are presented vocoded speech, by tracking their eye movements while they are listening to sentences.

The expected outcomes are that older listeners with AH and listeners with CIs (both younger and older) will process sentences differently than younger listeners with AH, and that an increased reliance on sentence context is beneficial for older listeners (relative to younger listeners) to achieve correct speech recognition when the sentence context is congruent with the target word. An increased weight on context may compensate for slower real-time speech processing speeds for older listeners when disambiguating highly degraded words in congruent contexts compared to the faster processing of younger listeners. A reliance on sentence contexts when the context is incongruent would increase the discrepancy between categorization in the predictive sentence context compared to the neutral sentence context. Listeners with CIs may be less confident in their choice of the target word than peers with AH and may demonstrate a wait-and-see approach. Cognitive measures of processing speed may be more predictive of gaze patterns and response times on this task, while measures of working memory and inhibition may be more predictive of overall accuracy and context effects. The results of this experiment are expected to reveal performance strategies that may vary according to listener age, signal degradation, cognitive abilities, or interactions between these factors.

1.5 General Methodological Approaches

1.5.1 Phonemic Categorization

Phonemic categorization is a method in which stimuli are created that straddle the acoustic space between two phonemes. In a simple example, the phonemes /t/ and /d/ are both alveolar plosives. The difference between these phonemes is voicing (the vibration of the vocal cords). Acoustically, this difference is realized as a relative difference in the onset of voicing between the initial burst and the subsequent vowel following the plosive, referred to as voice-onset time (VOT). In the case of /t/, no vocal cord vibrations are used to produce this sound and there is a period of aspiration before the voicing from the subsequent vowel begins, resulting in a VOT of approximately 86 ms. In the case of /d/, the vocal cords begin vibrating following a relatively short delay after the plosive, blending smoothly into the voicing of the subsequent vowel and resulting in a VOT of approximately 26 ms. This difference in the relative onset of voicing, or VOT, is the main acoustic cue that distinguishes these two phonemes (Lisker & Abramson, 1964).

A stimulus continuum is created with two endpoints that ideally differ along a single acoustic parameter. Intermediate tokens that vary in a series of equal step sizes along the acoustic parameter of interest make up the rest of the continuum. Ideally, the endpoints of the continuum are always perceived as different words. A listener is asked to categorize the target words from each step along the continuum as either one endpoint word or the other. At a certain point along the continuum, a listener will switch

from categorizing the target words of the intermediate steps as one endpoint word to the opposite endpoint word. Several trials of each step along the continuum are presented in a method of constant stimuli to create a psychometric function of an individual listener's categorical boundary between the two phonemes.

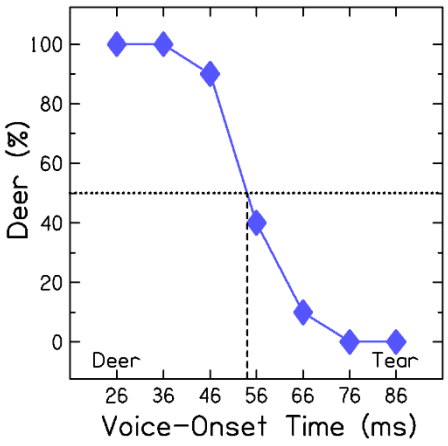

Figure 1-3. An example psychometric function showing performance on a phonemic categorization task.

An example of a psychometric function is shown in Figure 1-3. This example shows performance on a phonemic categorization task on a seven-step continuum between the /d/ in the word "deer" and the /t/ in the word "tear." The x-axis shows the VOT of each stimulus, corresponding to each step along the continuum of the acoustic cue. The y-axis shows the percentage of responses that the individual identified the word as "deer." The horizontal line at 50% shows the crossover point where the listener's categorization shifted from mostly "deer" responses to mostly "tear" responses. The crossover point corresponds to roughly 54 ms of voice-onset time as indicated by the dashed line from the intercept between the psychometric function and the 50% line to the x axis. This value (crossover point) has been shown to change with age, demonstrating age-related temporal processing deficits in listeners with acoustic hearing (Gordon-Salant et al., 2008; Gordon-Salant et al., 2006; Goupell et al., 2017) and in listeners with CIs (Xie et al., 2019).

The phoneme contrast chosen for the experiments in this dissertation, VOT, is mostly temporal in nature. An analysis of how far along the continuum an individual's categorization switches from one word to the other, the crossover point, can provide insight into the temporal processing abilities of that individual. Another feature of interest from the psychometric function is the slope of the function. A steeper slope implies a clear boundary between the two perceived endpoint words. A shallower slope implies more ambiguity and uncertainty about the phonemic boundary.

1.5.2 Eye Tracking

Eye tracking is a method used to provide insight into processing functions that may not be conscious. It has been used in language research to explore lexical processing as speech unfolds over time. Over the years, it has provided support for the Neighborhood Activation Model (NAM) of word recognition (Luce & Pisoni, 1998). The NAM states that as a word unfolds over time, the acoustic-phonetic properties of the word cause the activation of potential lexical matches to the incoming stimulus in the brain (all the words that differ from the target word by a single phoneme). Word recognition is dependent on a decision-making process that combines acoustic-phonetic information with higher-level lexical information (e.g., word frequencies, semantic relationships, and recent memory items that might enhance word recognition) to decide which of the activated neighborhood of words is the target word. The word-recognition process starts as soon as the initial acoustic information reaches the brain. Evidence for this theory comes from eye tracking. When listeners are looking at

pictures depicting possible word choices, listeners' eyes reflexively look at the picture representing the word they think they are hearing. This reflexive movement of a listener's eyes, as options are considered, can be measured in young children as well as adults using the visual world paradigm, or VWP (Allopenna et al., 1998; Cooper, 1974; Tanenhaus et al., 1995). The VWP will be used in this dissertation to assess the real-time processing of degraded words in congruent and incongruent sentence contexts. In the VWP, there are four pictures on the screen representing possible interpretations of the acoustic stimulus. The experimenter measures the time it takes for a listener's gaze to settle on a single picture representing the word the listener heard. The eye-tracking data can also be presented as a ratio of time spent looking at the target word compared to the distractor words as the auditory stimulus plays. The four pictures visible on the screen from which a listener can choose typically include the target word and three distractors that may share some of the properties of the target word. Listeners reflexively, and often subconsciously, look at the representation of the word they think they are hearing as it unfolds over time. An analysis of these reflexive eye movements can provide insight into real-time word and sentence processing without relying on subjective reports.

1.6 Overview of Chapters

Chapter 2 describes Study 1, which examined how the position of the target word in the sentence influenced the context effect as a function of age in participants with AH. The experiment presented target words from an acoustic continuum in

sentences that were predictive of one word choice (e.g., “deer”), the other word choice (e.g., the rhyming word, “tear”), or not predictive of one identification response over the other. Both context sentences and target words were spectrally degraded as a simulation of the signal utilized by CI listeners. Additionally, the target words were presented in background noise at a range of SNRs to further limit a participant’s access to acoustic information. Context reliance was measured behaviorally using categorization of ambiguous and unambiguous target words in context sentences.

Chapter 3 describes Study 2, which used the same experimental methodology as Study 1 with participants who use CIs. As in the first experiment, Study 2 evaluated how age impacted the use of context in spectrally degraded speech. Context reliance was measured behaviorally using categorization of ambiguous target words at the beginning or end of sentences with or without predictive context.

Chapter 4 describes Study 3, an experiment that used eye-tracking methodology to infer information about the real-time processing of degraded speech in younger and older participants with AH or CIs. Participants performed a limited phonemic categorization task on ambiguous target words presented at the beginning of sentences with congruent or incongruent context while a camera tracked the location of their eye gaze.

Chapter 5 contains a general discussion of the findings presented in the previous chapters and how those findings contribute to and complement the existing literature. In addition, the limitations of the current studies and ideas for future directions are discussed.

2 Study 1: Sentence position and context effects on voice onset time contrasts as a function of age in listeners with normal hearing

2.1 Introduction

Successful communication involves decoding sensory signals that carry intended messages between people. This decryption of sensory signals is accomplished by combining bottom-up information (e.g., peripheral inputs and central auditory processing) with top-down information (e.g., cognitive skills and effort) (Marslen-Wilson & Tyler, 1980; Morton, 1969). Top-down processing is most useful, and the benefit of context is greatest, when the sensory input is moderately degraded (Bhandari et al., 2021; Obleser et al., 2007; Obleser & Kotz, 2011; Sohoglu et al., 2012; Wild et al., 2012). This study presents target words at various SNRs in speech-shaped noise to cover a range of speech degradation conditions from mildly to highly degraded speech signals. Context cues trigger top-down processes that allow the brain to fill in missing bottom-up information for less effortful understanding of the intended message (Sohoglu et al., 2012). Knowledge of vocabulary, grammar, and the rules governing semantic roles narrows the range of possible intended messages and provides considerable context benefit (McClelland & Elman, 1986). In addition, some cognitive abilities have been shown to correlate positively with degraded speech understanding. Specifically, better working memory abilities have been shown to correlate with better

speech understanding in noise for older listeners with hearing loss (e.g., Kemper, 1992; Rönnberg et al., 2010; Rudner et al., 2011; Schubotz et al., 2021). In addition, older listeners with slower processing speeds and reduced inhibition abilities generally perform more poorly than older adults with better cognitive skills on tasks with degraded speech stimuli (Pichora-Fuller, 2003; Presacco et al., 2019; Sommers & Danielson, 1999; Stenbäck et al., 2021).

Older adults typically have extensive vocabularies and vast experiences with language that stay intact with age (Ben-David et al., 2015; Verhaeghen, 2003). It is possible that this knowledge of the language can be used to overcome sensory deficits – simulated in the current study with the use of vocoded speech and various levels of background noise – through the top-down use of context (Amichetti et al., 2018; Sommers & Danielson, 1999). Context usage may be sufficient to offset the declines in both sensory and cognitive abilities that often occur with advanced age.

2.1.1 Phoneme Categorization

Early work evaluating the effects of top-down influences on phoneme categorization was performed by Ganong (1980). Ganong found that when a VOT continuum was constructed such that one endpoint was a non-word and the other endpoint was a word (e.g., giss-kiss), listeners more often categorized an ambiguous word between the two extremes as belonging to the word category. Ganong also found that response times for the trials where the listener judged the ambiguous words as words were faster than when the listener judged the ambiguous words to be non-words.

These results have been interpreted as a lexical bias, in other words, listeners prefer words over non-words. This lexical bias has been shown in a graded manner in young listeners with normal hearing (NH) presented with spectrally degraded speech (Gianakas & Winn, 2019). Lexical bias was greater for more degraded signals than for less degraded and unprocessed speech.

Sentence context can affect the categorization of a sentence-final word. Connine (1987) showed that the 50% crossover point between categories was shifted along a phonetic continuum when the target words occurred after two different context sentences that each predicted a different phoneme. A context effect was also seen in the reaction times to the unambiguous contrasting words. Response times were significantly faster for unambiguous target words when the sentence context was congruent with the target word than when the context was incongruent. This was interpreted as tentative evidence toward Fodor's (1983) modular theory of mind and a post-perceptual role for semantic context. Fodor's modular theory of mind states that processes within the brain occur in domain-specific self-contained localized "modules" that do not have access to information from other modules while the inputs to the module are being processed. The post-perceptual role of semantic context, therefore, would mean that sounds are first perceived (in one domain-specific module), and then semantic context (a separate module) can alter the final decision about what word was perceived.

2.1.2 The effect of sentence position

Context can affect the categorization of a word, regardless of the relative location of target word and context in the sentence (e.g., Connine et al., 1991; Wingfield et al., 1994; Winn & Moore, 2018; Wotton et al., 2011). For target words on a phonetic continuum with subsequent context, the effects are greatest if the semantic predictive information in the context sentence occurs within three syllables after the target word. The effects of context for young listeners with normal hearing are almost completely eliminated if the semantic predictive information in the context sentence occurs six syllables or more after the target word (Connine et al., 1991). This implies that there is a window during which subsequent information in the context sentence can influence the categorization of a sentence-initial ambiguous target word. In another study, Wingfield et al. (1994) found that older listeners with AH showed less context benefit from subsequent context than younger listeners with AH in a word recognition task. These authors also showed that the benefits of preceding and subsequent contexts plateaued approximately two words away from the target word. To date, there are no studies showing a windowing effect of a context benefit for sentence-final words.

2.1.3 Signal degradation, aging, and context effects in listeners with AH

Signal degradation affects phoneme categorization. Listeners with AH who are presented vocoded speech show poorer phoneme categorization with decreasing levels of spectral resolution (Winn & Litovsky, 2015). Signal degradation also interacts with

age to disproportionately slow the response times of older adults compared to younger adults (Goy et al., 2013). However, it was recently shown that if listener groups were equated on speech understanding performance of sentences with no contextual information, the differences in the benefit of context between age groups were no longer significant (O'Neill et al., 2021). This confirms previous studies that showed that when speech was equally audible for younger and older listeners with AH, age-related differences in the benefit of context were minimal (Dubno et al., 2000; Humes & Dubno, 2010).

2.1.4 Experimental Questions

This study tested younger and older participants with AH on a phoneme categorization task where the spectrally degraded target word was presented at the beginning or end of spectrally degraded context sentences. The context sentences were presented in quiet, and the target words were presented in background noise. 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? and (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 smaller effects for subsequent context reported by Connine (1991) and Wingfield et al. (1994) compared to preceding context, 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. If the brain's prediction overrides the degraded acoustics of the target word in predictive context sentences, the same target word could be identified differently when presented in a predictive context sentence compared to when presented in a neutral context sentence. The identification of a noise-masked target word at the beginning of the sentence may be less affected by subsequent context because no internal prediction had been made prior to hearing the target word.

An additional research sub-question (Question 1a) is whether an individual's working memory ability is related to the extent to which the individual's responses are based on contextual information, and whether or not this relationship varies with the location of the target word in the sentence. The hypothesis was that individuals with poorer working memory ability might show larger context effects when the target word was at the beginning of the sentence more than when it was at the end of the sentence compared to individuals with better working memory abilities. An individual with better working memory abilities might be able to hold their initial interpretation of a target word located at the beginning of a sentence separate from the influence of the subsequent context sentence. Working memory ability was not hypothesized to have a

significant effect on the identification of target words at the end of the sentence because of the relatively simple and short sentences used.

The hypothesis for the second research question was that older participants' identification of target words would show similar patterns of context effects as younger participants, but the response times of older participants would show different patterns of context effects than those of younger participants. It was hypothesized that older participants would exhibit longer response times to target words occurring in incongruent sentence contexts than younger participants. This difference in response times 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 older participants would show smaller differences in the context effects between the two sentence positions than the younger participants. Wingfield et al. (1994) found smaller effects of subsequent context for older listeners with AH compared to younger listeners with AH. Thus, similar effects of sentence context on categorization of target words are expected between age groups when the target word is at the end of the sentence, and smaller effects of sentence context are expected for older listeners than younger listeners when the target word is at the beginning of the sentence.

2.2 Method

2.2.1 Participants

Thirty-six (36) adult participants (ages 20 - 76 years old, mean = 48.7 years) with nominally normal acoustic hearing were recruited. All participants had no more than a slight hearing loss (≤ 25 dB HL) at octave frequencies between 250 - 4000 Hz in at least one ear (re: ANSI, 2018). Participants were self-reported native speakers of American English to avoid variability on the speech identification tasks stemming from diverse language backgrounds. Participants had to be able to read moderately large print on a computer screen with or without corrective lenses. This ability was determined during the practice trials.

2.2.2 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 (f0=125 Hz). Recordings took place in a double-walled sound-treated booth. The talker was seated so that his mouth was positioned 8 inches away from the wind guard covering a condenser microphone set in omni-directional mode with zero gain and a flat frequency response (Shure KSM141, Niles, IL, USA). The microphone was connected to an audio recorder (Zoom H4n Pro, Hauppauge, NY, USA). The recording

sampling rate was 44,100 Hz with 16-bit depth. After a recording session, the recordings were transferred to a computer for excision and analysis.

The context sentences were read at least three times with each target endpoint word. The context sentence chosen for each condition was selected 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 VOT from 26 to 86 ms for a word-initial /d/ to a /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 (see Figure 2-1). The endpoints of the continuum, Step 1 with 26 ms of VOT and a total length of 310 ms and Step 4 with 86 ms of VOT and a total length of 370 ms, were clear tokens of the two target words, "deer" and "tear," respectively. The intermediate steps could be perceived as either of the two target words. Each continuum step was concatenated to the word "the" taken from a sentence where it preceded the word "tear." 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 units ("the dear" and "the tear") were equalized in root-mean-square (RMS) amplitude. The target words were saved in units containing the article "the" so that they could be presented either before or after the context sentence in the presentation software.

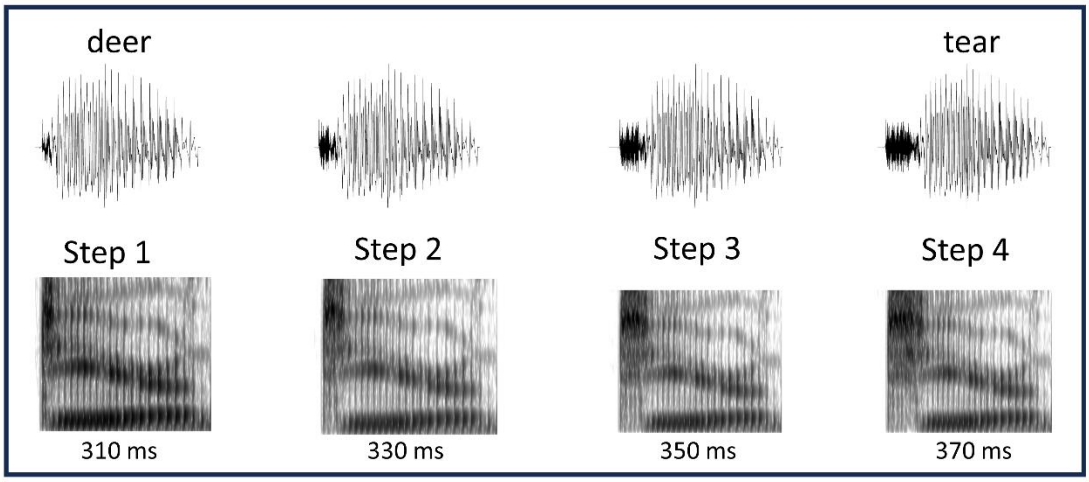

Figure 2-1. Waveforms and spectrograms of the four steps along the dear-tear continuum. The length of the word in milliseconds (ms) is provided under each step.

Identification of the target words was confirmed with behavioral judgments from young AH participants in a pilot study. The two target words at the endpoints were perceived with 100% accuracy, confirming that the "deer" stimulus was perceived as "deer," and the created "tear" stimulus was 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 2-I. When given the choice of the two endpoint words, the intended endpoint was chosen to complete the context sentences 100% of the time in pilot testing. All pilot testing was completed in quiet with no background noise.

All auditory stimuli were sine-vocoded. The auditory waveforms were divided into 8 logarithmically spaced channels between 250-4000 Hz using 3rd-order Butterworth filters. Envelopes were extracted by half-wave rectifying the contents of each channel and then using a 150-Hz 2nd-order Butterworth filter, applied forward and

backward (Shader, Yancey, 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). The VOT contrast in the deer/tear continuum does not critically rely on energy above 4000 Hz. The background noise used was speech-shaped noise that was created from the long-term average spectrogram 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 & 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 SNR before the entire signal was vocoded. This range of SNRs was chosen so that the "moderately degraded" level of background noise, known to cause maximum context effects (Bhandari et al., 2021; Boothroyd & Nittrouer, 1988), would be included for most participants.

Table 2-I. Sentences used to provide sentence-initial and sentence-final context.

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 ____."

2.2.3 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 (target words and context sentences) were presented over a loudspeaker (Yamaha HS5, Buena Park, CA, USA) approximately one meter directly in front of the participant's chair or over circumaural headphones (Sennheiser HD-650, Old Lyme, CT, USA) for seven participants because of equipment availability. The stimuli were calibrated in the sound field to be presented at 70 dB A using a sound level meter with a quarter-inch microphone (Brüel & Kjær Type 2250, Nærum, Denmark) to measure a segment of speech-shaped noise set to the same root-mean-square (RMS) amplitude as the target stimuli. A 1000-Hz tone with the same RMS 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. Stimuli were presented and responses were recorded using PsychoPy software v. 2022.1.4 (Open Science Tools Ltd., Nottingham, UK).

2.2.3.1 Practice

The visual stimuli were presented on a 23-inch computer monitor located 0.75 meters from the participant's chair. Prior to beginning the experiment, participants completed a short practice session during which they heard the context sentences that would be used in the experiment. During each self-initiated trial, the word "Ready…" showed on the screen for 1200 ms. The sound file of the context sentence and target word (both in quiet) played while the screen showed the printed sentence they were hearing with the target word represented by a blank. After the auditory presentation

was complete, the question “Which word did you hear?” appeared on the top half of the screen and two color photographs or drawings 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 one and four) were presented. Participants were informed that occasionally the word might not match the sentences. Two examples of incongruent contexts were included among the 12 context sentences presented during the first part of the practice. Correct answer feedback was provided.

The second part of the practice was a baseline measure of the participant’s response times. The 12 sentences presented were neutral (i.e., no predictive contexts) and the target words were presented in quiet. The third part of the practice introduced the background noise. Participants first heard three of the neutral sentences with the target words presented at +10 dB SNR. Then, two neutral sentences with target words were presented at each of the remaining four SNRs, becoming progressively more difficult. The last sentence in this section was presented in quiet. The total number of sentences in the third part of the practice was 12. The final part of the practice session measured baseline response times to 12 target stimuli at the most favorable SNR (+10

dB SNR). The sentences did not provide predictive context during these baseline measurements. The practice session in its entirety accomplished a number of objectives: 1) it ensured that the context sentences would be easily recognized and understood during the task; 2) it allowed some additional adaptation to the vocoded speech; 3) it provided familiarity with the task; 4) 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; and 5) it established individualized baseline response times for normalization during analysis.

2.2.3.2 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 along 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."

Each step along the continuum was presented 10 times at the beginning and end of each of the three types of sentence contexts at five different SNRs (4 steps × 3 sentence contexts × 2 sentence positions × 10 repetitions × 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 minutes to avoid fatigue. Performance was measured in two ways: 1) a percent "deer" response for each step along the continuum; and 2) response times for categorization choices.

2.2.3.3 Cognitive Tests

During the breaks between blocks of the experiment or after the last block, participants completed two cognitive measures: verbal processing speed and inhibition (Stroop task; Stroop, 1935) and working memory (NIH Toolbox List Sorting; Tulsky et al., 2014). The scores from these measurements were used in the analysis as predictors of the effect of context on speech recognition performance. Higher performance on these tasks was hypothesized to predict smaller context effects, where participants with higher scores might use the acoustic signal to categorize the words rather than rely completely on the context (Mattingly et al., 2018; Schubotz et al., 2021).

The protocol for the Stroop Color Word Task followed the widely used recommendations in the manual by Golden (2002). The task contained three sub-tasks. The first sub-task required the participant to read aloud, as quickly as possible, from a

list of words written in black font on a white background. The words were “red, blue, green, and yellow” in a random order. The words were presented on a computer screen from a PDF of the paper version of this task. The page contained 50 words in 10 rows of five words each. The page was sized so that the entire list was visible on the computer screen. Participants were instructed that if they reached the end of the page before the timer sounded, to start again at the top of the page. The timer was set for 45 seconds. Participants’ responses were recorded on a digital voice recorder for offline scoring. This first sub-task established a baseline reading speed for each individual, as well as a comparison for the next two sub-tasks. The second sub-task similarly required participants to report colors aloud as quickly as possible. However, the stimuli for the second sub-task were squares of the four colors used in the first sub-task. Participants were again instructed to name as many of the color squares aloud as possible within the 45-second time limit. If they reached the bottom of the page (again containing 10 rows of 5 colored squares) before the timer sounded, they were instructed to start again at the top of the page. The third sub-task required participants to say aloud the color the word was written in, not read the word, as quickly as possible. The same color words from the first task were used, but they were written in colors that did not necessarily match their lexical meaning. The stimulus page again contained 10 rows of five words, and participants were given the same 45-second time limit. According to the recommendations of Scarpina and Tagini (2017), the number of correct responses in each of the three subtasks was reported. A global variable was also calculated to represent the participant’s inhibition ability by calculating a predicted color-word score

with the formula (W × C)/(W + C). This number was then subtracted from the actual measured color-word score. A positive number suggested greater inhibition ability and less interference than a negative number which indicated an interference effect (Golden, 2002). This global variable was used in the analysis to indicate inhibition ability. The difference between the word-only score and the color-only score was used in the analysis to represent the lexical access/processing speed of an individual. The difference between the number of words reported in these two conditions is driven by the fact that the participant must identify the color of the block, access the word associated with it, and form that word to say it aloud (color-only condition) versus being given the word and simply forming the word to say it aloud (word-only condition). These measures of processing speed were mean-centered, standardized, and multiplied by negative one so that larger or more positive numbers represented faster processing speeds.

The NIH List Sorting task (Tulsky et al., 2014) was used to present participants with an ever-increasing number of pictures of animals or food items. The participant was presented the pictures on an iPad screen. They were required to remember the pictures in the list, sort them in size order from smallest to biggest, and then report the sorted list aloud. There were two practice trials containing first two and then three items in the list. Participants received feedback on their performance during the practice trials, but not during the experimental trials. The experimental trials began with two items in the list to be sorted. If the participant answered correctly, the number of items in the list increased by one. A second opportunity with the same number of items was

provided if a participant answered incorrectly. If they answered incorrectly for both trials at a single number of list items, the test was halted and moved to the second phase. The second phase was slightly more complex than the first phase. In the second phase, the participant was required to separate the items from the list into food and animal categories and then sort the items in those categories in size order from smallest to biggest. Again, there were two practice trials with feedback. After the practice, the first trial contained two items, and the participants reported the food category before the animal category. As the number of items in each list increased, the pictures were presented in an alternating pattern between the two categories. The stopping rule was the same: when a participant had responded incorrectly to two trials with the same number of items in the list, the test was over. The age-corrected standard score was used as a normalized measure of working memory in the analyses.

2.2.4 Analysis

Three analyses were conducted to assess the processing of phonetic cues with varying context. The analyses focused on the categorization of contrasting words, the derived context effect, and response times. Specific analyses varied with each processing measure, as described below.

The binomial distribution of the categorization choice (“deer” or “tear”) from each participant in each condition was modeled with a logistic mixed-effect model (MEM) containing fixed effects of age, VOT, context type (neutral, deer, or tear), sentence position (initial or final), SNR (5 levels), and their interactions. Age, VOT,

and SNR were treated as continuous variables. All three variables were standardized such that the mean of each variable (mean age = 49.2 years, mean VOT = 56 ms, mean SNR = 0 dB SNR) was zero (with a standard deviation of one) and was the reference value in the model. The reference values for the categorical variables of context type and sentence position were "neutral" and "final," respectively. The *buildmer* package, version 2.11 (Voeten, 2023), was used to create a model with minimal experimenter bias. The experimenter inputs the formula for a maximal model of both fixed and random effects. In this case, the five fixed effects described above were entered as fully interacting with each other and the random effects included a random intercept of participant with random slopes of SNR, sentence position, and VOT for each participant. The *buildmer* package tests 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 statistical analysis was conducted in R (version 4.3.3; R Core Team, 2024) using the *buildmer* package and *lme4* package version 1.1.35.3 (Bates et al., 2015).

The context effect for each participant was calculated at each SNR and target word position (see description of this calculation near Figure 2-4). The number of "deer" responses in each of the two predictive context sentences was compared to the number of "deer" responses in the neutral context sentence. The differences were added together to create an individual's context effect for that condition. This context effect value was log-transformed to reduce the skewness of the distribution. The log-

transformed value was used as the dependent variable in a linear MEM with a Gaussian distribution.

The standardized age-corrected score for the working memory task was provided by the NIH Toolbox. This value was then centered, scaled, and treated as a continuous variable. The variable used to represent an individual's processing speed was derived from the raw scores of the Stroop task. The number of color blocks correctly reported was subtracted from the number of black-ink words correctly reported (Word score – Color score). This value was then centered, scaled, and treated as a continuous variable. The variable used to represent an individual's ability to inhibit irrelevant information was also derived from the raw scores of the Stroop task. The product of the number of color blocks correctly reported and the number of black-ink words correctly read was divided by the sum of the number of color blocks correctly reported and the number of black-ink words correctly read. The result was then subtracted from the number of colored-ink words correctly reported {Color-Word score – [(Color score × Word score) / (Color score + Word score)]}. This value was then centered, scaled, and treated as a continuous variable. Larger, more positive, numbers indicate higher working memory ability, faster processing speed, and greater inhibition ability.

Once again, the *buildmer* package, version 2.11 (Voeten, 2023), was used to develop the model. The initial fixed effects were SNR, Sentence Position, Age, Working Memory, Processing Speed, and Inhibition ability. SNR and age were centered, scaled, and treated as continuous variables. Sentence Position was a

categorical variable with a reference value of sentence-final position. Processing Speed and Inhibition Ability were continuous variables. The initial random effects included random intercepts for each participant, with random slopes of SNR and Sentence Position.

The response times for all trials were modeled in a linear MEM (Gaussian distribution). Response times greater than 10 seconds and less than 0.1 second were removed from the analysis because response times less than 0.1 second are physically impossible after a decision, while response times greater than 10 seconds 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. The fixed effects were SNR, Target Word Acoustical Ambiguity, Age, Sentence Position, and Context Sentence Congruency. SNR was centered, scaled, and treated as a continuous variable. 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. Age was centered, scaled, and treated as a continuous variable. Sentence 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 incongruent 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 will account for individual differences in overall response times. The *buildmer* package, version 2.11 (Voeten, 2023), was used to identify the best-fit model as described previously. Post-hoc testing included re-leveling the Context categorical variable such that the deer-predictive context became the new reference level. This allowed a comparison between performance in the deer-predictive sentences and the tear-predictive sentences. The residuals of each model were checked to ensure there were no major violations of normality or homoscedasticity.

2.3 Results

2.3.1 Categorization Performance: "Deer" Responses

The results are presented first as average performance for all participants at each continuum step, at each SNR, and in each sentence position (Figure 2-2). The top row represents the number of "deer" responses at each continuum step when the target words were presented at the beginning of the sentences. The lower row shows the same data when the target words were presented at the end of the sentences. The various SNRs are presented in columns from the most challenging SNR on the left to the easiest SNR on the right. The red circles represent responses when the context sentence

predicted a “deer” target word. The blue squares represent responses when the context sentence predicted a “tear” target word. The black triangles represent responses when the context sentence was neutral and did not predict either target word.

A context effect can be seen when the red line, representing the deer-predictive sentences, is above the black neutral line and when the blue line, representing the tear-predictive sentences, is below the black neutral line. Higher values on the chart represent more “deer” responses and lower values represent more “tear” responses. The effect of context on responses is small until -5 dB SNR for the sentence-initial target words (Figure 2-2 top row) and is small but consistent across SNRs for the sentence-final target words (Figure 2-2 bottom row).

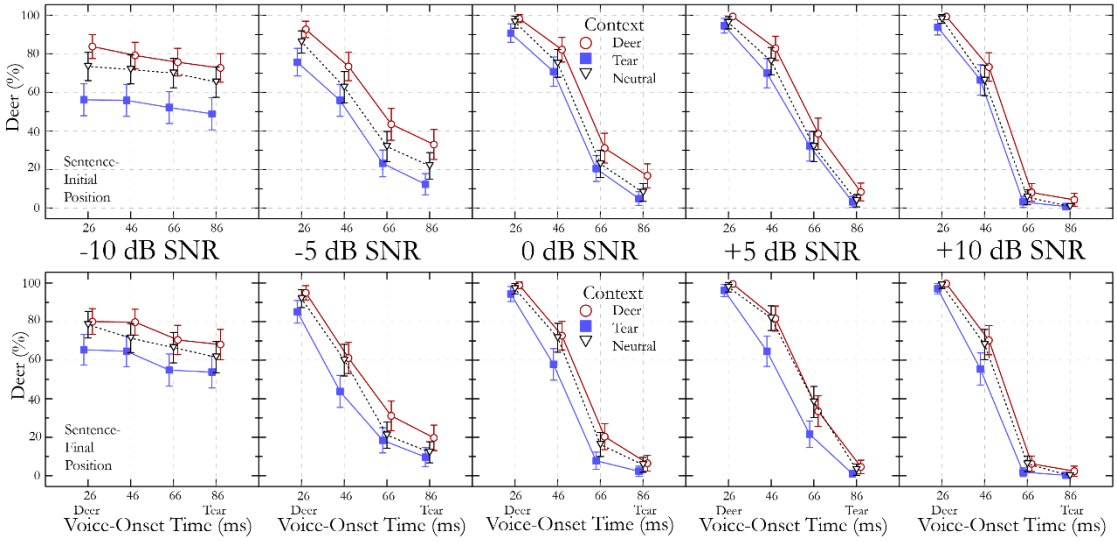

Figure 2-2. Percentage of "deer" responses for all listeners (n=36) divided by target word position within the sentence (top, bottom), SNR (columns), and sentence contexts (colors/symbols). Error bars are ± 1 standard error of the mean.

The output of the model that was fit to the trial level data is shown in Table 2-II. The model shows that the probability of choosing “deer” as the response is greatly

affected by the experimental manipulations of SNR and VOT duration ($p < 0.001$ for both) as expected. The probability of a "deer" response decreased significantly when the sentence context predicted the "tear" interpretation compared to the neutral sentence context (Context-Tear, $p < 0.001$), but did not change significantly when the sentence context predicted the "deer" interpretation (Context-Deer, $p > 0.05$). The context categories were re-leveled for post-hoc testing, which revealed a significantly lower probability of a "deer" response in the tear-predictive contexts compared to the deer-predictive contexts ($p < 0.001$).

In response to the VOT cue, the likelihood of a "deer" response should be higher for shorter VOTs and lower for longer VOTs. There were significant interactions such that the difference between the number of "deer" responses at shorter and longer VOTs was reduced at more difficult SNRs (VOT × SNR, $p < 0.001$), when the target word was presented in the sentence-initial position (VOT × Position-Initial, $p < 0.001$), and with increasing age (VOT × Age, $p < 0.001$).

There were two significant three-way interactions involving VOT (VOT × SNR × Position-Initial, $p < 0.001$ and VOT × SNR × Age, $p < 0.001$). These interactions should be viewed with caution given the vastly different performance in the -10 dB SNR condition compared to the other SNR conditions [see Figure 2-3 (A) and (B)]. Since most participants reported that they were guessing at -10 dB SNR, these interactions likely reflect the fact that participants were unable to hear the acoustic target words in that condition.

Table 2-II. Model output for a logistic mixed effects model. The dependent variable is the binary "deer" or "tear" choice, and the independent variables are SNR, Sentence Position, Context, VOT, and Age. Reference values are Final Sentence Position, Neutral Context, and average VOT, SNR, and Age.

Formula: Deer-or-Tear ~ VOT × SNR × (Sentence Position + Age) + SNR × Sentence Position × (Age + Context) + (VOT| Participant)

	Deer or Tear				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	0.099	0.089	1.108	0.268	
VOT	-2.501	0.075	-33.272	< **0.001**	***
SNR	0.452	0.037	12.318	< **0.001**	***
Position-Initial	0.003	0.051	0.051	0.959	
Age	-0.155	0.087	-1.779	0.075	
Context-Deer	-0.028	0.053	-0.531	0.595	
Context-Tear	-0.680	0.053	-12.920	< **0.001**	***
VOT × SNR	1.451	0.032	44.878	< **0.001**	***
VOT × Position-Initial	0.413	0.046	9.064	< **0.001**	***
VOT × Age	0.312	0.073	4.242	< **0.001**	***
SNR × Position-Initial	0.021	0.050	0.424	0.672	
SNR × Age	-0.190	0.023	-8.306	< **0.001**	***
SNR × Context-Deer	0.148	0.052	2.872	**0.004**	**
SNR × Context-Tear	0.176	0.051	3.461	< **0.001**	***
Position-Initial × Age	0.065	0.031	2.071	**0.038**	*
Position-Initial × Context-Deer	0.388	0.072	5.399	< **0.001**	***
Position-Initial × Context-Tear	0.341	0.072	4.771	< **0.001**	***
VOT × SNR × Position-Initial	-0.195	0.042	-4.617	< **0.001**	***
VOT × SNR × Age	-0.156	0.023	-6.780	< **0.001**	***
SNR × Position-Initial × Age	0.061	0.031	1.965	**0.049**	*
SNR × Position-Initial × Context-Deer	-0.088	0.071	-1.235	0.217	
SNR × Position-Initial × Context-Tear	-0.393	0.070	-5.622	< **0.001**	***
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
By-Participant effects	0.233	0.483			
By-Participant VOT slopes	0.156	0.395	-0.38		
Observations	43200				

While there was no overall difference in the likelihood of a "deer" response as a function of target word position, there was an interaction such that a significant difference in "deer" responses based on target word position (fewer in the sentence-final position compared to those in the sentence-initial position) emerged with increasing age (Position-Initial × Age, $p = 0.038$). In the neutral sentence context, the number of deer responses did not change between sentence-final and sentence-initial

presentations of the target word. In the two predictive sentence contexts, there were significantly more "deer" responses when the target word was at the beginning of the sentence compared to when the target word was at the end of the sentence (Position-Initial × Context-Deer, $p < 0.001$; Position-Initial × Context-Tear, $p < 0.001$). There was also a significant three-way interaction such that performance on sentence-initial target words that were presented in tear-predictive sentence contexts changed with the various SNRs differently than for target words presented in neutral contexts and deer-predictive contexts (SNR × Position-Initial × Context-Tear, $p < 0.001$) [see Figure 2-3 (D)].

In general, with increasing age, the difference in number of "deer" responses at different levels of background noise was reduced. That is, older participants had fewer "deer" responses in the most difficult SNR condition and more "deer" responses in the easiest SNR condition compared to younger listeners (SNR × Age, $p < 0.001$). There was also an interaction between SNR and context such that the likelihood of a "deer" response increased as the SNR became more difficult at a significantly different rate in the deer-predictive and tear-predictive sentence contexts compared to the neutral context (SNR × Context-Deer, $p = 0.004$; SNR × Context-Tear, $p < 0.001$). Post-hoc releveling revealed no significant difference between the deer-predictive and tear-predictive sentence contexts in the rate of increased likelihood of "deer" responses as SNR became more difficult (SNR × Context-Tear with Context-Deer as reference, $p > 0.05$). There was also a significant three-way interaction such that older participants showed less of an effect of SNR, especially for words presented in the sentence-final

position, compared to younger participants (SNR × Position-Initial × Age, $p = 0.049$) [see Figure 2-3 (C)].

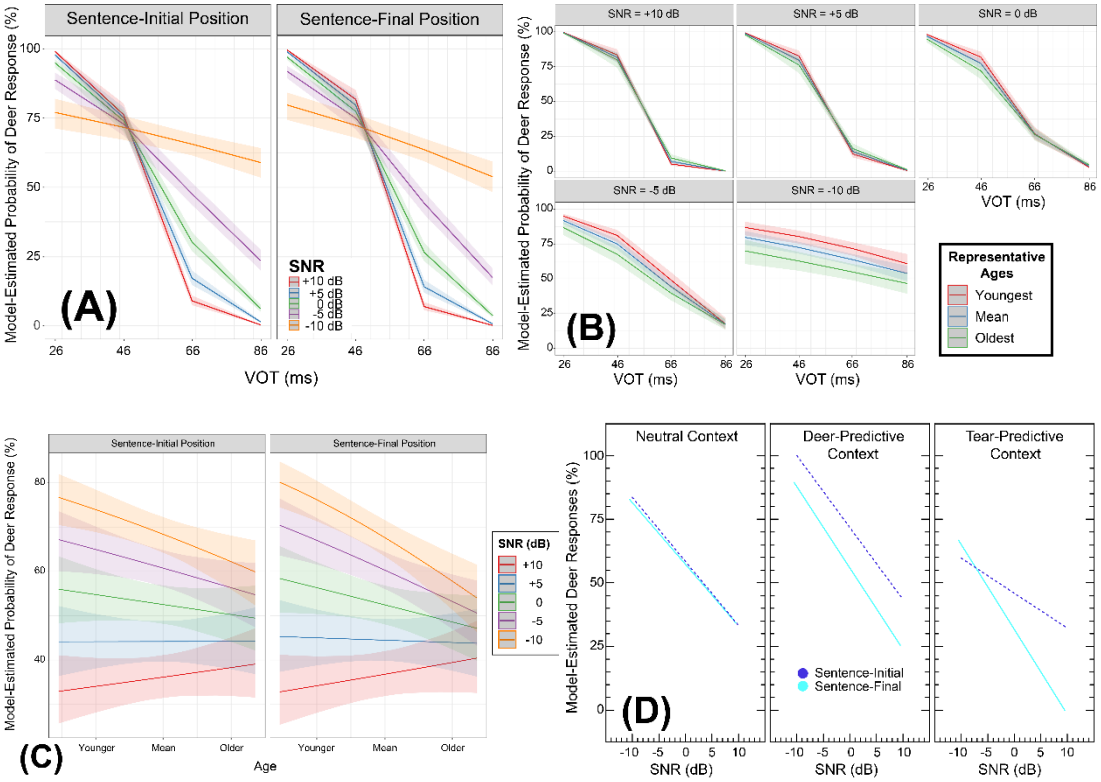

Figure 2-3. Panel (A) shows the three-way interaction between sentence position, voice-onset time (VOT), and signal-to-noise ratio (SNR) on the model's estimates of "deer" responses. The colors represent the different SNRs and the panels represent the target word's position. Panel (B) shows the three-way interaction between SNR, VOT, and age on the model's estimates of "deer" responses collapsed. The colors correspond to three representative ages and the panels represent the SNRs. Panel (C) shows the three-way interaction between sentence position, SNR, and age on the model's estimates of "deer" responses collapsed across VOTs. The colors represent the SNRs and the panels represent the target word's position. Panel (D) shows the three-way interaction between sentence position, SNR, and semantic context on the model's estimates of "deer" responses collapsed across VOTs. The dark dotted line represents estimates of performance on target words in the sentence-initial position. The lighter solid line represents estimates of performance on target words in the sentence-final position.

2.3.2 Cognitive Measures and the Context Effect

The raw and derived scores on the Stroop and working memory tasks are shown in Table 2-III for 35 participants (as per Golden, 2002; Scarpina & Tagini, 2017). One participant did not complete the cognitive tests during their visit and was unable to return. The score for the working memory task was the standardized age-corrected score provided by the NIH Toolbox. The raw scores for the Stroop are simply the number of words correctly reported by the participant within the allotted time. The derived scores represent an individual's processing speed (Color score – Word score) and an individual's ability to inhibit irrelevant information {Color-Word score – [(Color score × Word score) / (Color score + Word score)]}. Larger, more positive, numbers indicate higher working memory ability, faster processing speed, and greater inhibition ability. Pearson's product-moment correlations were performed between age and working memory [$r = 0.272$; $t(33) = 1.62$, $p > 0.05$], age and inhibition [$r = -0.495$; $t(33) = -3.27$, $p = 0.003$], and age and processing speed [$r = 0.194$; $t(33) = 1.14$, $p > 0.05$]. Inhibition was the only cognitive measure significantly correlated with age (negative).

*Table 2-III. Scores on cognitive tests, both raw and derived. Inhibition was calculated as Color-Word score – (Color score * Word score)/(Color score + Word score). Processing speed was calculated as (Word score – Color score)*(-1).*

Participant	Age (yrs.)	Gender	Standardized Working Memory Age-Corrected	Raw Words	Raw Colors	Raw Color-Words	Inhibition (Derived)	Processing Speed (Derived)
811	20	M	102	92	70	38	-1.75	-22
815	20	F	107	113	78	59	12.85	-35
816	20	F	96	133	83	65	13.89	-50
817	20	F	107	123	63	32	-9.66	-60
812	21	M	112	143	99	61	2.50	-44
818	26	F	81	95	62	32	-5.52	-33
800	28	F	119	78	69	50	13.39	-9
826	30	F	109	89	77	47	5.72	-12
833	30	F	98	108	71	45	2.16	-37
822	31	F	77	90	86	53	9.02	-4
824	31	F	109	89	76	47	6.01	-13
832	38	F	106	116	90	55	4.32	-26
820	43	F	113	70	70	45	10.00	-0
827	44	M	91	92	96	58	11.02	4
834	44	F	119	134	109	62	1.89	-25
837	44	F	113	81	63	38	2.56	-18
835	46	F	103	90	65	23	-14.74	-25
830	52	F	100	110	75	45	0.41	-35
819	53	F	101	89	70	37	-2.18	-19
821	53	M	107	107	52	34	-0.99	-55
829	55	F	85	94	75	46	4.28	-19
809	56	F	107	105	96	61	10.85	-9
831	56	F	119	101	66	38	-1.92	-35
825	57	F	108	115	87	44	-5.53	-28
814	65	F	146	114	80	41	-6.01	-34
823	65	F	112	99	96	29	-19.74	-3
804	68	F	108	80	59	25	-8.96	-21
806	68	F	102	94	65	33	-5.43	-29
803	69	F	91	110	73	52	8.12	-37
805	69	F	92	88	60	23	-12.68	-28
808	69	F	109	78	68	40	3.67	-10
801	71	F	116	103	64	32	-7.47	-39
810	72	F	134	92	66	26	-12.43	-26
807	74	F	135	103	93	45	-3.87	-10
813	76	F	96	85	71	29	-9.69	-14

The phoneme categorization data can be converted into a single number demonstrating the effect of context for an individual at each SNR and in each sentence position. 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. This approach collapses responses across all four steps (VOT durations) of the acoustic continuum. 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 context sentences than in the neutral sentences. A negative context effect would be the result of fewer "deer" responses in the deer-predicting context sentences than in the neutral sentences or fewer "tear" responses in the tear-predicting 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.

In the example in Figure 2-4, the graph shows the participant's number of deer responses at each of the four steps along the VOT continuum in each of the sentence contexts. The total number of deer responses in the deer-predictive context sentences (red circles) across the VOT continuum is 26. The total number of deer responses in the neutral sentence is 21. The total number of deer responses in the tear-predictive context sentences is six. The context effect is therefore calculated as the sum of the differences between the number of "deer" responses in the deer context (26) compared to the neutral

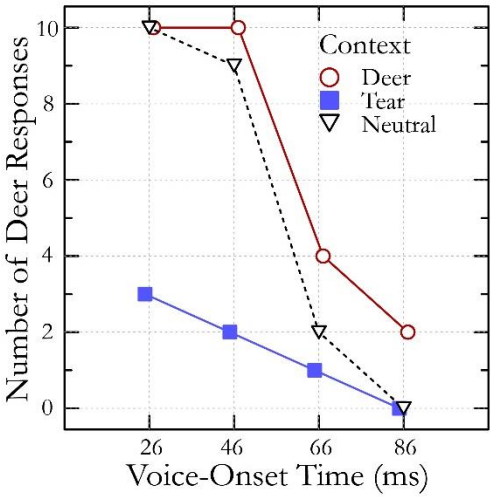

Figure 2-4. Example data from one participant to illustrate the calculation of the context effect.

context (21) and between the neutral context (21) compared to the tear context (6), which equals 20 [(26-21) + (21-6)].

In order to visualize the effects of context, Figure 2-5 shows the average performance of participants separated into two groups based on the median age (53 years old): a younger group (ages 20-52 years, mean = 32.7 years, n=18), and an older group (ages 53 – 76 years, mean = 64.8 years, n=18). Both age groups showed very similar patterns of performance across age groups and positive context effects at all SNRs tested for target words presented at both the beginning and the end of sentences..

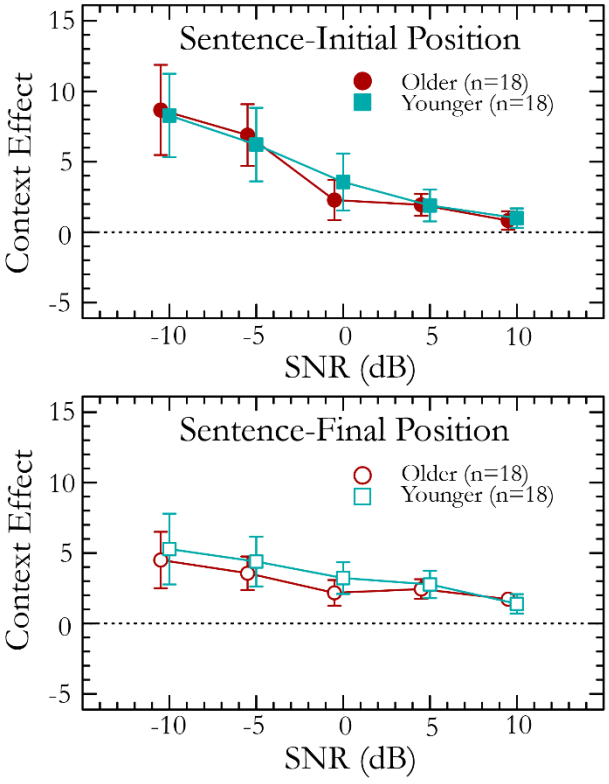

Figure 2-5. Average context effect compared to performance in the neutral context sentences of older participants (red circles) and younger participants (cyan squares) across SNRs for target words presented at the beginning of context sentences (top) and at the end (bottom). Error bars are ± 1 standard error of the mean.

At the most difficult SNRs, participants on average show an increase in the use of context. This is a logical response to an inaudible target word. When the target word was presented in the sentence-initial position, both younger and older participants demonstrated a greater context effect in the more difficult SNRs compared to when the target word was in the sentence-final position. Figure 2-6 shows the context effect at each SNR as a function of target word position collapsed across all participants.

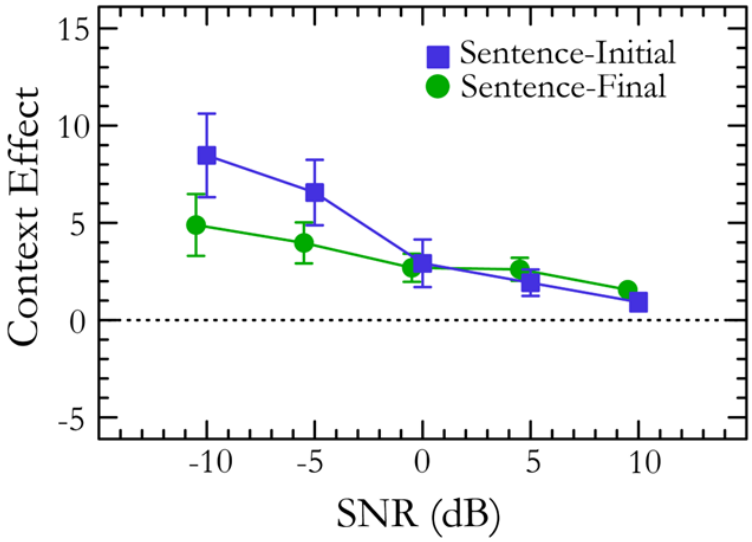

Figure 2-6. The average context effect across all participants at each SNR condition for target words presented at the beginning of the context sentence (blue squares) and at the end (green circles). Context effect is calculated as the difference in performance between the two predictive sentence contexts and the neutral sentence context (represented by the dotted line). See text for more detail.

A linear MEM was created to test the influence of the cognitive measures of inhibition ability, processing speed, and working memory on the context effect. To that end, 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 (see description of calculation of context effect near Figure 2-4). The log-transformed

context effect was used as the dependent variable in the linear MEM. The final model is presented in Table 2-IV. Both age and working memory were excluded from the final model by the *buildmer* package because they did not contribute significantly to the model fit. While inhibition ability was kept in the model, it did not contribute to any significant terms. Processing speed was the only cognitive measure that was both included in the final model and contributed to a significant term.

Table 2-IV. Model output for a linear mixed-effects model fit to log-transformed derived context effect. Reference values: Final Sentence Position, average SNR, processing speed, and inhibition ability.

Formula: Context Effect ~ SNR × Sentence Position × (Processing Speed + Inhibition Ability) + (SNR + Sentence Position | participant)

	Context Effect				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	2.543	0.085	30.044	**<0.001**	***
SNR	-0.016	0.065	-0.243	0.808	
Position-Initial	0.011	0.094	0.121	0.904	
Speed	0.002	0.003	0.612	0.545	
Inhibition	-0.005	0.005	-0.986	0.331	
SNR × Position-Initial	-0.225	0.058	-3.913	**<0.001**	***
SNR × Speed	0.001	0.002	0.371	0.712	
SNR × Inhibition	-0.003	0.004	-0.719	0.476	
Position-Initial × Speed	0.000	0.003	0.126	0.901	
Position-Initial × Inhibition	0.002	0.006	0.310	0.758	
SNR × Position-Initial × Speed	-0.006	0.002	-2.979	**0.003**	**
SNR × Position-Initial × Inhibition	0.006	0.004	1.808	0.072	
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
By-Participant effects	0.052	0.228			
By-Participant SNR slopes	0.024	0.154	-0.85		
By-Participant Sentence Position slopes	0.052	0.228	0.30	-0.71	
Observations	350				

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 Figure 2-6) (SNR × Position-Initial, $p<0.001$). There was a significant three-way interaction involving processing speed (SNR × Position-Initial × Speed, $p=0.003$). Post-hoc plots revealed

that, for most SNR conditions, increased processing speed was associated with slightly greater context effects for target words in the sentence-final position compared to those in the sentence-initial position. In the -10 dB SNR condition, however, the opposite trend can be seen. Increased processing speed appears to be associated with greater context effects for target words in the sentence-initial position compared to those in the sentence-final position. Given that these effects only appear in the -10 dB SNR condition, this interaction should be interpreted with caution. Overall, measures of working memory, processing speed, and inhibition ability did not significantly predict the context effect experienced by participants in this experiment.

2.3.3 Response Time Measures

Response times were collected for all trials. Response times greater than 10 seconds and less than 0.1 second were removed from the analysis because response times less than 0.1 second are physically impossible after a decision, while response times greater than 10 seconds are likely to represent lapses in attention or other factors not related to decision making. This left 43,023 response times to analyze. The remaining values were log-transformed to reduce the skewness of the distribution and approximate a Gaussian distribution. A category for 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. A linear MEM was fit to the data with SNR, Target Word Acoustical Ambiguity, Age, Sentence Position, and Context Sentence Congruency as fixed effects.

The model shows that response times were significantly predicted by the interactions between SNR, Sentence Position, VOT, Age, and Context Congruency (Table 2-V). In general, response times were shorter when the SNR was easier, when the target word occurred at the beginning of the sentence, and when the context sentence was congruent (SNR, $p < 0.001$; Position-Initial, $p < 0.001$; Context-Congruent, $p < 0.001$). Response times were generally longer when the target words were acoustically ambiguous and with increasing age (TargetWords-Ambiguous, $p < 0.001$; Age, $p < 0.001$).

Table 2-V. Model output for a linear mixed-effects model fit to log-transformed response time data. Reference values: Final Sentence Position, Unambiguous Target Words, Incongruent Sentence Contexts (both neutral and opposite contexts), average SNR and Age.

Formula: Response Time ~ SNR * Sentence Position * Target Word Acoustic Ambiguity + Sentence Position * Age * (SNR + Context Congruency) + SNR * Congruency + (SNR + Position + Ambiguity | participant)

	Response Time				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	-0.210	0.058	-3.592	**0.001**	*
SNR	-0.178	0.017	-10.201	**<0.001**	***
Position-Initial	-0.179	0.024	-7.416	**<0.001**	***
TargetWords-Ambiguous	0.109	0.009	12.462	**<0.001**	***
Age	0.237	0.060	3.957	**<0.001**	***
Context-Congruent	-0.022	0.006	-3.401	**<0.001**	***
SNR × Position-Initial	0.082	0.006	13.702	**<0.001**	***
SNR × TargetWords-Ambiguous	0.063	0.006	10.480	**<0.001**	***
Position-Initial × TargetWords-Ambiguous	-0.090	0.008	-10.592	**<0.001**	***
Position-Initial × Age	0.013	0.025	0.516	0.609	
Position-Initial × Context-Congruent	0.002	0.009	0.178	0.858	
SNR × Age	-0.024	0.018	-1.353	0.185	
Age × Context-Congruent	-0.011	0.007	-1.605	0.109	
SNR × Context-Congruent	0.008	0.004	1.840	0.066	
SNR × Position-Initial × TargetWords-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-Congruent	0.019	0.009	2.012	**0.044**	*
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
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 Word Ambiguity slopes	0.001	0.038	-0.19	0.07	0.14
Observations	43023				

Response times to acoustically ambiguous target words appear to be longer than to acoustically unambiguous target words mainly at the more favorable SNRs especially when the target word was at the end of the sentence [see Fig 2-7(A) and (B)] (SNR × Position-Initial × TargetWords-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 30-60 years old than it did for those younger than 30 years old or older than 60 years old, except at -10 dB SNR, where the oldest listeners showed the largest effect [see Figure 2-7 (C)](SNR × Position-Initial × Age, $p < 0.01$). This pattern of the difference in response times based on the target word's sentence position and the listener's age is also affected by the congruency of the context sentence with the target word (Position-Initial × Age × Context-Congruent, $p < 0.05$). Delayed response times for incongruent sentence contexts appear to be present in the data of participants around age 40 and again around age 70 [see Figure 2-7(D)]. The oldest participants showed greater delays in response times when the target word was at the end of a sentence than when it was at the beginning of the sentence. For illustration purposes in Figure 2-7 (C) and (D), participants were grouped by age into decades and their performance averaged across the members of the group (5-7 participants/group).

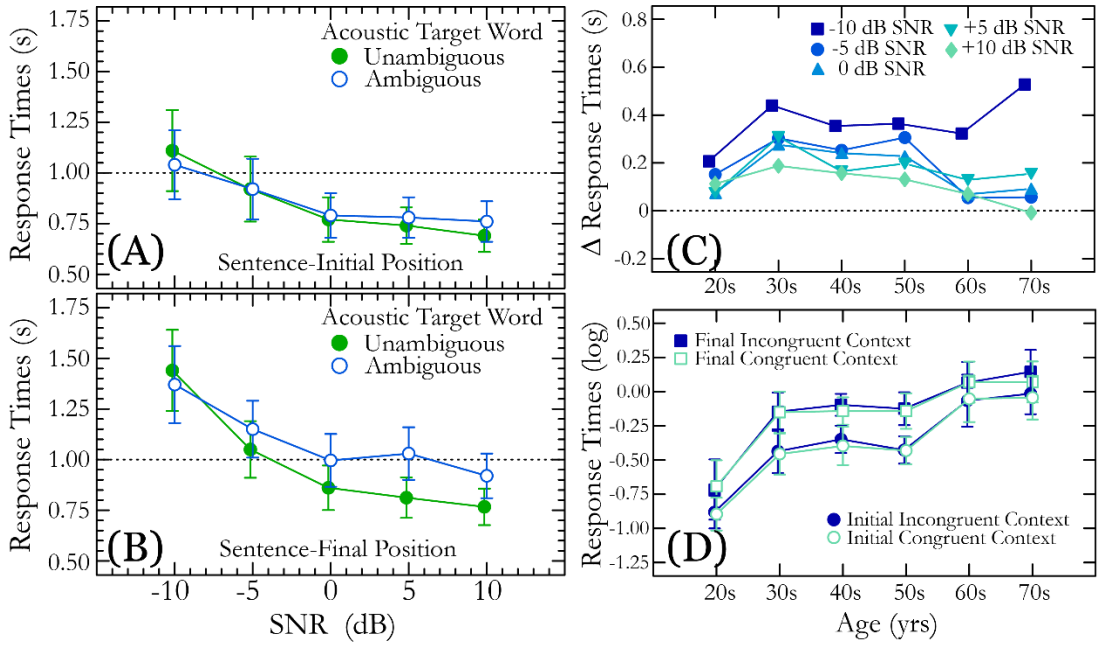

Figure 2-7. Figure 2-8. Panel (A) shows raw response times plotted as a function of SNR when the acoustic target word was unambiguous (solid symbols) compared to ambiguous (open symbols) and was presented at the beginning of the sentence. Panel (B) shows raw response times plotted as a function of SNR when the acoustic target word was unambiguous (solid symbols) compared to ambiguous (open symbols) and was presented at the end of the sentence. Error bars represent ± 1 standard error of the mean. The dotted line at 1.00 second is a reference for comparing across panels. Panel (C) shows the difference between response times when the target word is presented at the end of the sentence compared to the beginning of the sentence across ages as a function of SNR. Panel (D) shows log-transformed response times across ages as a function of target word position and context congruence. Error bars represent ± 1 standard error of the mean.

2.4 Discussion

This study tested adult participants with AH with a wide range of ages on a phoneme categorization task where the target word was presented in background noise at the beginning or end of spectrally degraded context sentences. The broad goal was to examine the interaction of age, signal degradation, and sentence position on the

perception of speech sounds. The central hypothesis was that increased age would result in a smaller difference between context effects found at the beginning vs the end of a sentence, and that participants would show a smaller effect of context for sentence-initial target words compared to target words in the sentence-final position.

The first research question was: 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? The hypothesis 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, given the lesser effects for subsequent context compared to preceding context reported in Connine (1991) and Wingfield et al. (1994). Results showed that the position of the target word had a significant effect on the context effect (Figure 2-6). The position of the target word also interacted with context and SNR to differentially affect categorization performance (Figure 2-3, Table 2-II). Contrary to previous research and the hypothesis, there was a greater effect of context when the target word was at the beginning of the sentence compared to when it was at the end of the sentence. The difference in the current findings compared to previous studies may lie in the differences in methodology and the calculation of context effects. Connine et al. (1991) analyzed the total number of /d/ responses between a context sentence predicting the /d/ response and a context predicting the /t/ response to determine the context effect. The method of calculating the effect of context by comparing the psychometric functions of a phonemic categorization task to performance with a neutral baseline is unique to this study. The current method corrects for any inherent response

biases on an individual basis. Also contrary to the initial hypothesis, a measure of working memory was not correlated with the context benefit when the target word was at either end of the sentence.

The finding that the tear context seemed to cause a larger difference in performance from the neutral context than the deer context may reflect the greater difficulty in detecting the acoustics of the voiceless /t/ sound. Participants appeared to be more willing to categorize a phoneme as /t/ that they might otherwise have categorized as /d/ in the presence of the tear context than to do the reverse in the presence of the deer context.

The second research question was: Is the age of the listener related to the extent to which listeners use context when identifying a spectrally degraded noise-masked target word in a spectrally degraded sentence? It was hypothesized that older participants' identification of target words would show similar patterns of context effects as younger participants, but the response times of older participants would show different patterns of context effects than those of younger participants. In support of this hypothesis, Figure 2-5 shows that older participants demonstrated similar patterns and a similar magnitude of the context effect as younger participants overall. This matches previous literature that shows little to no age effect on speech perception when audibility is assured (e.g., Dubno et al., 2000). There were, however, significant interactions between age, SNR, VOT, and target word position within the sentence (Table 2-II). The differences in performance across VOT by age were only seen in the more difficult SNR conditions [Figure 2-3(B)] and may reflect a combination of age-

related processing deficits (e.g., Fitzgibbons & Gordon-Salant, 1996) and age-related differences in strategy when listening in noise. The small but significant differences in performance across age by target word position occurred across positive and negative SNRs. Thus, while there were no large differences in the context effects seen across age (Table 2-IV and Figure 2-5), there were small differences across ages in the categorization performance (Table 2-II).

Table 2-V shows the expected effect of age-related slowing on response times. The hypothesized interaction between age, context congruency, and sentence position on response times was statistically significant [see Figure 2-7(D)]. While most participants showed little to no difference in response times based on congruency across sentence positions, the oldest participants showed a greater difference in response times (longer for incongruent than congruent sentence contexts) for sentence-final target words compared to sentence-initial target words. Except for the most difficult SNR condition (-10 dB SNR), older participants also showed very little difference in response times for sentence-initial compared to sentence-final target words. The largest effects of target word position were seen in the middle of the age range of participants for all SNR conditions except -10 dB SNR [see Figure 2-7(C)].

In addition, response times were longer for more difficult SNRs compared to easier SNRs, for sentence-final target words compared to sentence-initial target words, for ambiguous target words compared to unambiguous target words, and for incongruent sentence contexts compared to congruent sentence contexts. It is not surprising 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. However, there was a significant interaction between SNR and sentence position, such that the difference in response times to target words presented at the end of the sentence compared to those presented at the beginning of the sentence was significantly greater at more difficult SNRs compared to easier SNRs [Figure 2-7(A) and (B)]. There was also a significant interaction between SNR and acoustic ambiguity, such that response times to the acoustically ambiguous words were significantly slower than the response times to the unambiguous endpoint words for the easier SNRs, but not for the more difficult SNRs [Figure 2-7(A) and (B)]. Despite possible age-related reductions in the ability to inhibit the influence of incongruent information (e.g., Hasher et al., 1991), data from the current study showed that the response times of participants of all ages were similarly affected by the congruency of the context sentences with the target words.

The third research question was: 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? It was hypothesized that older participants would show smaller differences in the context effects between the two sentence positions than the younger participants. This hypothesis was not supported in the current findings. Participants across the adult ages tested used context to a similar degree in both sentence positions. Age was not a significant predictor of the effect of context, nor did it interact significantly with congruency to affect response times. Previous research found smaller effects of subsequent context for older listeners with AH compared to

younger listeners with AH, but not for target words at the end of sentences (Wingfield et al., 1994). The lack of any significant age effect in this study could indicate that there truly is no significant difference in the use of context by younger and older adults when the speech is spectrally degraded and the target words are partially obscured by background noise. As stated in the introduction, context effects are greatest when the speech is moderately degraded. The combination of vocoded speech and background noise, even at the easier SNRs may still have been too difficult to elicit maximal context effects, thus minimizing any age-related differences.

The context effects as seen in Figure 2-2 show minimal differences in performance based on the context sentences for most of the easier SNRs tested, especially when the target word was in the sentence-initial position. This indicates that participants are using the acoustic information available in the more positive SNRs to categorize the target words without relying on the context sentences. At the more difficult SNRs, however, clear evidence of the effect of context can be seen (see also Figure 2-6). All participants stated that at least one of the more difficult SNRs was "impossible." They were instructed to respond with their best guess for the target word. This led to a variety of strategies being used in the most difficult SNRs. On average, participants most often guessed the target word that was congruent with the context sentence or simply guessed the word "deer," although these were not the only strategies for all participants.

The additional question of whether an individual's working memory ability, inhibition ability, or processing speed were related to the context effects experienced

by that individual was not conclusively answered in the current study. It is possible that the measures used to quantify working memory, inhibition ability, and processing speed were not sensitive enough to adequately define an individual's abilities separate from their age. It is also possible that the cognitive abilities specifically tested in this study do not differentially predict the effects of context and target word sentence position. The only significant effect of any of the cognitive measures, processing speed, occurred in an interaction with SNR such that processing speed was only predictive of context effects in the -10 dB SNR condition. As previously discussed, performance in this SNR condition likely better represent guessing strategies rather than true speech understanding performance. Therefore, there were effectively no significant effects of the three cognitive measures on performance in this experiment.

2.4.1 Conclusion

This study has implications for people who are constantly listening to spectrally degraded speech, such as those who use CIs. CI processors convey a spectrally degraded representation of sound to the auditory nerves of listeners. 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. 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 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 experienced by listeners with normal hearing or with CIs in typical conversations.

It is possible that listeners of all ages can learn to benefit from context cues, given that age was not a significant predictor of context effects in the current study. Future studies that 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 CIs 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 et al., 2021).

In the current study, 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 -10 and -5 dB SNR. Also contrary to expectations, the effects of background noise level, target word position, and context sentence-target word

congruency were consistent across a wide range of ages. 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. This study supports the use of context cues to aid understanding of degraded speech for adults of all ages.

3 Study 2: Age and background noise alter the benefit of context for perceiving voice onset time contrasts in listeners with cochlear implants

3.1 Introduction

Cochlear-implant (CI) processors provide partial access to acoustic information. This partial information results in a representation of speech that is sometimes unclear. All listeners, including those who use CIs, can use the context of the surrounding words to help clarify unclear word(s) when the context is clear. An optimal strategy for understanding speech is to combine bottom-up processing of the acoustic information available to the listener with top-down processing. An application of top-down processing involves the use of cognitive abilities (e.g., processing speed and working memory), knowledge of the language, and experience to predict the message of the unclear or unheard pieces of the original utterance (Marslen-Wilson & Tyler, 1980; Morton, 1969).

The effects of age on speech perception with a CI are of vital interest to individuals who have been implanted with the device and are now growing older, as well as individuals who are over the age of 65 years and are currently contemplating receiving an implant. On the one hand, older adults typically have extensive language experience that remains intact with age (Ben-David et al., 2015; Verhaeghen, 2003) and is needed for the top-down use of context (Amichetti et al., 2018; Sommers &

Danielson, 1999). On the other hand, older adults may experience an age-related decline in the cognitive abilities associated with understanding degraded speech, such as working memory (e.g., Zekveld et al., 2013), processing speed (e.g., Rush et al., 2006; Wingfield, 1996), and inhibition (e.g., Sommers & Danielson, 1999; Stenbäck et al., 2021). At present, the effect of aging on the use of context to aid speech perception in older adults with CIs is unknown. In addition, it is unknown if the target word's position in a sentence interacts with predictive and non-predictive sentence contexts to influence how listeners with CIs perceive speech sounds and if those influences change with increasing age.

3.1.1 Phoneme Categorization

Phoneme categorization is a task that exploits the tendency of listeners to identify an ambiguous speech sound that is acoustically between the canonical versions of two phonemes, as one of those phonemes. The perceptual boundary between two phonemic categories is not usually a sharp step function. Instead, the same ambiguous speech sound can be categorized as either of the two phonemes, depending on such factors as the spectrum of the preceding speech (e.g., Shorey & Stilp, 2023) and the semantic context of the preceding speech (e.g., Amichetti et al., 2018; Connine, 1987; Connine & Clifton, 1987; Wingfield, 1996; Wingfield & Tun, 2001). The context effect can be observed in the location of the 50% crossover point between categories or in the response times. While response times are generally faster when the sentence context is congruent with the target word than when the sentence context is incongruent, this

difference is greatest for the endpoint words that are acoustically unambiguous (Connine, 1987). The difference in response times likely represents the cost of overcoming the prediction of the incongruent sentence context to select the clear example of the target word.

3.1.2 Signal degradation, aging, and context effects in listeners with CIs

Listeners with CIs may rely on context more than their peers with normal hearing (NH) (e.g., Amichetti et al., 2018; Dingemanse & Goedegebure, 2019; Winn, 2016). These prior studies showed a larger context benefit for listeners with CIs than for listeners with NH, using highly vs. minimally predictive contexts during an identification task for noise-gated words. In these word-onset gating tasks, the target words were presented with an initial gate size of 50 ms (i.e., all but the initial 50 ms of the target word was obscured by noise). After each presentation, the participant was asked to identify the target word. If they answered incorrectly, the same stimulus was presented, but the gate size was increased by 50 ms (i.e., 100 ms of the word onset were presented before the remainder of the word was obscured by noise). The gate sizes continued to be increased in 50-ms steps until either the word was correctly identified, or the entire word was presented without noise. The difference in the size of the gate (in ms) needed to identify the same target words presented in a highly predictive context compared to a minimally predictive sentence context was called the context benefit. Both groups performed very well with highly predictive context sentences. The larger context benefit for listeners with CIs is attributed to their poorer performance on

the minimally predictive context sentences compared to the listeners with NH. When listener groups were matched on their baseline performance (sentences without context) and age, similar context benefits for noise-gated words were found across listener groups (O'Neill et al., 2021). In a similar study with groups of acoustic hearing (AH) listeners who had different amounts of hearing loss, when the speech signal was equally audible, differences in the effect of context by age group were minimal (Dubno et al., 2000; Humes & Dubno, 2010).

3.1.3 The effect of sentence position

Studies on the use of sentence context have traditionally focused on the effect of context on the identification of a sentence-final word (e.g., Amichetti et al., 2018; Boothroyd & Nittrouer, 1988; Kalikow et al., 1977; Miller et al., 1951; Pichora-Fuller, 2008). Less is known, however, about how listeners with CIs use sentence context when the target words are at the beginning of sentences in which subsequent words could provide cues that would allow a listener to infer the missing or unclear word. In fact, there are no previous studies that have compared the effects of context for preceding vs. subsequent semantic context.

When processing a word in isolation, listeners with CIs have shown a tendency to delay committing to an identification response for the word as it unfolds over time compared to listeners with normal hearing (McMurray et al., 2017, 2019; F. Smith & McMurray, 2018). When a word near the beginning of a sentence is missing or unclear, listeners with CIs may decide to use this strategy, sometimes referred to as "wait-and-

see," to wait until they hear more of the subsequent sentence before committing to an identification response of the missing or unclear word. It is unknown whether this strategy, if used in sentences, would interact with age to impact the effect of semantic context on speech understanding.

3.1.4 Research questions and hypotheses

This study tested adults with CIs on a phoneme categorization task. The target word was presented in background noise at the beginning or end of sentences that were presented in quiet. The sentences contained congruent, incongruent, or neutral semantic context cues. The effect of context was calculated as the difference between target word categorizations in a neutral sentence context and categorizations of the same target words in predictive sentence contexts. There were three main research questions: (1) Does the age of the CI listener interact with location of the target word to change phoneme categorization? (2) Is working memory ability inversely related to the size of the context effect, especially when the target word is at the beginning of the sentence? and (3) Do listeners with CIs display a greater context effect than listeners with AH?

Older listeners may be contending with age-related cognitive changes that could interact with the loss of hearing ability to increase the effort needed to understand speech (e.g., Degeest et al., 2015; Lemke & Besser, 2016; Phillips, 2016; Tun et al., 2009). Given these potential increases in listening effort compared to younger listeners, older listeners may form strategies, such as always interpreting a word as congruent with the sentence context, that differ from those of younger listeners with different

experiences. The hypothesis for the first research question, therefore, was that older listeners would show larger context effects overall compared to younger listeners. Furthermore, the size of the context effect for older listeners would not be dependent on the location of the target word, while younger listeners would show larger context effects when the target word was at the sentence-final position compared to when it was at the sentence-initial position.

An additional research sub-question (Question 1a) is: Does an individual's speed of lexical access (i.e., processing speed) or inhibition ability predict the effect of semantic context on their categorization of target words? An individual's processing speed could affect the number and quality of the predictions they make based on degraded acoustic information alone. In addition, an individual's ability to inhibit the predictions from incongruent semantic context in favor of only using the acoustic cues to identify a target word could affect an individual's final identification of that target word. The hypothesis was that a measure of processing speed, as well as a measure of inhibition ability, would be predictive of an individual's context effect, especially when the target word was in the sentence-initial location. Higher performance on these cognitive tasks would likely be predictive of smaller context effects, because participants with faster processing and greater inhibition may be able to exclude confusing information and perform well with just the acoustic cues of the target words (Mattingly et al., 2018; Schubotz et al., 2021).

When an unclear word occurs at the beginning of a sentence, the potential interpretations of that word must be held in memory, and subsequent context must be

used to retroactively decide on the target word's identity. Assuming that CI processors can provide sufficient access to the acoustics of the target word to provide clues to the target word's identity, this suggests that a listener with better working memory may retain more acoustic cues to the identity of a sentence-initial target word than a listener with poorer working memory. A listener with poorer working memory might forget the acoustic cues of the early target word, and would likely be swayed by subsequent context cues to a greater degree than a listener who is able to hold the acoustic cues of the target word in their working memory. When the target word occurs at the end of the sentence, listeners will have access to the prediction formed by the semantic context as well as the acoustic cues to the target word's identity. The hypothesis for the second research question, therefore, was that better working memory ability would be inversely related to the size of the effects when the target word was at the beginning of the sentence. When the target word was at the end of the sentence, similar context effects would be expected for all levels of working memory ability among listeners with CIs.

The daily experience that listeners with CIs have with degraded speech would likely predispose them toward a reliance on context. The hypothesis for the third research question, therefore, was that listeners with CIs would exhibit a larger context effect than listeners with AH. It was also possible that the location of the target word would interact with the listener group if listeners with CIs employed a "wait-and-see" strategy (as described in: Farris-Trimble et al., 2014; McMurray et al., 2017, 2019), unlike listeners with AH. In this case, listeners with CIs might show greater context

effects when the target word was at the sentence-initial position, while listeners with AH might show greater context effects when the target word was at the sentence-final position.

3.2 Method

3.2.1 Participants

Thirty-two (32) adult CI participants (ages 21-86 years old, mean = 56.5 years) participated in this study. Most had post-lingual-onset hearing loss (i.e., hearing loss onset at age 5 or older). Five participants had pre-lingual onset of profound hearing loss, and two others had hearing loss prelingually, but used hearing aids to learn language. Participants had one or more CIs from FDA-approved manufacturers. They had at least one year of experience with their CI(s). They used their default everyday programs on their own processors during testing. If participants reported residual acoustic hearing, thresholds in that ear were measured. If acoustic thresholds were 60 dB HL or better at any audiometric frequency (re: ANSI, 2018), participants wore an earplug in that ear. The use of hearing aids during the task was not allowed. Only two participants needed to use an earplug (CI02 and CI04).

The participants of the study described in the previous chapter (who had AH) were used as the control group for comparisons across listener groups. All participants of both listener groups were native speakers of English, based on self-report, to avoid variance in speech understanding performance from diverse language backgrounds. In

addition, all 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.

Table 3-I. Participant information: including Duration of Severe-to-Profound Hearing Loss (DoSPHL), duration of CI use, CI processor(s), and Onset of Severe-Profound Hearing Loss with respect to language acquisition (pre- or post-lingual). The processors' brand names are indicated with "AB" for Advanced Bionics or "C" for Cochlear.

Participant	Age	Gender	DoSPHL (yrs)	Duration of CI use (yrs)	CI processor(s)	Pre-/Post-Lingual Onset
CI01	22	F	7	15	C Nucleus 7	Pre
CI02	22	F	12	8	C Nucleus 8	Post
CI03	25	F	21	4	C Nucleus 7	Post
CI04	25	M	13	7	C Nucleus 8	Post
CI05	31	M	1	15	C Nucleus 7	Post
CI06	33	F	2	32	C Nucleus 6	Pre
CI07	36	F	1	34	C Nucleus 7	Pre
CI08	38	M	2	36	C Nucleus 6	Pre
CI09	49	M	40	9	AB Naida	Pre
CI10	50	M	1	22	C Nucleus 7	Post
CI11	52	M	3	9	C Nucleus 7	Post
CI12	54	F	<1	15	C Nucleus 6	Post
CI13	55	F	21	13	C Kanso 2	Post
CI14	55	F	3	13	C Nucleus 8	Post
CI15	60	F	1	17	C Nucleus 6	Post
CI16	61	M	2	16	C Nucleus 8	Post
CI17	62	M	<1	13	C Nucleus 6	Post
CI18	63	M	11	19	AB Naida M	Post
CI18	65	F	4	16	C Nucleus 6	Post
CI20	65	M	13	8	C Nucleus 7	Post
CI21	66	F	19	13	C Nucleus 8	Post
CI22	66	F	3	18	C Nucleus 6	Post
CI23	69	M	17	16	AB Naida M	Post
CI24	70	F	13	12	C Nucleus 6	Post
CI25	70	F	2	14	C Nucleus6	Post
CI26	72	F	2	18	C Nucleus7	Post
CI27	75	M	4	1	AB Naida M	Post
CI28	77	M	<1	20	C Nucleus 6	Post
CI29	77	F	4	23	C Nucleus 7	Post
CI30	78	F	10	18	C Nucleus 7	Post
CI31	79	M	5	26	AB Naida Q	Post
CI32	86	F	<1	16	C Nucleus 6	Post

3.2.2 Stimuli

The stimuli were the same as those used in Study 1, except that the non-vocoded acoustic waveforms were presented to the participants with CIs. All other features were the same. The target words were taken from an acoustic continuum of four steps. The endpoint words were "deer" and "tear" (/dɪɹ/ and /tɪɹ/). The stimuli were low-pass filtered at 4000 Hz and had equal RMS intensities. Speech-shaped background noise was added to the target stimuli at +10, +5, 0, -5, and -10 dB SNR. This range of SNRs was chosen so that the degradation provided by the combination of listening through the CI processor 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 & Nittrouer, 1988). The sentences used for context are presented in Table 3-II.

Table 3-II. Sentences used to provide sentence-initial and sentence-final context.

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 ____."

3.2.3 Procedure

Participants provided written consent and were compensated for their time and travel. The Institutional Review Board of the University of Maryland approved the experimental procedures. Participants were seated comfortably in a sound-attenuating

booth. Auditory stimuli (context words and target sentences) were presented over a single loudspeaker (Yamaha HS5, Buena Park, CA, USA) approximately one meter in front of the participant's chair at 0 degrees azimuth. The stimuli were calibrated in the sound field to be presented at 70 dB A using a sound level meter (Brüel & Kjær Type 2250, Nærum, 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. Stimuli were presented and responses were recorded using PsychoPy software v. 2022.1.4 (Open Science Tools Ltd., Nottingham, UK).

3.2.3.1 Practice

The visual stimuli were presented on a 23-inch computer monitor located 0.75 meters from the participant's chair. Before the main experiment, participants completed a short practice session. The purpose of the practice session was to provide familiarity with the context sentences, establish baseline response times, ensure that participants could correctly identify the end-point target words with greater than 90% accuracy, and introduce participants to the task. The practice session followed the same procedures and format as described in Study 1. One participant did not meet the 90% criterion during the practice session (instead scoring 83%), but this fact was not noticed at the time of testing. Three other participants did not meet the 90% criterion during their first attempt and repeated the practice until the criterion was met.

3.2.3.2 Main Experiment

The experiment had the same six conditions for the continuum as the previous experiment: the target words presented at the beginning or end of a sentence that did not provide predictive context (neutral context), the target words presented at the beginning or end of a sentence that predicted one end-point of the continuum, and the target words presented at the beginning or end of a sentence that predicted the other end-point of the continuum. Each of the four target words was presented 10 times in the five different signal-to-noise ratios (SNRs) in the three types of context sentences and in the two sentence positions. Only the target words were presented in background noise. The context sentences were presented in quiet.

The target words, SNRs, and conditions were presented in random order for a total of 10 blocks of 120 trials (10 repetitions × 4 target stimuli along the VOT continuum from "deer" to "tear" × 5 SNRs × 6 conditions = 1200 trials in grand total). The participants categorized the words using a mouse to select a picture on the computer screen corresponding to one of the target words. The words of the target sentences were not presented on the screen during the experiment. Participants were encouraged to take a break between any of the blocks of 120 trials. Each block took approximately 15 minutes to complete. Performance was measured in two ways: 1) a percent "deer" response for each step along the continuum and 2) response times.

3.2.3.3 Cognitive Tests

During the breaks between blocks of the experiment or after the last block, participants completed a measure of verbal processing speed and inhibition, the Stroop task (Stroop, 1935), and a measure of working memory, the NIH Toolbox List Sorting Task (Tulsky et al., 2014). The same procedures described in Study 1 were followed for the administration and scoring of these tasks in Study 2. The scores from these measurements were used in the analysis as predictors of the effect of context on speech recognition performance.

3.2.4 Analysis

Data analysis followed the same procedures described previously. 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) and the *buildmer* package version 2.11 (Voeten, 2023). The "deer" categorizations from each participant in each condition were modeled with a logistic mixed-effect model (binomial distribution) containing fixed effects of age, VOT, context type (neutral, deer, or tear), sentence position (initial or final), SNR (5 levels), and their interactions. Age, VOT, and SNR were treated as continuous variables. All three variables were standardized such that the mean of each variable (mean age = 56.5 yrs, mean VOT = 56 ms, mean SNR = 0 dB SNR) was zero (with a standard deviation of one) and was the reference value in the model. The neutral sentence context was used as the reference value for context

type, and the sentence-final position were used as the reference value for sentence position. The random effects included a random intercept of participant with random slopes of SNR, sentence position, and VOT for each participant. The maximal model with all fixed effects interacting and maximal random effects was input into the *buildmer* package to determine the best-fitting model. The maximal formula was used as input with a binomial distribution for the dependent variable and the default stepwise elimination procedure. *Buildmer* first ranks terms by their contribution to the model while ensuring that the model converges, and then performs backward elimination using likelihood-ratio testing.

A separate linear MEM (Gaussian distribution) was used to analyze log-transformed response times with the same random effects as the previous analysis. The response times were log-transformed in order to obtain a better approximate of a Gaussian distribution for the model. The fixed effects of SNR, sentence position, congruency, age, and acoustic ambiguity were used in the model. Given that the acoustic ambiguity of some of the target words could potentially affect response times, both endpoint words of the VOT continuum were categorized as acoustically unambiguous. The intermediate target words were categorized as acoustically ambiguous. 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 acoustically unambiguous target words, the incongruent sentence context, and the sentence-final target word

position were used as the reference conditions for the acoustic ambiguity, congruency and sentence position variables in the model. Random effects of participant were included in the model to account for differing baseline response times for each individual. Random slopes of target word position, SNR, and Ambiguity were also included.

A final linear MEM (Gaussian distribution) was used to analyze the contribution of other predictive factors, such as processing speed, inhibition ability, and working memory ability, to the log-transformed context effects experienced by the participants. Context effects were calculated as described in Study 1, and used as the dependent variable in this model.

Post-hoc testing included re-leveling the Context categorical variable such that the deer-predictive context became the new reference level. This allowed a comparison between performance in the deer-predictive sentences and the tear-predictive sentences. The residuals of each model were checked to ensure there were no major violations of normality or homoscedasticity.

3.3 Results

3.3.1 Categorization Performance: "Deer" Responses

Results are first presented as the average performance for all participants at each continuum step, at each SNR, and in each sentence location (Figure 3-1). Each point represents the average percentage of "deer" responses at each of the four steps along

the VOT continuum. The left-most point in each subplot represents performance on the "deer" endpoint target word, while the right-most point in each subplot is the performance on the "tear" endpoint target word. The top row shows performance when the target words were presented at the beginning of the sentences. The lower row shows performance when the target words were presented at the end of the sentences. The various SNRs are presented in columns from the most challenging SNR on the left to the easiest SNR on the right. The red circles represent responses when the context sentence predicted a "deer" target word. The blue squares represent responses when the context sentence predicted a "tear" target word. The black triangles represent responses when the context sentence was neutral and did not predict either target word.

If the red line representing the phoneme categorization performance in the deer-biased context sentences is above the black neutral line, then a context effect can be observed. Similarly, if the blue line representing the phoneme categorization performance in the tear-biased context sentences is below the black neutral line, this also is considered a context effect. In other words, the addition of tear-biased context caused the participants to identify the word as "tear" more often than in the neutral context sentences.

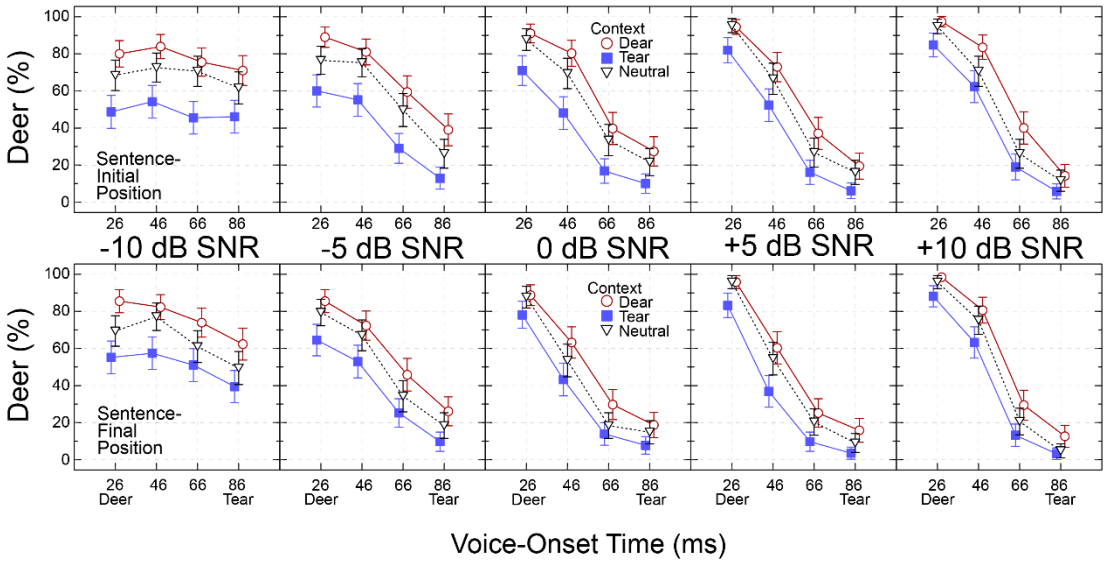

Figure 3-1. Performance of 32 CI participants when target words were presented in the sentence-initial (top) and sentence-final (bottom) positions at five signal-to-noise ratios (SNRs; columns). Responses are shown as red open circles when the sentence predicted "deer", as blue solid squares when the sentence predicted "tear," and as open black triangles when the sentence context was neutral. Error bars are ± 1 standard error.

An examination of Figure 3-1 reveals that when the target word is almost completely masked by background noise (-10 dB SNR, left column), participants have difficulty performing the phoneme categorization task (neither endpoint word is identified consistently). Indeed, most participants reported that they were guessing in the -10 dB SNR condition. As the SNR becomes more favorable, moving to the right in the figure, participants' identification of the endpoint target words improves. The effect of sentence position can be seen in a comparison of the upper and lower rows. Context effects, comparing the red and blue lines to the black dotted line, appear to be somewhat greater when the target words were presented at the beginning of the context sentences.

The output of the model that was fit to the trial level data is shown in Table 3-III. The model shows that the experimental manipulations of VOT and SNR significantly affected the probability of a "deer" response ($p < 0.001$ for both). There was a significant interaction between VOT and SNR (VOT × SNR, $p < 0.001$). The difference in performance in the -10 dB SNR condition compared to the other SNR conditions seems to contribute the most to this interaction (see leftmost column in Figure 3-1). Since most participants reported that they were guessing at -10 dB SNR, this interaction likely reflects the fact that participants were unable to hear the acoustic target words in that condition. The model also shows that VOT, SNR, and sentence context interact significantly with target word position. Overall, there were more "deer" responses for target words presented at the beginning of the sentence compared to those presented at the end of the sentence. The difference in "deer" responses between the two sentence positions is smaller for the more difficult SNR conditions compared to the easier SNR conditions [Fig. 3-2(A); SNR × Position-Initial, $p < 0.001$], smaller for the unambiguous "deer" endpoint word compared to the other words [Fig. 3-2(B); VOT × Position-Initial, $p < 0.001$], and smaller for words presented in the tear-predictive contexts [Fig. 3-2(C); Context-Tear × Position-Initial, $p < 0.001$]. Post-hoc testing with re-leveling revealed that there was a significant difference in the effect of target word position between the deer-predictive sentences and the tear-predictive sentences (Context-Tear × Position-Initial with Deer context as reference, $p < 0.001$), but not between the neutral sentences and the deer-predictive sentences (Context-Deer × Position-Initial with neutral context as reference, $p > 0.05$).

Table 3-III. Model output for a logistic mixed effects model. The dependent variable is the binary "deer" or "tear" choice, and the independent variables are VOT, Context, SNR, Age, and Sentence Position. Reference values: Neutral Context, Final Position, average SNR, VOT, and Age.

Formula: Deer-or-Tear ~ VOT × Context × SNR × Age + Sentence Position × VOT × SNR + Context × Sentence Position + (VOT × SNR + Context + Sentence Position | participant)

	Deer or Tear				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	-0.234	0.124	-1.884	0.060	
VOT	-1.997	0.181	-11.052	**<0.001**	***
Context-Deer	0.605	0.151	3.996	**<0.001**	***
Context-Tear	-0.634	0.190	-3.327	**<0.001**	***
SNR	-0.523	0.123	-4.264	**<0.001**	***
Age	-0.095	0.106	-0.897	0.370	
Position-Initial	0.405	0.088	4.626	**<0.001**	***
VOT × Context-Deer	0.027	0.043	0.624	0.533	
VOT × Context-Tear	0.104	0.042	2.458	**0.014**	*
VOT × SNR	-1.128	0.109	-10.320	**<0.001**	***
Context-Deer × SNR	-0.021	0.036	-0.604	0.546	
Context-Tear × SNR	0.041	0.035	1.179	0.238	
VOT × Age	0.328	0.178	1.839	0.066	
Context-Deer × Age	0.243	0.136	1.786	0.074	
Context-Tear × Age	-0.184	0.179	-1.024	0.306	
SNR × Age	-0.009	0.122	-0.075	0.941	
VOT × Position-Initial	0.281	0.034	8.357	**<0.001**	***
SNR × Position-Initial	0.114	0.028	4.016	**<0.001**	***
Context-Deer × Position-Initial	0.012	0.068	0.183	0.855	
Context-Tear × Position-Initial	-0.295	0.066	-4.442	**<0.001**	***
VOT × Context-Deer × SNR	0.072	0.042	1.732	0.083	
VOT × Context-Tear × SNR	0.027	0.041	0.666	0.506	
VOT × Context-Deer × Age	-0.023	0.044	-0.522	0.602	
VOT × Context-Tear × Age	-0.077	0.044	-1.756	0.079	
VOT × SNR × Age	0.176	0.108	1.625	0.104	
Context-Deer × SNR × Age	-0.049	0.036	-1.377	0.168	
Context-Tear × SNR × Age	0.023	0.035	0.642	0.521	
VOT × SNR × Position-Initial	0.049	0.033	1.503	0.133	
VOT × Context-Deer × SNR × Age	0.015	0.042	0.349	0.727	
VOT × Context-Tear × SNR × Age	-0.083	0.042	-2.004	**0.045**	*

Random Effects:	*Variance*	*St. Dev.*	*Corr.*					
By-Participant intercept	0.457	0.676						
By-Participant VOT slope	1.006	1.003	0.66					
By-Participant SNR slope	0.457	0.676	0.57	0.69				
By-Participant Context-Deer slope	0.654	0.809	0.38	0.22	0.15			
By-Participant Context-Tear slope	1.093	1.046	-0.43	-0.40	-0.17	-0.91		
By-Participant Position-Initial slope	0.175	0.418	-0.52	-0.16	-0.06	-0.43	0.33	
By-Participant SNR × VOT slope	0.343	0.586	0.64	0.97	0.69	0.15	-0.30	-0.08
Observations	38400							

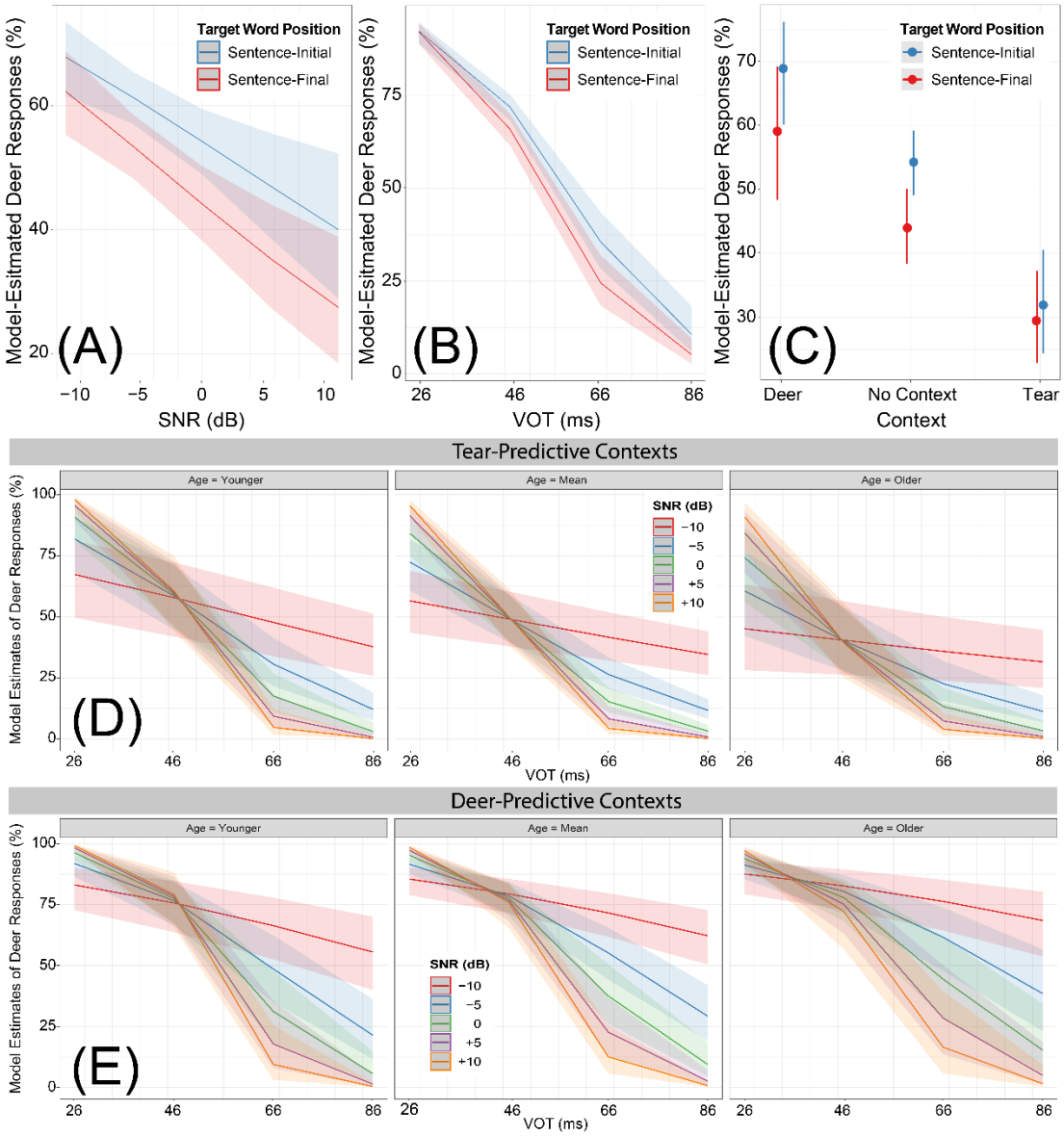

Figure 3-2. Panel (A) shows the interaction between sentence position and signal-to-noise ratio (SNR). Blue represents the model's estimates of deer responses when the target word is in the sentence-initial position, and red represents the same for the sentence-final position. Panel (B) shows the interaction between sentence position and voice-onset time (VOT) on the model's estimates of deer responses. Panel (C) shows the interaction between sentence position and semantic context on the average percentage of "deer" responses collapsed across VOTs. Panel (D) shows the interaction between VOT, SNR, and age when the target word is presented in a tear-predictive context sentence. Panel (E) shows the interaction between VOT, SNR, and age when the target word is presented in a deer-predictive context sentence. Error bars and shaded regions represent the 95% confidence interval.

There was a significant four-way interaction (VOT × Context-Tear × SNR × Age, $p = 0.045$) such that the difference in the number of deer responses by SNR is greater for VOTs of 66 and 86 ms compared to the other target words for older participants when the target word is presented in deer-predictive contexts and greater for target words with 26-ms VOT compared to the other target words for older participants when the target word is presented in tear-predictive contexts [see Fig. 3-2 (D) and (E)].

3.3.2 Cognitive Measures and the Context Effect

Working memory was the only cognitive measure that had a significant impact on the derived context effect. Table 3-IV shows the raw and derived scores on the cognitive tests for 28 of the participants who used CIs. Three participants did not complete the cognitive tests during their visits and were unable to return. One participant was older than the NIH Toolbox norms and therefore their age-corrected working memory score could not be calculated. Despite using the age-corrected norms for the working memory task, scores for the working memory task and the derived inhibition scores were significantly correlated with age using Pearson's product-moment correlation [working memory: $r = 0.596$; $t(26) = 3.786$, $p < 0.001$; inhibition: $r = -0.647$; $t(26) = -4.327$, $p < 0.001$]. The derived measure of processing speed was not significantly correlated with age for this group [processing speed: $r = -0.087$; $t(26) = -0.444$, $p > 0.05$].

*Table 3-IV. Scores on cognitive tests, both raw and derived. Participant information includes the duration of severe-to-profound hearing loss (DoSPHL) and age. Inhibition was calculated as Color-Word score – (Color score * Word score)/(Color score + Word score). Processing speed was calculated as (Word score – Color score) * (-1).*

Participant	Age (yrs.)	Standardized Working Memory Age-Corrected	Word Score	Color Score	Color-Word Score	Inhibition (Derived)	Processing Speed (Derived)
CI01	21	102	89	79	49	7.15	-10
CI02	22	97	108	83	60	13.07	-25
CI03	25	92	112	93	64	13.19	-19
CI04	25	70	80	75	50	11.29	-5
CI05	31	88	101	88	47	-0.03	-13
CI06	33	94	118	85	60	10.59	-33
CI07	36	78	106	74	50	6.42	-32
CI08	38	100	100	82	45	-0.05	-18
CI09	49	82	107	75	37	-7.09	-32
CI10	50	94	97	75	25	-17.3	-22
CI11	52	106	108	66	39	-1.97	-42
CI12	54	107	83	59	34	-0.49	-24
CI13	55	102	87	79	43	1.60	-8
CI14	60	109	90	75	48	7.09	-15
CI15	61	87	79	51	29	-1.99	-28
CI16	63	134	127	84	60	9.44	-43
CI17	65	113	87	65	30	-7.20	-22
CI18	65	124	85	75	33	-6.84	-10
CI18	66	113	91	72	36	-4.20	-19
CI20	66	118	100	76	47	3.82	-24
CI21	69	115	117	82	45	-3.21	-35
CI22	70	104	94	58	23	-12.87	-36
CI23	72	116	75	65	28	-6.82	-10
CI24	75	142	56	54	34	6.51	-2
CI25	77	96	79	52	27	-4.36	-27
CI26	77	114	107	94	37	-13.04	-13
CI27	78	120	91	59	26	-9.79	-32
CI28	79	92	87	70	21	-17.79	-17

As in the previous study, the data can be converted into a single number demonstrating the effect of context for an individual at each SNR and in each sentence position. The context effect was calculated by adding the difference between the number of "deer" responses in the deer context and the neutral context to the difference between the number of "deer" responses in the neutral context and the tear context.

This method was described in more detail in Study 1. The same formulae were used to calculate context effects for listeners with CIs.

The context effects are plotted in Figure 3-3. This figure shows the average context effects on the performance of participants with CIs at all five SNRs in separate panels for sentence-initial and sentence-final target words. While age was treated as a continuous variable in the analysis, the participants are separated into three age groups for purposes of visualizing the age effects. These age groups do not directly correspond to traditional definitions of middle-age or old-age – rather, they are a way to evenly divide the participant pool. As a result, some of the effects of aging may be obscured by the chosen groupings. The older participants (those > 65 years old) demonstrated relatively large context effects that appear to gradually decrease with increasing SNR and are only different across sentence positions in the -5 and 0 dB SNR conditions [see red open circles in Figure 3-3 (C)]. The middle-aged participants (those 45-65 years old) showed context effects that were similar to those shown by the older participants [see Figure 3-3 (A) and (B)]. The middle-aged participants also had the largest difference in context effects based on sentence position with sentence-initial context effects appearing to be greater than sentence-final context effects – especially at -5, 0, and +5 dB SNRs [see filled blue triangles in Figure 3-3 (C)]. The younger participants, in stark contrast, appear to experience a relatively small context effect across all SNRs [see Figure 3-3 (A) and (B)]. The context effects appear to be larger for sentence-initial target words than for sentence-final target words in SNR conditions ≤ 0 dB [see open cyan squares in Figure 3-3 (C)]. The differences between the middle-aged and older

participant groups should be interpreted with caution because some of the participants who are plotted in the middle-aged group are old enough to be classified as "older" by many definitions.

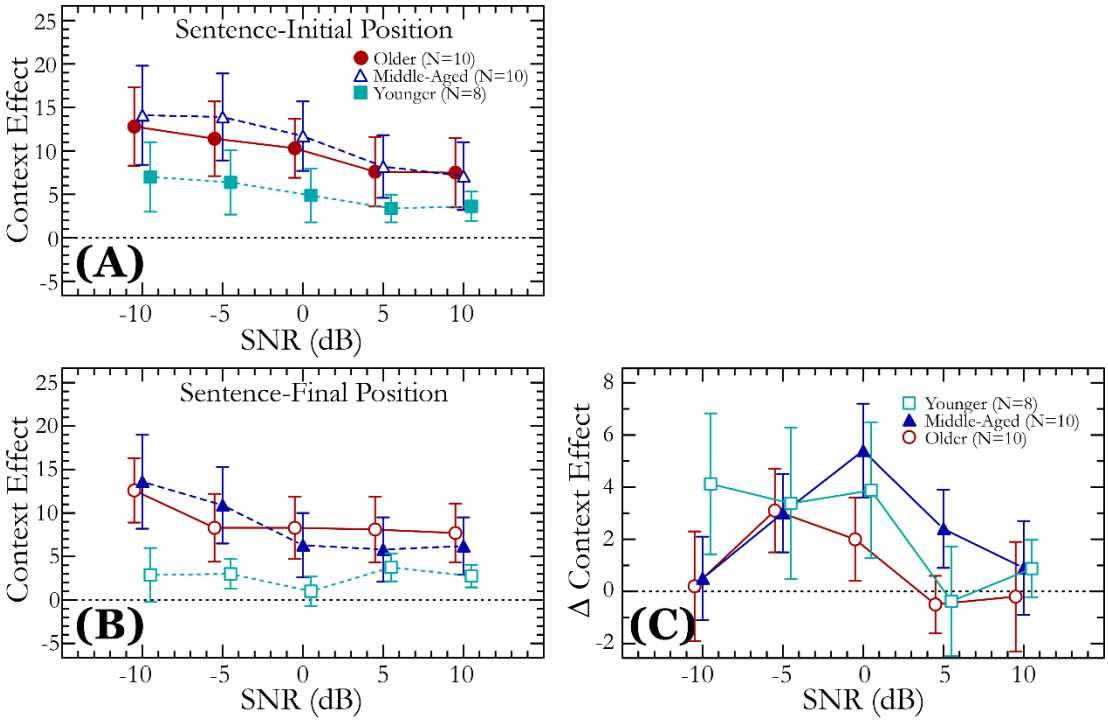

Figure 3-3. The plots on the left show the average context effect compared to performance in the neutral context sentences of older participants (>65 years; red circles), middle-aged participants (40-65 years; dark blue triangles), and younger participants (<40 years; cyan squares) with CIs. This effect is shown across SNRs for target words presented at the beginning of context sentences (A) and at the end of context sentences (B). Panel (C)shows the difference between the context effect an individual experienced for target words at different positions (sentence-initial – sentence-final). Error bars represent ± 1 standard error of the mean.

The context effect for these 28 participants was log-transformed to reduce the skewness of the distribution and used as the dependent variable in a linear MEM (Gaussian distribution). The initial fixed effects included sentence position, SNR, age, duration of severe-to-profound hearing loss, inhibition, processing speed, and working

memory. The standardized working memory score was centered at zero to better match the scale of the other predictors. The *buildmer* package 2.11 (Voeten, 2023) eliminated measures of inhibition, processing speed, and duration of severe-profound hearing loss as significant predictors of the context effect. However, the measure of working memory was retained in the final model, as shown below (Table 3-V).

Table 3-V. Model output for a linear mixed effects model fit to log-transformed context effect. Reference values: Final Sentence Position, average SNR, Age, and Working Memory

Formula: Context Effect ~ SNR * Sentence Position * (Working Memory + Age) + (SNR + Sentence Position | participant)

	Context Effect				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	2.672	0.080	33.307	**<0.001**	*******
SNR	-0.050	0.042	-1.190	0.242	
Position-Initial	0.086	0.047	1.821	0.081	
Working Memory	-0.192	0.100	-1.922	0.066	
Age	0.261	0.100	2.611	**0.015**	*****
SNR × Position-Initial	-0.012	0.037	-0.318	0.751	
SNR × Working Memory	-0.001	0.053	-0.020	0.984	
SNR × Age	-0.034	0.053	-0.655	0.516	
Position-Initial × Working Memory	0.019	0.059	0.314	0.756	
Position-Initial × Age	-0.057	0.059	-0.960	0.346	
SNR × Position-Initial × Working Memory	-0.126	0.046	-2.726	**0.007**	******
SNR × Position-Initial × Age	0.108	0.046	2.345	**0.020**	*****
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
By-Participant effects	0.161	0.402			
By-Participant SNR slopes	0.031	0.177	-0.67		
By-Participant Sentence Position slopes	0.025	0.157	0.35	-0.60	
Observations	280				

The size of the context effect for participants was generally greater at SNRs ≤ 0 dB and greater for sentence-initial target words at -5, 0, and +5 dB SNR, especially for middle-aged participants [see Figure 3-3 (C); SNR × Position-Initial × Age, p = 0.02]. There was also a significant three-way interaction involving the age-corrected standardized scores on the working memory task. Because age and working memory

were so highly correlated (i.e., older participants had higher working memory scores), this interaction could be more reflective of the age of the participants than the influence of working memory on this task. Overall, age interacted with the target word's sentence position and SNR to determine the size of the context effect.

3.3.3 Response Time Measures

Response times were collected for all trials. Response times greater than 10 seconds and less than 0.1 second were removed from the analysis because response times of less than 0.1 second are physically impossible after a decision, while response times greater than 10 seconds are likely to represent lapses in attention or other factors not related to decision making. This left 38,278 response times to analyze. The remaining values were log-transformed to reduce the skewness of the distribution. A category for acoustic ambiguity was created using the steps along the VOT continuum. The two endpoint steps were categorized as unambiguous (clearly "deer" and "tear"), while the two middle steps were categorized as acoustically ambiguous (could be interpreted as either "deer" or "tear"). A linear MEM (Gaussian distribution) was fitted to the response time data with SNR, Acoustic Ambiguity, Age, Sentence Position, and Congruency as fixed effects (Table 3-VI).

Table 3-VI. Model output for a linear mixed-effects (LME) model fit to log-transformed response time data. Reference values: Final Sentence Position, acoustically Unambiguous Target Words, Incongruent Sentence Contexts (a combination of the neutral contexts and opposite contexts), and average SNR and Age.

Formula: Response Time ~ SNR * Sentence Position * (Target Word Acoustic Ambiguity + Age) + Ambiguity * Age * Context Congruency + (1 + Position + SNR + Ambiguity | participant)

	Response Time				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	0.037	0.070	0.532	0.599	
SNR	-0.135	0.016	-8.424	**<0.001**	***
Position-Initial	-0.183	0.027	-6.875	**<0.001**	***
TargetWords-Ambiguous	0.062	0.012	5.318	**<0.001**	***
Age	0.155	0.070	2.215	**0.035**	*
Context-Congruent	-0.042	0.008	-5.311	**<0.001**	***
SNR × Position-Initial	0.038	0.007	5.142	**<0.001**	***
SNR × TargetWords-Ambiguous	0.070	0.007	9.502	**<0.001**	***
SNR × Age	0.002	0.016	0.131	0.897	
Position-Initial × TargetWords-Ambiguous	-0.054	0.010	-5.410	**<0.001**	***
Position-Initial × Age	0.003	0.026	0.104	0.918	
TargetWords-Ambiguous × Age	-0.011	0.010	-1.052	0.299	
TargetWords-Ambiguous × Context-Congruent	-0.011	0.011	-1.017	0.309	
Age × Context-Congruent	-0.010	0.008	-1.330	0.184	
SNR × Position-Initial × TargetWords-Ambiguous	-0.044	0.010	-4.244	**<0.001**	***
SNR × Position-Initial × Age	0.000	0.005	0.054	0.957	
TargetWords-Ambiguous × Age × Context-Congruent	0.020	0.011	1.792	0.073	
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
By-Participant effects	0.155	0.394			
By-Participant Sentence Position slopes	0.021	0.145	-0.41		
By-Participant SNR slopes	0.007	0.086	-0.42	0.37	
By-Participant Target Word Ambiguity slopes	0.002	0.047	0.02	0.12	-0.04
Observations	38278				

The model shows that response times were significantly predicted by interactions between Sentence Position, Target Word Acoustic Ambiguity, and SNR. In general, response times were shorter when the SNR was more favorable (SNR, $p < 0.001$), when the target word occurred at the beginning of the sentence (Position-Initial, $p < 0.001$), and when the context sentence was consistent with the acoustic target word (Context-Congruent, $p < 0.001$). Response times were generally longer (i.e., slower)

for ambiguous target words (TargetWords-Ambiguous, $p < 0.001$) and with increasing age (Age, $p = 0.035$). Response times to acoustically ambiguous target words appear to be longer than to acoustically unambiguous target words in the more favorable SNR conditions, mainly when the target word occurred at the end of the sentence [see Fig 3-4(A) and (B)] (SNR × Position-Initial × TargetWords-Ambiguous, $p < 0.001$).

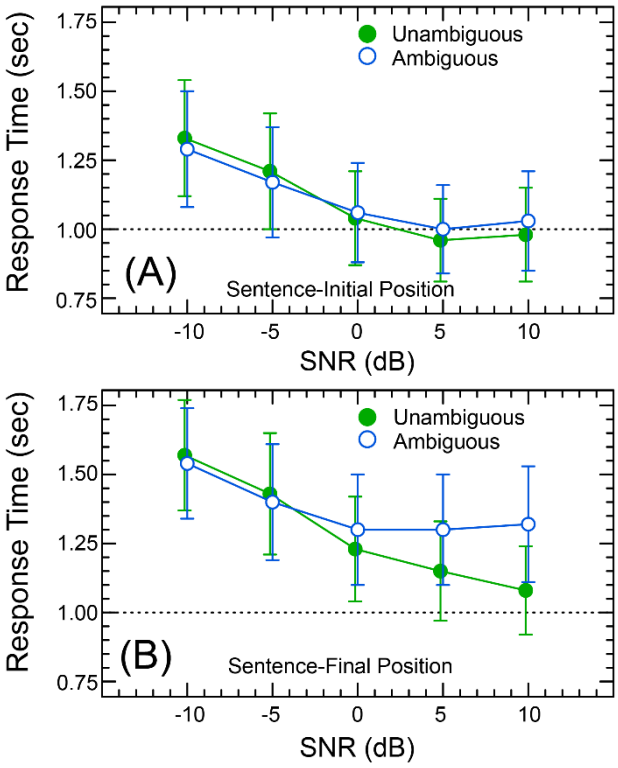

Figure 3-4. Panel (A) shows response times plotted as a function of signal-to-noise ratio (SNR) and acoustic ambiguity when the target word was in the sentence-initial position. Panel (B) shows response times plotted as a function of SNR and acoustic ambiguity when the target word was presented in the sentence-final position. Solid symbols represent unambiguous target word conditions, and open symbols represent ambiguous target word conditions. The dashed line at one second was included to aid comparison across plots. Error bars represent ± 1 standard error of the mean.

3.3.4 Comparison to Listeners with AH

As noted in the methods, the data from Study 1, which tested only adults with AH, were used as a "control group" to determine if participants who use CIs are differentially affected by SNR, target word position, and context compared to participants with AH who listened to a CI simulation. The context effect for all 63 participants (AH and CI) was log-transformed to reduce the skewness of the distribution and used as the dependent variable in a linear MEM with a Gaussian distribution. The initial fixed effects included sentence position, SNR, listener group, age, inhibition, processing speed, and working memory. The standardized working memory score was centered at zero to better match the scale of the other predictors. The *buildmer* package version 2.11 (Voeten, 2023) eliminated measures of inhibition and working memory as significant predictors of the context effect. However, age and the measure of processing speed were retained in the final model (Table 3-VII).

The average context effect for participants with CIs is greater than the average context effect for participants with AH across all SNRs (Group-CI, $p = 0.008$, Figure 3-5). The context effects shown for sentence-initial target words are larger than the context effects shown for sentence-final target words at SNRs ≤ 0 dB (SNR × Position-Initial, $p = 0.019$). Additionally, there is a three-way interaction showing that the difference in context effect between sentence-initial and sentence-final target word presentations is slightly larger for those with faster processing compared to slower processing — only in the most-difficult SNR conditions (SNR × Position-Initial ×

Speed, $p = 0.015$). In this data set that contains participants from both listener groups, processing speed is significantly correlated with age [$t(628) = 2.643$, $p = 0.008$]. Therefore, this significant interaction could reflect the effect of age rather than processing speed. The unique role that processing speed plays in this task, if any, remains unclear.

Table 3-VII. Model output for a linear mixed-effects (LME) model fit to the log-transformed context effect and compared across listener groups. Reference values: Final Sentence Position, AH listener Group, and average SNR, Processing Speed, and Age.

Formula: Context Effect ~ SNR * Sentence Position * Processing Speed + Group + Age + (SNR + Sentence Position | participant)

	Context Effect				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	2.479	0.057	43.306	**<0.001**	*******
SNR	-0.040	0.026	-1.516	0.133	
Position-Initial	0.040	0.034	1.186	0.240	
Speed	0.003	0.046	0.069	0.945	
Group-CI	0.213	0.078	2.740	**0.008**	******
Age	0.057	0.039	1.469	0.147	
SNR × Position-Initial	-0.056	0.024	-2.356	**0.019**	*****
SNR × Speed	-0.002	0.026	-0.085	0.932	
Position-Initial × Speed	-0.009	0.034	-0.254	0.800	
SNR × Position-Initial × Speed	-0.057	0.024	-2.434	**0.015**	*****
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
By-Participant effects	0.114	0.337			
By-Participant SNR slopes	0.026	0.160	-0.63		
By-Participant Sentence Position slopes	0.036	0.190	0.24	-0.63	
Observations	280				

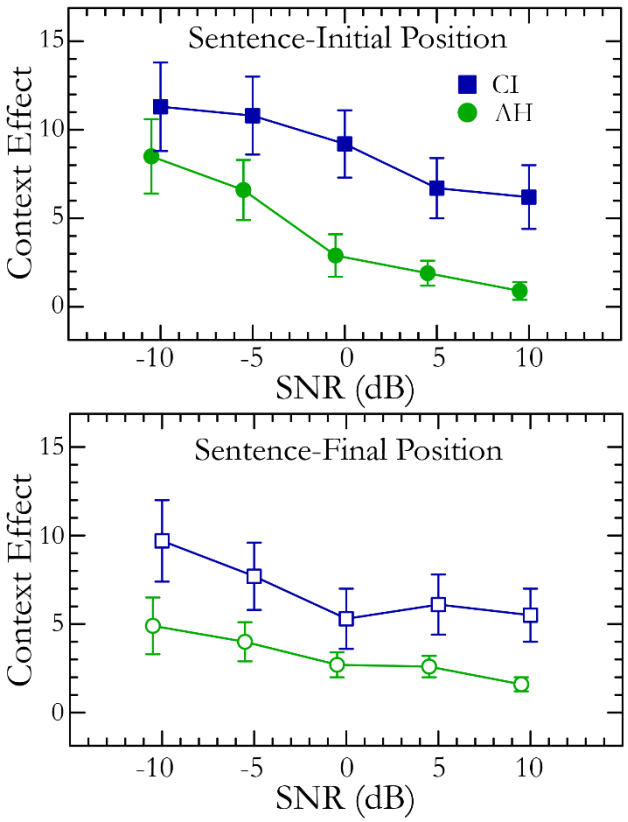

Figure 3-5. Average context effect compared to performance in the neutral context sentences of participants with AH (green circles) and participants with CIs (blue squares) across SNRs for target words presented at the beginning of context sentences (top) and at the end of context sentences (bottom). Error bars are ± 1 standard error of the mean.

3.4 Discussion

This study tested adults with CIs on a phoneme categorization task with the target word presented in background noise at the beginning or end of a sentence. The main goal was to determine if the participants' age interacted with the level of background noise, the target word's sentence position, and measures of various cognitive abilities to change how an individual perceived speech sounds from a VOT

continuum. Overall, background noise level and the target words' sentence position had the greatest effects on phoneme categorization.

The effects of background noise level (described as SNR in Figure 3-1) on phoneme categorization occur mainly in the -10 dB SNR condition, when the target words were largely inaudible to participants. The difference in the effect of context is greater in the sentence-initial position than in the sentence-final position and interacts with the target words, such that participants were slightly less sure of their categorization responses when the target word was presented in the sentence-initial position [Figure 3-2 (A)]. In both the neutral context sentences and the deer-predictive context sentences, target words presented in the sentence-initial position were more likely to be categorized as "deer" than when presented in the sentence-final position. The target word's sentence position did not appear to affect the categorization of words in the tear-predictive context sentences [Figure 3-2 (B)].

The response time analysis confirmed that participants' responses were slightly slower with increasing age and when the target words were ambiguous. Participants' responses were faster when the SNR was more favorable, when the target word occurred in the sentence-initial position, and when the context sentence was congruent with the target word (Table 3-VI and Figure 3-4). The clear effects of the experimental manipulations confirmed the validity of the stimuli for eliciting context effects.

3.4.1 Listeners with CIs

The hypothesis for the first research question, regarding the age of the listener, was that older listeners would show larger context effects overall compared to younger listeners. This hypothesis was confirmed in the analysis of context effects (Figure 3-3 and Table 3-IV). Older listeners who wear CIs showed larger context effects than younger listeners across all SNRs tested. A further hypothesis was that the sentence position of the target word would not alter the context effect experienced by older listeners, while younger listeners would show larger context effects when the target was at the sentence-final position compared to when it was at the sentence-initial position. Results showed that, on average, participants showed larger context effects for target words in the sentence-initial position rather than those in the sentence-final position [see points above the dotted line in Figure 3-3 (C)]. However, participants in the younger (< 40 years old) and middle (40-65 years old) age ranges showed the largest influence of the target word's sentence position depending on the level of background noise (SNR), while older listeners (> 65 years old) showed the least difference in context effect based on the target words' sentence position in the illustration of age effects shown in Figure 3-3 (C). Despite these results not supporting the hypothesis that older listeners would experience the same context effect for target words in sentence-initial and sentence-final positions, the findings do support the hypothesis that younger listeners are more affected by the target word's sentence position than older listeners. The hypothesis that participants with higher performance on the processing speed and

inhibition tasks would show smaller context effects, especially for sentence-initial target words was not supported. Both of these measures were significantly correlated with age in the current study and age was a significant predictor of the context effect. This correlation could have obscured the true effects of processing speed and inhibition ability on the use of context.

The hypothesis for the second research question, regarding the influence of working memory ability on context effects, was that better working memory would correlate with smaller context effects when the target word was at the beginning of the sentence, while working memory ability would not correlate with context effects when the target word was at the end of the sentence. The age-corrected scores on the working memory task were still highly correlated with age in this study. The correlation was positive, meaning that older participants demonstrated better working memory than younger participants. This could reflect that the highly educated and experienced older individuals who took part in the study truly had better working memory than the younger individuals, or it could merely indicate that older participants were more motivated to perform well on the working memory task than younger participants. Therefore, the significant interaction between working memory, SNR, and sentence position could partially reflect an age effect in addition to a working memory effect.

3.4.2 Comparison between listeners with CIs and those with AH

The hypothesis for the third research question, regarding the differences in context effect between listeners with CIs and those with AH, was that listeners with

CIs would exhibit a larger context effect than listeners with AH. The results support this hypothesis. Overall, listeners with CIs showed greater context effects than listeners with AH (Figure 3-5 and Table 3-VII). The current results confirm the findings of greater context effects for listeners with CIs than for listeners with AH while using a different task than used by previous studies (e.g., Amichetti et al., 2018; Dingemanse & Goedegebure, 2019; Winn, 2016).

These results appear to contrast with those of O'Neill et al. (2021), who found no difference in the size of the context effect for listeners with CIs and listeners with AH. The authors of that study argued for the importance of equating baseline speech understanding across listener groups. The current experiment was designed to roughly equate listener groups' access to the acoustic signal using background noise and signal processing so that these factors would be unlikely to contribute to differences in performance between those with CIs and those with AH. In addition, the current study required that all participants (those with CIs and those with AH) were able to identify the continuum endpoint target words with > 90% accuracy in the +10 dB SNR. A sub-analysis of performance in only the +10 dB SNR condition of the current study found that listeners with CIs still experienced significantly larger context effects than listeners with AH (Group-CI: $p = 0.014$). This was especially true for older listeners (Group-CI × Age: $p = 0.009$). It is possible that despite similar performance on the endpoint target words in the +10 dB SNR condition, the two listener groups' access to the acoustic cues was not as well-equated as in O'Neill et al.'s (2021) study. Therefore, their

conclusions of similar uses of context across listener groups when access to the acoustic signal is equated may still be valid.

The difference in findings could also be attributed to differences in the methods of determining and quantifying context effects. The current study evaluated performance on a phoneme categorization task with multiple repetitions of the same context sentences. O'Neill et al. (2021) compared word recognition in sentences from the Basic English Lexicon (BEL: Calandruccio & Smiljanic, 2012) speech corpus to word recognition in a nonsense version of the same corpus (O'Neill et al., 2020). The word-recognition task that O'Neill et al. (2021) employed is more similar to the struggle of every-day communication. It is possible that habits (e.g., relying on context cues) learned during daily experience listening with a CI affected the strategies displayed in the current study. Thus, even though the participants with CIs could likely perform at least some of the current task without relying on context, they relied on context cues in all parts of the task (maximizing the measured context effects), while the participants with AH did not.

In addition to an overall effect of listener group, it was hypothesized that the location of the target word in the sentence would differentially affect performance by the different listener groups, such that listeners with CIs would show greater context effects when the target word was at the sentence-initial position, while listeners with AH would show greater context effects when the target word was at the sentence-final position. The basis for this hypothesis was informed by the idea that listeners with CIs might use a "wait-and-see" strategy and listeners with AH might not (e.g., Farris-

Trimble et al., 2014; McMurray et al., 2017, 2019). The "wait-and-see" strategy is based on observations during visual world paradigm (VWP) eye-tracking studies where listeners are instructed to look at the object that represents the word they are hearing. In these VWP studies, it was observed that listeners with CIs often delay fixating their gaze on an object until more of the target word is presented, compared to listeners with AH. If listeners with CIs are waiting to interpret the sentence-initial target word until the sentence is completed, they may be more likely to interpret the word as consistent with the context, compared to listeners with AH who might decide on their interpretation of the target word before hearing the whole context sentence. There was no significant difference between the context effects shown by each listener group that was dependent on the location of the target word in the sentence, contrary to the hypothesis. Both listener groups showed a smaller context effect for sentence-final target words than for sentence-initial target words (Figure 3-5). This discrepancy might stem from the spectral degradation of the words and sentences presented to the listeners with AH. McMurray et al. (2017) provided evidence of the same "wait-and-see" strategy in listeners with AH when listening to highly degraded speech. It is possible that the speech in the current study was degraded enough to elicit this strategy in both listener groups.

3.4.3 Conclusion

The findings of this study strongly indicate that context plays a critical role in assisting listeners with CIs to identify ambiguous words in sentences. The target word's

sentence position affected performance such that participants showed larger context effects for sentence-initial target words compared to sentence-final target words. This finding is contrary to the previous literature in listeners with AH and hearing loss, however, it is consistent with the findings of Study 1. Overall, participants with CIs demonstrated similar patterns of performance compared to participants with AH. The effects of age may be amplified in the participants with CIs, because there are fewer redundancies in the speech signal that could assist in acoustic categorization compared to the speech signal heard by participants with AH.

The ability to test a larger sample of listeners with CIs could result in more clarity regarding the effects of age and cognitive abilities on the use of context in everyday listening. It is possible that the levels of background noise used in this experiment may have overestimated the difficulty of everyday listening environments for listeners with CIs. Future research should expand the testing groups and investigate additional cognitive factors that may impact the use of context. There should also be an effort to measure how individuals use semantic context in more realistic environments. This could take the form of using conversation-style stimuli, individualized levels of background noise or spectral degradation, or even virtual reality and speaker arrays to simulate listening environments. These attempts to test in a more realistic environment could be revealing when assessing an individual's speech understanding ability and reliance on context cues in an ecologically valid manner.

The findings of this study indicate that listeners with CIs rely on context to improve speech understanding. While listeners across the adult age range are affected

by context, the magnitude of the effect is not linearly related to age. In the current study, the magnitude of the context effect was smallest for the younger listeners and greatest for participants middle-aged and older. The use of context by listeners with CIs in everyday conversations and the individual factors that contribute to the use of context remain fruitful avenues for future research.

4 Study 3: Aging and Eye Movements: Strategies for the Use of Context

4.1 Introduction

A growing number of people who use cochlear implants (CIs) are over the age of 65 years and may be affected to greater or lesser degrees by physical and cognitive changes that accompany aging. Much is unknown about how aging influences the real-time communication abilities of an individual using a CI. Older listeners with and without hearing loss are expected to have more difficulty overriding contextual predictions that conflict with degraded acoustic information (Harel-Arbeli et al., 2021), and CI users have been shown to use a "wait-and-see" strategy during sentence comprehension (McMurray et al., 2017). The current study assessed the influence of age on real-time processing of degraded speech in background noise by adults who use CIs and adults with acoustic hearing (AH) who listened to spectrally degraded speech (i.e., a simulation of CI processing).

Compared to younger adults, older adults typically have extensive vocabularies and greater language knowledge and experience (Kavé & Halamish, 2015; Verhaeghen, 2003) that collectively can overcome age-related sensory deficits (Amichetti et al., 2018; Sommers & Danielson, 1999). Eye tracking has been used to explore prediction generation about forth-coming speech (Altmann, 1988; Huang et al., 2017), and it can provide information about real-time language processing as speech

unfolds over time (Allopenna et al., 1998; Farris-Trimble et al., 2014; Tanenhaus et al., 1995). One compelling reason to measure eye movement is that it allows for fine-grained analysis of real-time processing that can reveal processing or strategy differences even when behavioral performance is the same. In Studies 1 and 2, The current study employed a measure of eye tracking to compare the effects of context on the interpretation of sentence-initial target words between listeners with CIs and listeners with AH.

Performance in a phoneme categorization task, such as that used for Studies 1 and 2, does not provide insight into the timing of the decision-making process. It is impossible to know whether a participant changed their decision about the identity of a target word when that word was followed by an incongruent sentence context. Eye tracking is a tool that allows the researcher to measure such decisions that were obscured in the phoneme categorization task. Eye tracking also has the potential to reveal differences in the timing of such decisions. These timing differences could be indicative of differences in strategy, cognitive ability, or even motor planning ability among groups of listeners.

4.1.1 Eye Tracking for Real-Time Language Processing: Age

Several studies have investigated age-related changes in word recognition across the lifespan using the Visual World Paradigm (VWP) (Ben-David et al., 2011; Colby & McMurray, 2023; Revill & Spieler, 2012; Van Engen et al., 2020). These studies have focused on lexical factors, such as word frequency and rhyming, or

acoustic factors, such as background noise or accented talkers, to explain the deficits in word recognition observed in the older listeners. Some of these studies also explored cognitive abilities that might have predicted performance, such as working memory and inhibition abilities. In general, these prior investigations found that measures of working memory and inhibition abilities were highly correlated with the age of their adult participants, and they were thus unable to dissociate the contribution of these two factors to word recognition as measured by the VWP. To overcome these limitations, the current study uses advanced statistical modeling to compensate for age-related changes in motor speed, randomizes the presentation of context sentences, allows breaks to prevent fatigue, and includes a scientifically rigorous number of trials.

4.1.2 Eye Tracking for Real-Time Language Processing: CIs

Eye tracking within the VWP has been used to explore the real-time language processing abilities of listeners with CIs (e.g., Blomquist et al., 2021; Farris-Trimble et al., 2014; Harel-Arbeli et al., 2021; Huang et al., 2017; Kapnoula & McMurray, 2021; Martin et al., 2022; McMurray et al., 2017, 2019; F. Smith & McMurray, 2018). These studies, some in children with CIs and others in adults with CIs, have demonstrated the usefulness of the paradigm for this population. These studies have also revealed some differences in speech processing between listeners with AH and listeners with CIs. For example, when testing children 5-10 years of age, Blomquist et al. (2021) found that children with CIs were less efficient than their age-matched and vocabulary-matched peers at using semantic context cues to predict up-coming words in a sentence. Thus,

it is possible that adult listeners with CIs will show similar differences in the use of semantic contexts compared to adults with AH.

Studies of real-time language processing in listeners with CIs have revealed later commitments to target word choices compared to listeners with AH (McMurray et al., 2017, 2019; F. Smith & McMurray, 2018). This has been referred to as a "wait-and-see" strategy. McMurray et al. (2017) examined real-time lexical competition in prelingually deafened adolescent listeners with CIs and AH controls. They presented the target stimulus in the carrier phrase, "He said X." The sentences and target words were presented without background noise. Adolescent listeners with CIs demonstrated three effects compared to peers with AH: 1) a large delay before fixating on any picture on the screen; 2) less competition from words with similar onsets; and 3) more competition from words that rhymed with the target word. These results led the authors to propose that the "wait-and-see" strategy might be used by listeners with CIs and not by listeners with AH. However, when young listeners with AH were presented 4-channel vocoded speech (a highly degraded signal that led to matched word recognition accuracy with the CI group), the young listeners with AH demonstrated the same patterns of performance as the listeners with CIs. Delayed responses, less competition from words with the same onset, and more confusion with words that rhymed were hallmarks of performance with degraded speech for both groups. The current study expands these previous findings by: (1) using sentences containing semantic context rather than isolated words; (2) increasing the generalizability of the findings by testing across the adult lifespan; (3) substituting mostly post-lingually deafened listeners with

CIs for the prelingually deafened listeners; and (4) manipulating the degradation of the target words for both groups of listeners by testing in two levels of background noise.

4.1.3 Location of target words and context effects

Studies on the use of sentence context have traditionally focused on the effect of context on the identification of a sentence-final word (e.g., Amichetti et al., 2018; Connine & Clifton, 1987; Dubno et al., 2000; Kalikow et al., 1977; Pichora-Fuller, 2008). Less is known, however, about how listeners use sentence context when the target words are at the beginning of sentences, in which subsequent words could provide cues that would allow a listener to infer the missing or unclear word. Therefore, the current study will focus on sentence-initial target words that are presented in background noise and are followed by a context sentence that provides congruent or incongruent context cues for interpreting the target word. This is a novel paradigm that builds on the preliminary results of Studies 1 and 2 to further quantify the effects of sentence contexts on target word categorization in listeners with CIs and listeners with AH. The use of eye-tracking will allow for a fine-grained analysis of the timing of listeners' decisions that may or may not be influenced by the following semantic context, and thus may prove useful in determining if listeners with CIs use more time to process a degraded signal and use subsequent semantic context than AH listeners. A group difference in processing time may lend support to the theory of a "wait-and-see" strategy used by listeners with CIs. It is also possible that differences in processing time

could be driven by differences in the speed of decision making or motor planning that do not reflect strategic differences.

4.1.4 Experimental Questions and Hypotheses

The objective of this study was to measure the extent to which real-time speech understanding is affected by sentence context for listeners with CIs or with AH as a function of age. The experimental question was whether or not older listeners employ different processing strategies than younger listeners when a spectrally degraded sentence-initial word is unclear. The working hypothesis was that older participants would be more influenced by the incongruent contexts than younger participants. Specifically, older participants would provide more responses that were congruent with the context sentence than the acoustics of the target words. Study 2 demonstrated that older listeners with CIs relied heavily on context when the target word was in the sentence-initial position. This was partially explained by age-related changes to working memory. It was also hypothesized that older listeners would show longer delays before fixation to the target word than younger participants due to age-related slowed motor movements.

4.2 Method

4.2.1 Participants

A total of 24 adult (20-85 years old) CI participants with mostly post-lingual-onset hearing loss (i.e., hearing loss onset at age 5 or older) and 33 adult control participants with AH, were recruited (Table 4-I). Some of the data for this experiment were collected at the same time as the behavioral data collected in the first two experiments, on the same groups of participants with the same inclusion criteria. Individuals were allowed to participate in either Study 1 or Study 2 and Study 3, but were not required to do so. CI participants had one or more devices from one of two FDA-approved manufacturers with at least one year of experience with their CI(s). They used their default everyday programs on their own processors that had been checked for functionality prior to testing. If a participant with a CI reported acoustic hearing, thresholds were measured in that ear using an audiometer. If any thresholds were less than 60 dB HL, an earplug was worn in that ear. The use of hearing aids was not allowed. Two participants used hybrid devices with components for both electric and acoustic hearing. These participants took the acoustic receivers out of their ear canals and wore ear plugs. They relied solely on the electric hearing provided by their CIs (CI02 and CI03).

Table 4-I. Demographic information for participants. Participants are shown by age with listeners with acoustic hearing (AH) on the left of the bold line and those with CIs on the right.

Participants with Acoustic Hearing			Participants with Cochlear Implants		
Participant	Age	Gender	Participant	Age	Gender
20's			20's		
AH01	20	F	CI01	22	F
AH02	20	F	CI02	22	F
AH03	21	F	CI03	25	M
AH04	21	F	CI04	26	F
AH05	23	M			
30's			30's		
AH06	30	F	CI05	33	F
AH07	31	F	CI06	33	M
AH08	32	F	CI07	38	M
AH09	36	F			
AH10	38	F			
40's			40's		
AH11	43	F	CI08	49	M
AH12	44	F			
AH13	44	F			
AH14	46	F			
AH15	49	F			
50's			50's		
AH16	56	F	CI09	51	M
AH17	57	F	CI10	52	M
AH18	57	F	CI11	55	F
AH19	58	F			
60's			60's		
AH20	61	M	CI12	61	M
AH21	61	F	CI13	63	M
AH22	62	M	CI14	65	M
AH23	63	F	CI15	66	F
AH24	66	F	CI16	66	F
AH25	66	F	CI17	66	F
AH26	67	F	CI18	69	M
AH27	68	F			
70's			70's		
AH28	70	F	CI19	71	F
AH29	70	M	CI20	74	M
AH30	72	F	CI21	75	F
AH31	72	F	CI22	78	F
AH32	77	F	CI23	79	M
80's			80's		
AH33	88	F	CI24	83	F

The AH participants had no more than a slight hearing loss (≤25 dB HL) at octave frequencies between 250-4000 Hz (re: ANSI, 2018). All participants were native

speakers of American English, based on self-report, to avoid variance from diverse language backgrounds. Any participants who reported conditions that might preclude their participation in an eye-tracking study (e.g., large cataracts or amblyopia) were excluded from testing. In addition, all 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.

4.2.2 Stimuli

The auditory stimuli were sentences, predicting one target word or the other, played after sentence-initial target words from a /d/-/t/ contrast continuum (deer-tear) and a /b/-/p/ contrast continuum (beak-peak). The deer-tear contrast was chosen to be parallel to Studies 1 and 2. The beak-peak contrast was chosen as a second stimulus set because, like deer-tear, it also represents a voice-onset timing (VOT) contrast in single-syllable words that can be visualized and are roughly matched in word frequency and familiarity. A continuum was created from recordings made by the same talker as in Studies 1 and 2. The endpoint word “beak” had a VOT of 10 ms, while the endpoint word “peak” had a VOT of 70 ms. The beak-peak continuum was used in another categorization study in the lab, where participants clearly identified the endpoints and categorized the steps along the continuum as expected.

The specific target words for this experiment were steps 2 and 3 from the 4-step deer-tear continuum used in the previous study (VOTs of 46 and 66 ms) and steps 2 and 3 from a 4-step beak-peak continuum (VOTs of 30 and 50 ms). These target words

were partially ambiguous, with one slightly more likely to be identified as one of the endpoint words in the contrast and the other slightly more likely to be identified as the other endpoint word of the contrast. The target words chosen were on either side of most participants' 50% crossover points on the categorization psychometric functions of previous experiments (i.e., Studies 1 and 2 and another study performed in the lab that used the beak/peak stimuli in a categorization task). The sentences used for predictive contexts for each of the target words are presented in Table 4-II.

Table 4-II. Sentences used to provide context.

Sentence Position	Context Type	Sentence Stimuli
Initial	Deer	The ___ lost its antlers in the forest.
Initial	Tear	The ___ fell from her eye down her cheek.
Initial	Beak	The ___ of the chicken was full of twigs.
Initial	Peak	The ___ of the mountain was covered with snow.

The target visual stimuli were a photograph of a deer, a drawing of an eye with a tear falling from it, a photograph of the head of an eagle with a prominent beak, and a photograph of a mountain peak (see Figure 4-1). The images were high-resolution, converted to grayscale, and equalized in size and lumen so that they would be equally likely (or unlikely) to draw a participant's attention away from the target. The images were 8.2 degrees of visual angle in size. The images were each presented in one of the four corners of the screen with their most-central corners located 14.6 degrees of visual angle from the center. The center of the monitor screen contained a fixation point at the beginning of each trial.

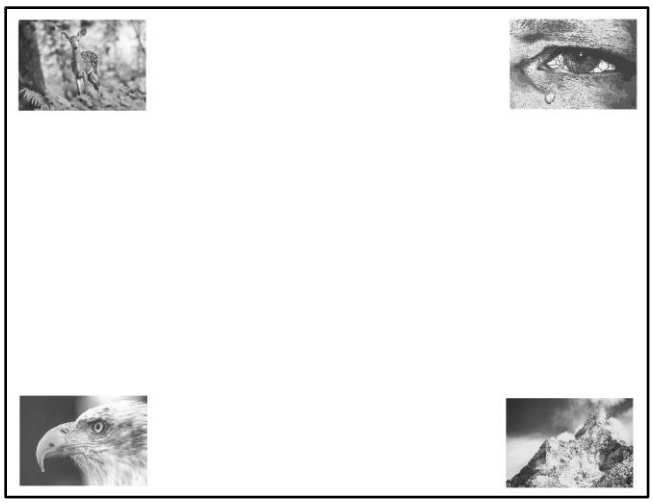

Figure 4-1. Visual stimuli representing deer, tear, peak, and beak in clockwise order starting from the upper left

All auditory stimuli for AH participants were presented as a simulation of CI-processed speech (vocoded) using an 8-channel sine vocoder. The vocoder divided the frequencies between 250-4000 Hz into 8 channels using 3rd-order Butterworth filters. The resulting signals were half-wave rectified and filtered with an envelope cut-off of 150 Hz using a 2nd-order Butterworth filter applied forward and backward (Shader, Yancey, et al., 2020). This type of vocoding is commonly used to compare performance across participants with AH and participants with CIs. It is possible that participants who have AH might show a measurable effect of context on ambiguous target words in background noise even if the target words were not vocoded. However, the differences in signal degradation would make comparisons to listeners with CIs problematic. 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).

4.2.3 Procedure

Participants were seated comfortably in a sound-attenuating booth. Auditory stimuli (context words and target sentences) were presented over one loudspeaker (Yamaha HS5, Buena Park, CA, USA) located approximately one meter away from the participant's chair at 0 degrees azimuth and above the computer monitor. The level of stimulus presentation was calibrated to 70 dB A by playing a speech-shaped noise created from the long-term average spectrum of the stimuli used in this experiment through the loudspeaker. This calibration noise was measured by a B&K 2250 sound level meter with a quarter-inch microphone at the location of the participant's head. The 23-inch computer monitor for presenting visual stimuli and collecting responses was located on a desk approximately 0.75 meters away from the participant's chair at 0 degrees azimuth. The desk-mounted eye-tracking camera (EyeLink 1000 Plus; SR Research Ltd., Ottawa, Canada) was located in front of the computer monitor directly in front of the participants. The camera recorded the eye movements and tracked the gazes of participants. Participants placed their chins on a table-mounted chin-rest that maintained their eyes at a fixed distance from the camera and monitor. Before the practice session and before each block of the experiment, the eye-tracking software was calibrated for more accurate tracking of each participant's eye movements. The experiment presentation software, Experiment Builder (SR Research Ltd., Ottawa, Canada), was used to calibrate gaze location, present stimuli, and record responses. Eye-gaze position was recorded as a series of x-y coordinates at a rate of 1000 Hz

throughout the training and experimental blocks. In addition, categorization responses and reaction times were also recorded.

Prior to beginning the experiment, participants completed a short practice session. The four context sentences used in the experiment were presented in the practice session with unambiguous versions of the target words (steps 1 and 4 from each continuum). During the practice session, the context sentences were presented in quiet, and the practice target words were presented once in +10 dB SNR background noise and once in +5 dB SNR background noise. The practice target words played at the beginning of each context sentence were the endpoint words that were congruent with the individual sentence (i.e., the endpoint word "deer", step 1, was played at the beginning of the sentence, "The ____ lost its antlers in the forest"). The practice session had several objectives: 1) to ensure that the context sentences were easily recognized and understood during the task; 2) to familiarize the participants with the visual stimuli; 3) to allow listeners with AH time to adapt to 8-channel vocoded speech; and 4) to provide familiarity with the task.

During the practice session and the experimental session, each trial began with a fixation point at the center of the screen. Participants were asked to start the trial by looking at the fixation point. When the participant's gaze was fixated on the center of the screen, the researcher initiated the trial. A screen that showed the four target pictures in the four corners of the screen appeared, and the auditory stimulus, consisting of a noise-masked target word followed by an unmasked context sentence, was played. Participants were instructed to use the mouse to click on the picture representing the

target word that they heard. The pictures of the four possible target words were visible on the screen while the target word and sentence played, but the triangle representing the mouse's location did not appear until the end of the sentence. The triangular pointer always appeared centered on the screen after the final word of the sentence. Thus, the participant had to wait until the end of the sentence to move the mouse toward the picture representing the target word.

The experiment had two context conditions (congruent and incongruent), two levels of masking noise during the target words (+5 and -5 dB SNR), and two target words from each continuum (steps 2 and 3 from the 4-step deer-tear and beak-peak continua). The four target words were presented eight times in each context condition at each level of background noise for a total of 64 trials. These trials were presented in two blocks of 32 trials with all conditions randomized within blocks. In addition to the eye-tracking task, participants completed the Stroop task and the NIH Toolbox List-Sort Working Memory task, following the procedures outlined in Study 1. If participants had completed either Study 1 or Study 2, they did not need to repeat these cognitive measures.

Behavioral performance was measured in two ways: 1) accuracy for choosing each acoustic target word in each SNR and each context condition was calculated; and 2) response times for categorization choices were recorded. The scores from the cognitive tasks were used in the analysis as predictors of the effect of context. The data from the eye-tracker were analyzed from the onset of the target word until the

participant clicked the mouse on the picture representing their identification of the target word.

4.2.4 Analysis

Data analysis for the behavioral data followed the same procedure described previously. Target word categorization accuracy based on the acoustics of the target stimuli from each participant on each trial was modeled with a logistic mixed-effect model (MEM) using a binomial distribution. The model contained fixed effects of age, modality (CI or AH), context type (congruent or incongruent), SNR (+5 or -5 dB), and their interactions. Age was standardized such that the mean was zero, with a standard deviation of one, and treated as a continuous variable. The mean age (52.8 years old) was used as the reference in the model. The reference values for the categorical variables of modality, context type and SNR were "AH," "Congruent," and "+5 dB," respectively. The log-transformed response times were modeled in a linear MEM (Gaussian distribution) with the same fixed effects and random effects structure as the previous analysis. The statistical analysis was conducted in R (version 4.3.3; R Core Team, 2024) using the *buildmer* package version 2.11 (Voeten, 2023) and *lme4* package version 1.1.35.3 (Bates et al., 2015). The residuals of each model were checked to ensure there were no major violations of normality or homoscedasticity.

The trial-by-trial gaze-tracking data were analyzed using the *eyetrackingR* package version 0.2.1 (Forbes et al., 2023). The data were first down-sampled from 1000 Hz to 100 Hz. Each sample thus represented 10 samples recorded from the eye-

tracker. For each sample, the proportion of samples during which the participant's gaze was fixated on or near the picture representing the acoustic target was recorded. This same summary was performed for each of the items on the screen as well as the white space between the items of interest. The samples from each trial were divided into 20 bins representing performance during the target word presentation, 20 bins representing performance during the context sentence, and 20 bins representing performance after the context sentence until the participant used the mouse to select their response. Once again, looks to the various items on the screen were summarized for each bin. The 60 time bins that covered the whole trial were then analyzed sequentially with t-tests (using Bonferroni corrections for multiple comparisons) to determine time periods of statistically different performance on each of the predictive factors in the study (listener group, SNR, congruency, and age group). This method does not allow factors to interact. For the eye-tracking analysis, participants were separated into two age groups, older (n=21) and younger (n=26). Participants 65 years old and older were categorized into the "older" age group. The dependent variables were the empirical logit transformations of the proportion of looks to target data because this transformation best approximated a Gaussian distribution.

4.3 Results

4.3.1 Accuracy

Results are first presented as the accuracy of participants' responses identifying the acoustic target words as a function of age (Figure 4-2). Responses are divided into panels representing performance when the target words were presented in sentences that matched the acoustic target word (congruent) and in sentences that did not match the acoustic target word (incongruent). Listeners with CIs are represented by the blue lines and listeners with AH are represented by the red lines.

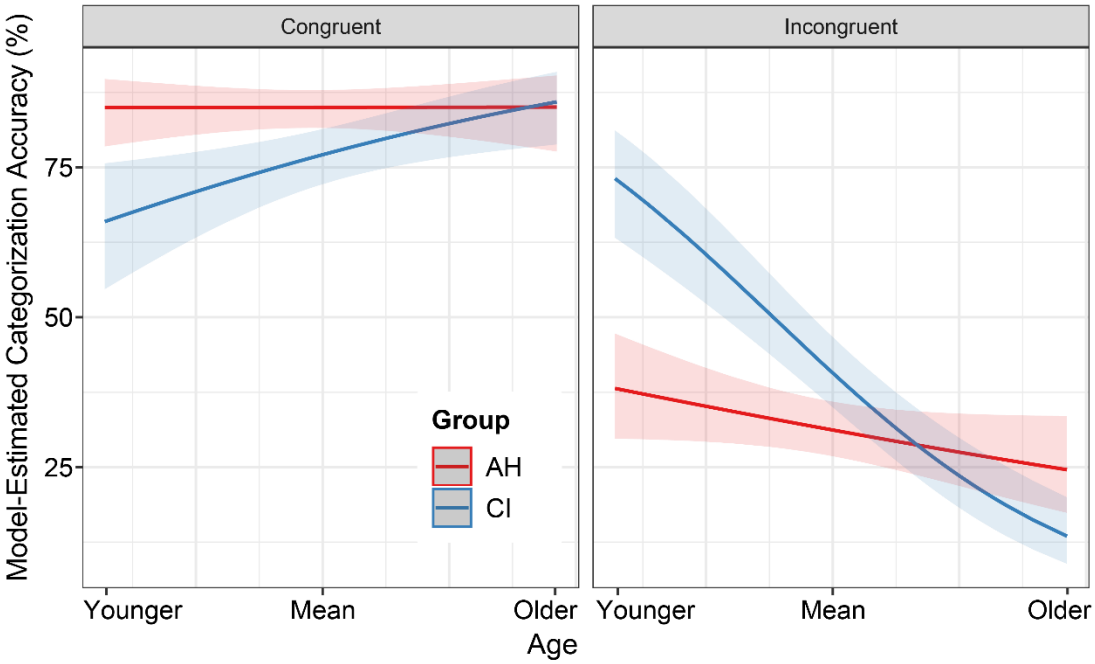

Figure 4-2. Model-estimated accuracy by congruency and listener group collapsed across SNR conditions as a function of age. Shaded regions represent 95% confidence intervals.

An examination of Figure 4-2 reveals that participants maintained high accuracy when the sentence context matched the acoustic target word. When the

context sentence was incongruent with the acoustic target word, older participants were less likely than younger participants to respond with the acoustic target word. The effect of age appears to be especially strong in the participants with CIs. The data were modeled in a logistic MEM with accuracy as the dependent variable in a binomial distribution, and listener group, context congruency, age, and SNR as independent variables. The cognitive measures of working memory, processing speed, and inhibition ability were eliminated from the model by the *buildmer* package. The output of the model is shown in Table 4-III.

Table 4-III. Output of a logistic mixed-effects model to predict accuracy as a function of listener group, context congruency, age, and SNR during the eye-tracking task. Reference values are the AH listener group, Congruent context, +5 dB SNR, and mean age.

Formula: Accuracy ~ Listener Group * Context Congruency * Age + SNR * Context Congruency + (1 | participant)

Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	1.734	0.126	13.778	**<0.001**	***
Group-CI	-0.520	0.163	-3.185	**0.001**	**
Congruent-Incongruent	-2.525	0.140	-18.08	**<0.001**	***
Age	0.002	0.112	0.019	0.985	
SNR(-5)	0.347	0.129	2.684	**0.007**	**
Group-CI × Context-Incongruent	0.935	0.172	5.424	**<0.001**	***
Group-CI × Age	0.319	0.162	1.970	**0.049**	*
Context-Incongruent × Age	-0.181	0.119	-1.525	0.127	
Context-Incongruent × SNR(-5)	-1.088	0.171	-6.366	**<0.001**	***
Group-CI × Context-Incongruent × Age	-0.941	0.171	-5.505	**<0.001**	***
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
By-Participant effects	0.136	0.369			
Observations	3648				

The model shows a significant three-way interaction (seen in Figure 4-2) such that there is a significantly greater effect of age for participants with CIs than for participants with AH in the incongruent context condition (Group-CI × Context-Incongruent × Age; $p < 0.001$). The model also shows a significant interaction between

SNR and congruency (Figure 4-3), such that accuracy was lower in the +5 dB SNR condition compared to the -5 dB condition when the context was congruent, but higher in the +5 dB SNR condition than the -5 dB SNR condition when the context was incongruent [Context-Incongruent × SNR(-5); $p < 0.001$].

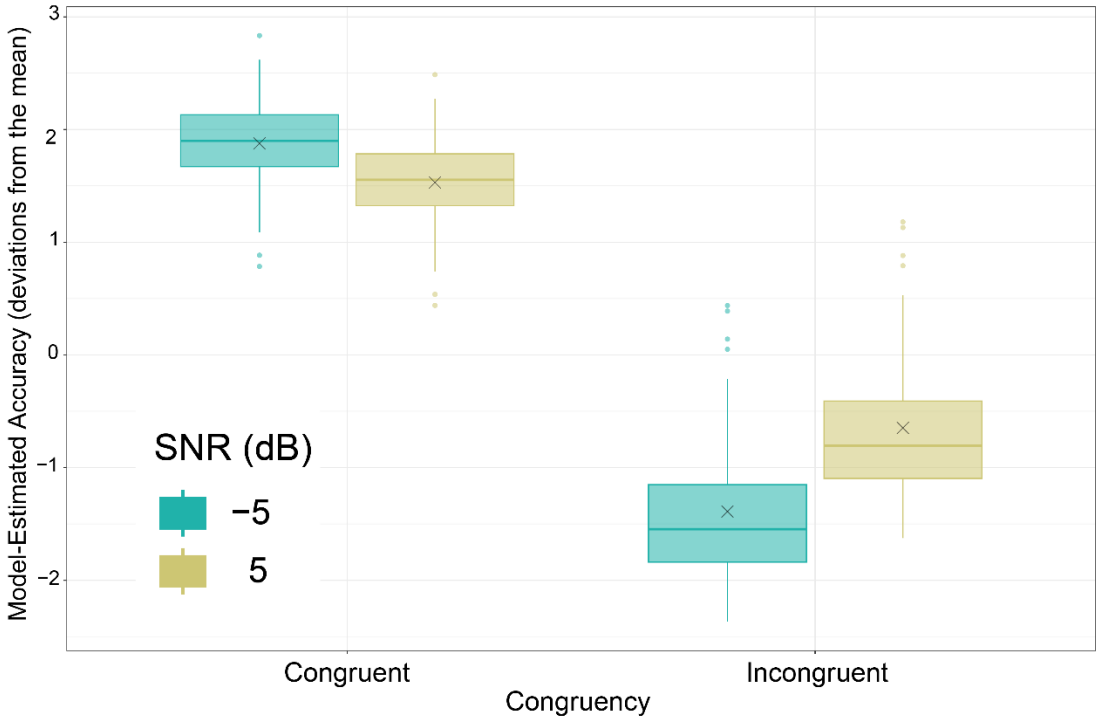

Figure 4-3. Model-estimates of accuracy in standard deviations from overall mean accuracy collapsed across listener groups as a function of context congruency and SNR. The "x" marks the mean accuracy in each condition.

These results show that congruency, SNR, age, and listener group all play a role in the accuracy of participants' performance in this task. Age seems to affect the performance of participants with CIs more than that of participants with AH. This could reflect differences in strategy across listener groups or differences in access to the acoustic signal.

4.3.2 Gaze Trajectories

The plots of gaze trajectories are divided into three time sections: during the presentation of the target word, during the context sentence, and after the context sentence until the participant clicked on their choice for the target word. The time sections were divided into normalized bins for plotting purposes; depending on the condition, the length of the Target Word section was 936-996 ms, the length of the Context Sentence section was 1520-1818 ms, and the length of the Response section was 100-10000 ms (*M* = 1460 ms, *SD* = 1042 ms).

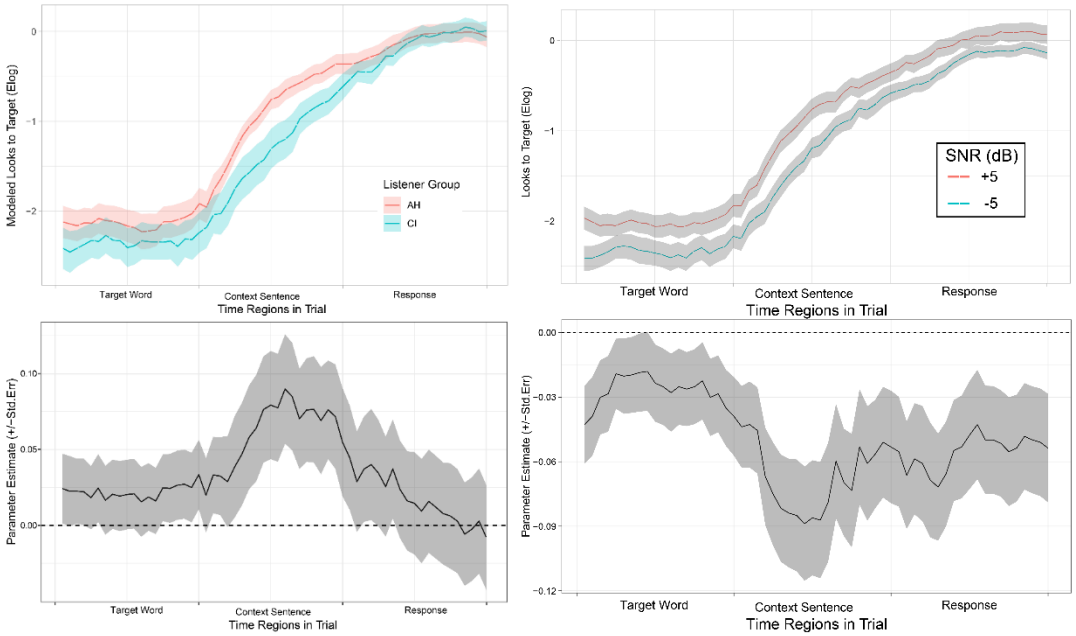

Figure 4-4. The top left panel represents modeled looks to target over the course of the trial divided by listener group (AH participants in red, CI participants in blue). The bottom left panel represents the difference between the gaze trajectories of the two listener groups. The top right panel represents looks to target over the course of the trial in each of the SNR conditions. The bottom right panel represents the difference between the gaze trajectories for the two SNR conditions shown in the top panel. Each time bin was checked for a significant difference by t-test with a Bonferroni correction applied. There were no significant differences overall for either listener group or SNR condition.

The overall trajectory of looks to target by each listener group is shown in Figure 4-4 in the top left panel. Preliminary tests for significance using t-tests across time bins with the highly conservative Bonferroni correction applied did not reveal a statistically significant difference in looking patterns between listener groups ($p > 0.05$ for all). A similar plot of looks to target in each SNR condition is shown in Figure 4-4 in the top right panel. Preliminary tests for significance using t-tests across all time bins in the trial with the highly conservative Bonferroni correction applied also revealed no significant differences in looks to target as a function of SNR ($p > 0.05$ for all). These results can be interpreted as non-significant effects of listener group and SNR on gaze trajectories across time.

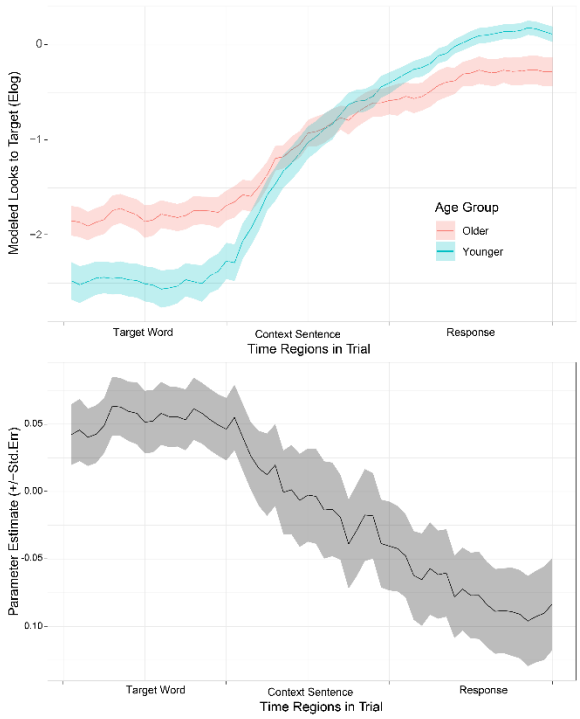

Figure 4-5. The top panel represents modeled looks to target over the course of the trial by each age group (older participants in red, younger participants in blue). The bottom panel represents the difference between the gaze trajectories of the two age groups. Each time bin across the whole trial period was checked for a significant difference by t-test with a Bonferroni correction applied. There were no bins where the difference between age groups was significantly different from zero.

The overall trajectory of looks to target by age group is shown in Figure 4-5 in the top panel. Preliminary tests for significance using t-tests across time bins with the highly conservative Bonferroni correction applied did not reveal a statistically significant difference in looking patterns between age groups ($p > 0.05$ for all). These results can be interpreted as a non-significant effect of age group on gaze trajectories across all time bins.

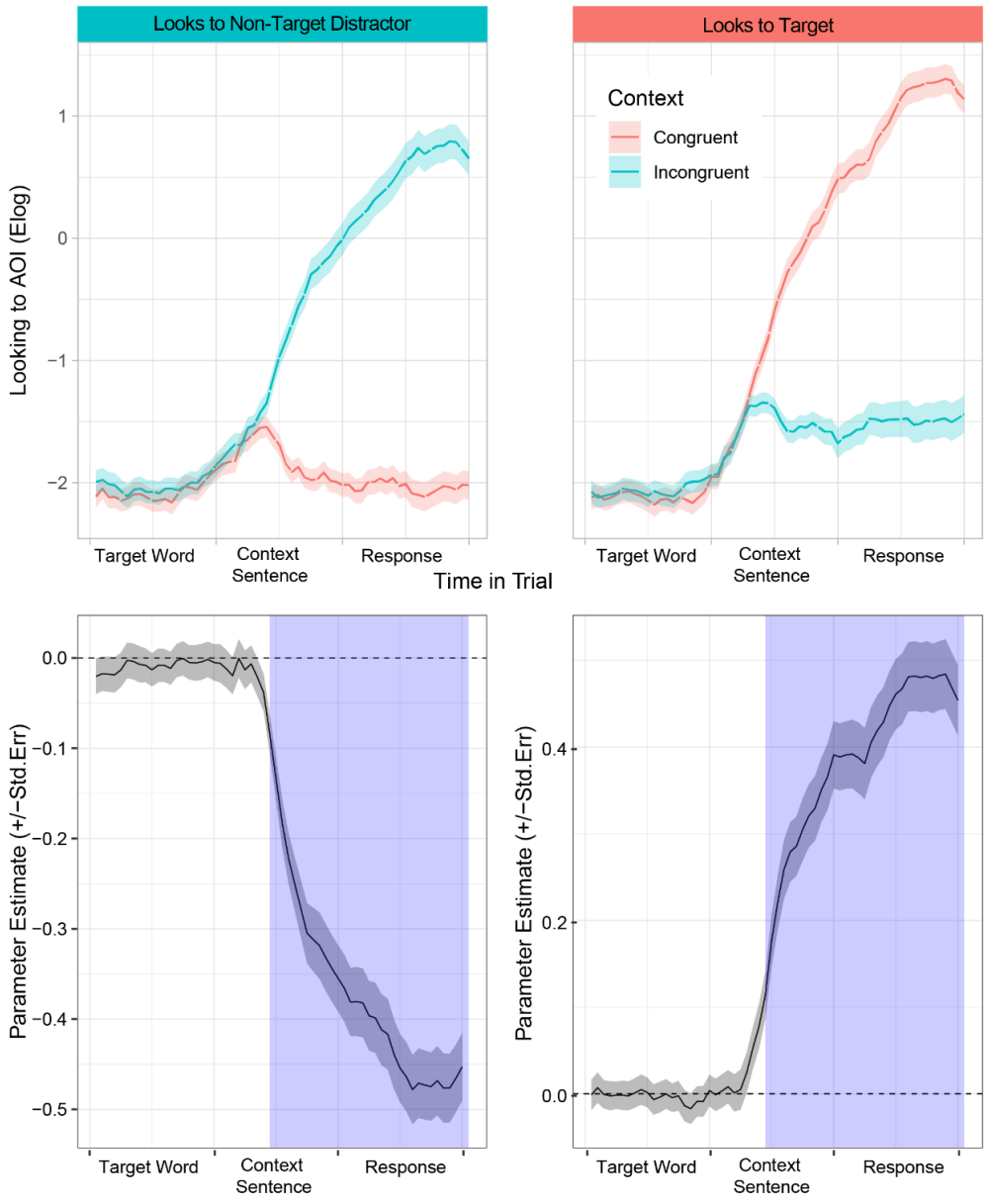

Figure 4-6. Top panels show looks to Non-Target Distractor (left) and Target (right) as a function of context across time. The lower panels show the difference in gaze trajectories between congruent and incongruent contexts for looks to the distractor and looks to the target. The time regions where there is a statistically significant difference between the two gaze trajectories are highlighted with purple shading.

Figure 4-6 compares gaze trajectories when participants were looking at the target or the non-target distractor in congruent and incongruent context conditions. The

overall trajectory of looks to the non-target distractor is shown on the left in Figure 4-6. The data show that looks to the distractor occur when the sentence context is predictive of that distractor (blue-ish line). Looks to the distractor when the context predicts the target disappear rapidly (reddish line). The overall trajectory of looks to the target is shown on the right side of Figure 4-6. The data show that participants looked to the target when the context predicted the target significantly more than when the context did not predict the target. Preliminary tests for significance using t-tests across time bins with the highly conservative Bonferroni correction applied revealed statistically significant differences in looking patterns between both looks to the distractor and looks to the target as a function of context ($p < 0.05$ for all purple shaded regions in the bottom panels of Figure 4-6).

These preliminary tests of the effects of individual factors on the eye-tracking data provide a glimpse into real-time processing of speech across a whole sentence. These results show no difference between the gaze trajectories of participants with CIs compared to participants with AH, gaze trajectories at a +5 dB SNR compared to a -5 dB SNR, nor between gaze trajectories of older compared to younger participants. The small differences that are visible in Figures 4-4 and 4-5 may be significant when tested with more sophisticated techniques. This basic analysis does show the large effect of context on the gaze patterns of all listeners. To determine differences in the shape of the gaze trajectories across the trial, a growth-curve analysis would be appropriate. To determine the onset of any significant differences across trajectories, generalized

additive mixed models or divergence point analyses would be appropriate (as suggested in Ito & Knoeferle, 2022).

4.3.3 Response Times

Response times for choosing the target word were modeled in a linear mixed-effects model (MEM) with log-transformed response times as the dependent variable and listener group, context congruency, age, and SNR as independent variables. The scores from the cognitive measures representing working memory ability, processing speed, and inhibition ability were also entered in the model. The model was built with the *buildmer* package in R. Working memory was eliminated from the model during the building process. The output of the model is shown in Table 4-IV.

Response times were significantly longer when the target word was presented in the -5 dB SNR condition (SNR(-5), $p < 0.001$). There was considerable individual variability in the response times (as seen in the spread of the light grey lines in Figure 4-7). While the effect of SNR was statistically significant, the actual difference in means between the two SNR conditions was 175 ms, which may not be practically significant.

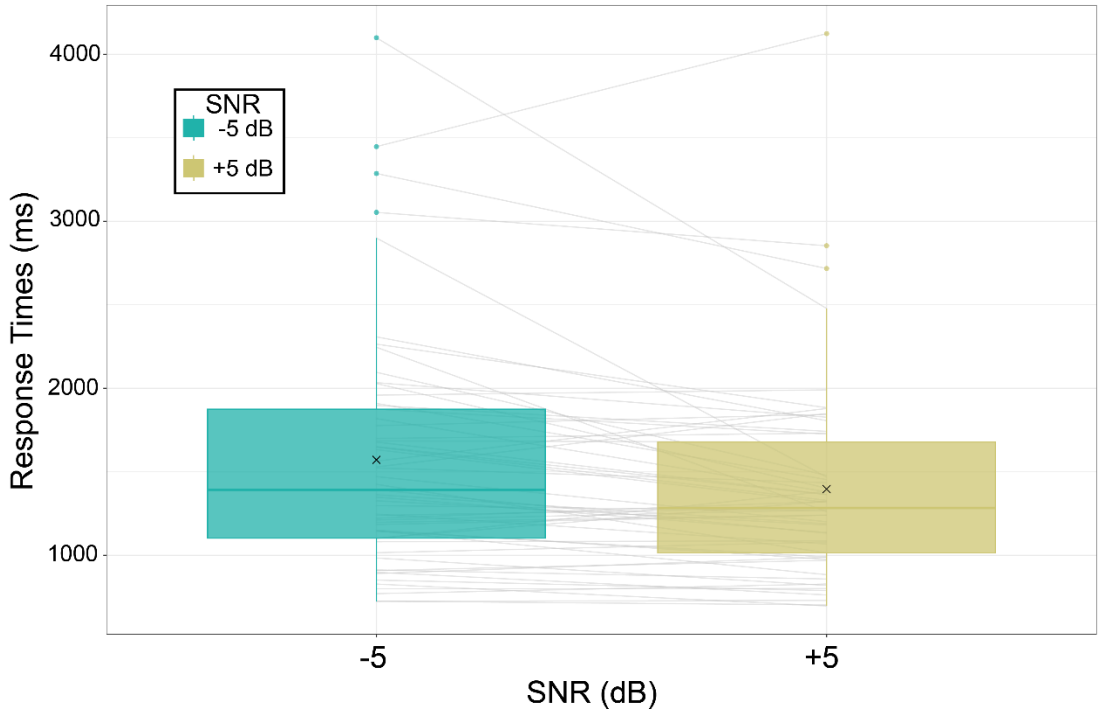

Figure 4-7. Response times in milliseconds (ms) in each SNR condition. Data are shown in boxplots with an "x" representing the mean response time and the light grey lines representing mean individual data for both listener groups, all ages, and both congruency conditions.

The output of the model also shows that the measure of inhibition ability was predictive of response times. There was a significant interaction between inhibition ability and context congruency such that the response times of individuals with greater inhibition ability were faster for congruent contexts than incongruent contexts, while the response times of individuals with poorer inhibition ability were faster for incongruent contexts than congruent contexts (Figure 4-7; Inhibition × Context-Incongruent, $p < 0.001$). In this task, age did not significantly predict response times. There was also no significant interaction between context and listener group on response times.

Table 4-IV. Model output of linear mixed-effect model on normally distributed log-transformed response times. Reference values are AH for Listener Group, +5 for SNR, and Congruent for Context.

Formula: Response Time ~ Inhibition Ability + Processing Speed + SNR + Context Congruency * (Age + Inhibition Ability + Listener Group) + (1 + SNR | participant)

	Response Time				
Predictors	*Estimate*	*Std. Error*	*z-value*	*p*	
(Intercept)	7.037	0.054	131.27	**<0.001**	***
Inhibition	0.165	0.050	3.280	**0.002**	**
Processing Speed	0.077	0.042	1.834	0.073	
SNR(-5)	0.099	0.024	4.203	**<0.001**	***
Context-Incongruent	0.022	0.018	1.196	0.232	
Age	-0.011	0.051	-0.220	0.827	
Group-CI	0.135	0.083	1.627	0.110	
Context-Incongruent × Age	0.031	0.017	1.845	0.065	
Inhibition × Context-Incongruent	-0.071	0.017	-4.231	**<0.001**	***
Context-Incongruent × Group-CI	-0.017	0.028	-0.597	0.551	
Random Effects:	*Variance*	*St. Dev.*	*Corr.*		
By-Participant effects	0.083	0.289			
By-Participant SNR slopes	0.020	0.143	0.08		
Observations	3648				

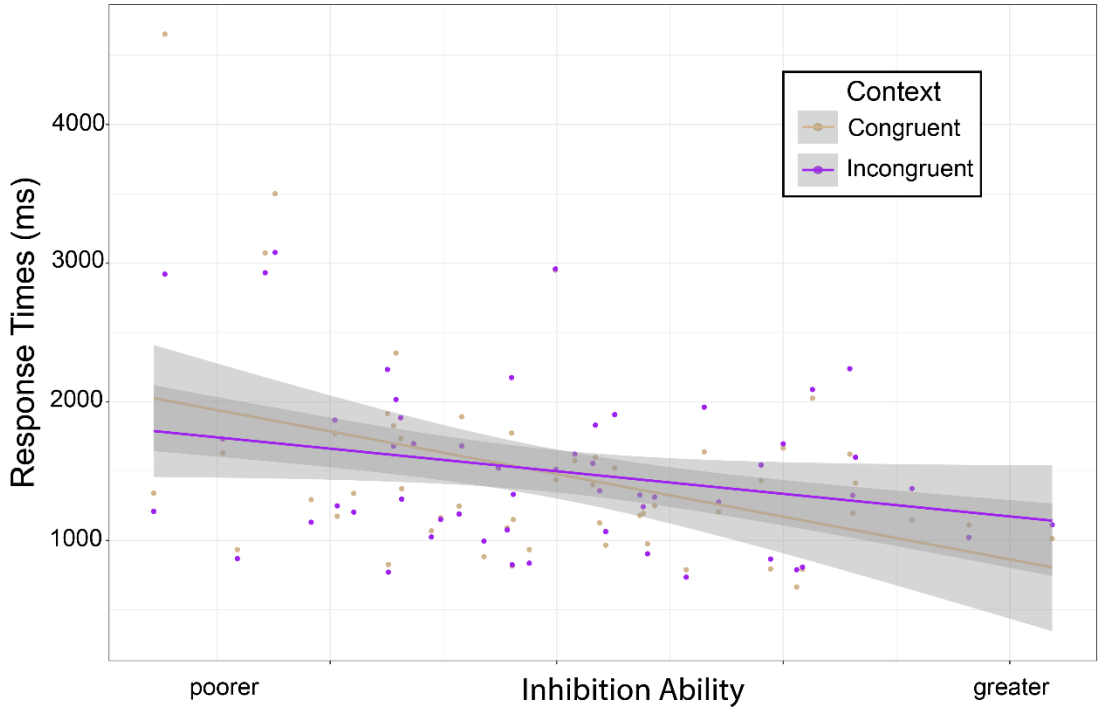

Figure 4-8. Response times in milliseconds (ms) as a function of inhibition ability and context congruency.

4.4 Discussion

This study tested adults with AH and adults with CIs on a phoneme categorization task where the target word was presented in background noise at the beginning of a sentence. The sentence contained lexical and semantic information that was either congruent with the acoustic target word or incongruent. The broad goal was to examine the interaction of listener group, age, and sentence context on real-time speech processing. The hypothesis was that participants with CIs would demonstrate a wait-and-see strategy compared to participants with AH, and that increased age would result in a greater reliance on context, especially in participants with CIs.

Analysis of the accuracy of participants' responses in the congruent condition showed that participants with CIs were less accurate overall than participants with AH (Figure 4-2). Younger participants of both listener groups were less accurate than older participants. There was an interaction between age and listener group such that the performance of older participants with CIs and with AH was more similar than the performance of younger participants with CIs and with AH. Given that the sentences provided supportive context, if the participants understood the context sentence and chose the target word predicted by it, their accuracy would be 100%. The performance of the older participants approaches this value. It is possible that this is an indication that older participants of both listener groups are more likely to trust the context to inform them of the identity of the target word than younger participants.

Analysis of the accuracy of participants' responses in the incongruent condition showed that participants with CIs were more accurate overall than participants with AH. However, there was an interaction between age and listener group; the performance of younger participants with CIs was more accurate than that of younger participants with AH, and the performance of older participants with CIs was less accurate than that of older participants with AH (Figure 4-2). It should be noted that the younger participants with CIs had the lowest accuracy in the congruent condition and the highest accuracy in the incongruent condition. The highest accuracy, it should also be noted, was roughly chance level (i.e., 50%). If participants were guessing on all trials, accuracy would be at 50% for both congruent and incongruent context sentences. Thus, younger participants with CIs appear to be using context for some, but not all, of their responses. Similar to the older participants' performance in the congruent condition approaching 100%, the performance of older participants in the incongruent condition approached 0%. Such a pattern could indicate that older participants are heavily relying on context to inform their decisions – more so than younger participants. Thus, younger and older participants of both listener groups (AH and CI) appear to be demonstrating clear differences in the use of subsequent semantic context.

It is not surprising that participants with CIs had lower accuracy on a speech-in-noise task than participants with AH. However, the increased reliance on context exhibited by the older participants in both listener groups provides more evidence for the age-related use of context that appeared in Studies 1 and 2. The performance of the younger participants with CIs could indicate that they had more difficulty

understanding the context sentences than the other participants or that they adopted a strategy that resulted in more responses overall to words that were not predicted by the context sentences. The accuracy analysis does not provide enough information to determine whether this group of participants was spreading their responses among all potential target words equally (e.g., perhaps not understanding the context sentences), or using a strategy that rewarded responses to the word that was specifically opposite the target word (e.g., responding "tear" when the target was "deer").

Eye-tracking studies provide researchers with insight into the real-time processing of speech. The timing of a participant's eye movements can indicate the decisions that they are making. For example, in this task, the participant hears the target word and within 150 ms begins to look at the picture representing that target word. However, while the participant is deciding where to look and moving their eyes, the subsequent context is presented. If the subsequent context is congruent, it supports their initial interpretation of the target word and should increase their confidence in the target word. This would be seen as faster more exclusive looks to the target. If the subsequent context in incongruent, the participant must decide if they are going to keep looking at the word they initially thought they heard or move their eyes to the picture representing the word predicted by the sentence. This indecision might be shown as a slower convergence of looks to target, a momentary decrease in looks to target, or even an abandonment of looking to target in favor of looking at the non-target distractor that matches the context sentence.

A “wait-and-see” strategy would be demonstrated by a delay in looking to the target word. This could be shown as a shallower slope of the gaze trajectory and as a later crossing into a non-chance proportion of looks to the target. Someone who was not employing a “wait-and-see” strategy would immediately begin to look at their choice of the target word and potentially change course later in the trial if they decide that their initial choice of target was incorrect. This would lead to more reversals in the gaze trajectory than someone who waited until they were certain before looking to their choice of the target.

All of the predictive context sentences contained a semantically related word within 2-3 syllables of the target word and another semantically related word at the end of the sentence. It is possible that some listeners may initiate looks to target either during or immediately after they hear the target word, but then keep periodically looking at the other pictures before making up their minds (a shallow slope in gaze trajectory). Another possibility is that some listeners may wait until they have heard one or both of the semantically related words in the sentence before they start looking at the target word (a steep slope with a crossover later in the sentence compared to others). Both of these strategies could be interpreted as a “wait-and-see” strategy.

The statistical analysis of the gaze trajectories (plotted in the bottom panel of Figure 4-4) shows that SNR does not affect the timing of looks to target despite the consistent differences visible in the accompanying plot of gaze trajectories (top panel of Figure 4-4). Figure 4-4 also shows the proportion of looks to the acoustic target by listener groups, with data collapsed across sentence contexts, age groups, and SNRs. In

the plot, it looks as though participants with AH looked to the target word sooner overall than participants with CIs (the reddish line is to the left of the blueish line when the lines begin to slope upward). However, the preliminary statistics show no significant difference across listener groups in their gaze trajectories across time. Therefore, the findings currently provide no support for the "wait-and-see" strategy proposed by McMurray et al. (2017, 2019). Future analyses may shed more light on any potential differences in real-time processing.

Age was treated as a categorical variable in this analysis. The gaze trajectories of participants in each age group were plotted in the top panel of Figure 4-5. The gaze trajectory of the older participants has a shallower slope than that of the younger participants. However, the basic statistical analysis performed in this study found no significant differences between the gaze trajectories of these two groups across time. This is contrary to the hypothesis that older participants would demonstrate delays in their eye movements compared to younger participants as a result of age-related motor slowing. It is possible that future analyses, using more sophisticated statistical techniques, may be able to determine whether or not there was a true difference in gaze trajectories between younger and older participants.

Analysis of the response times showed an effect of an individual's inhibition ability. In this task, greater inhibition ability was related to faster response times when the context sentence was incongruent, compared to those with poorer inhibition ability. The finding that inhibition ability and processing speed (to a lesser extent) were predictive of response times provides some evidence that these cognitive abilities

contribute to real-time processing of speech in the presence of context (as in Grindrod & Raizen, 2019; Sommers & Danielson, 1999). If greater inhibition ability is related to faster processing of incongruent contexts, as evidenced by performance in this experiment, then those with greater inhibition ability likely have an advantage in conversational speech in challenging listening conditions over those with poorer inhibition ability (e.g., Wingfield, 1996; Wingfield et al., 2003). A misunderstanding in conversational speech can have far-reaching consequences both for conveying information and for interpersonal relationships. It should be noted that greater inhibition ability was not related to greater accuracy in target word categorization. Therefore, inhibition ability is only one component of the complex mechanism that supports the use of context in speech understanding.

The findings of this study appear to confirm the profound effects of sentence context on speech processing in listeners with CIs and listeners with AH, even when the context sentence is presented after the target word. There were no significant differences in the gaze trajectories across listener groups, SNRs, or age groups. Therefore, this study was unable to provide evidence to support a "wait-and-see" approach among listeners with CIs or among older listeners. All participants benefited from a congruent context. This fact lends support to the importance of context as a tool for understanding degraded speech for listeners with AH and listeners with CIs.

5 General Discussion

The overall objective of this dissertation was to examine the impact of age and acoustic degradation on the use of semantic context. The target word's position in the sentence and the amount of background noise were manipulated to determine the influence of these factors on an individual's use of semantic context. Two groups of participants were tested, both of whom heard spectrally degraded speech: adults with CIs, and adults with AH who listened to a simulation of CI-processed speech. Age, working memory, processing speed, and inhibition ability were all measured to determine the influence of these demographic and cognitive factors on an individual's use of semantic context. Performance on the categorization of phonetically distinct target words and response times were analyzed to determine whether the age and cognitive abilities of the participant influenced the size of the context effect on the participant's perception. Eye-tracking was also used to assess real-time speech processing when context followed the target word.

5.1 Context Use

Context affected the performance of adults with CIs and adults with AH, as revealed in Studies 1 and 2. The size of the context effect was influenced by the target word position and background noise level. The age of the participants and the cognitive abilities measured for these studies explained some of the variance in performance,

especially for listeners with CIs. In the third study that used eye-tracking, a strong context effect was also observed.

5.1.1 Target word position

Contrary to expectations, both adults with CIs and adults with AH were found to demonstrate a larger effect of context for sentence-initial target words than for sentence-final target words. The majority of studies evaluating the effect of context only tested preceding context's impact on sentence-final target words in listeners with CIs and AH (e.g., Amichetti et al., 2018; Connine, 1987; Kalikow et al., 1977; Patro & Mendel, 2016; Wagner et al., 2016; Winn, 2016). One previous study that assessed subsequent context found greater context effects from preceding context on phoneme categorization compared to following context in young listeners with AH (Connine et al., 1991). Another previous study that assessed subsequent context as a function of age found larger context effects on target words with preceding context cues compared to target words with subsequent context cues only for older listeners with AH (Wingfield et al., 1994). Differences in methodology and the calculation of context effects are the likely sources of inconsistent findings between the current study and the two earlier studies in listeners with AH (Connine et al., 1991; Wingfield et al., 1994). There have been no previous studies comparing the effects of preceding vs. subsequent context in listeners with CIs.

Connine et al. (1991) used a phoneme categorization task to assess the impact of context in young listeners with AH. The effect of preceding context that they

reported was consistent with the literature. They used sentences of increasing length and complexity to provide context-providing words three or six syllables away from the target word. They found small effects of subsequent context when it occurred within three syllables of the target word, but no effects of subsequent context when it occurred six syllables after the target word. They concluded that there may be a finite temporal window in which perception of an ambiguous word could be affected by subsequent context. However, the methodology of their study introduced a confound that likely affected their conclusions. They accepted responses at any time after the target word was played. This meant that participants could respond before the semantically related words in the sentence were played. The authors report that 84% of responses in their six-syllable subsequent context condition occurred prior to the context-providing key words. The current studies only accepted responses after all stimuli were presented to avoid the confound of participants potentially responding prior to hearing the context. This difference in methodology may be why the current studies found larger effects of subsequent context than for preceding context, in contrast to the findings of Connine et al. (1991).

Another way that the current studies differed from Connine et al. (1991) was in the method of calculating the context effect. Connine et al. calculated the context effect as the difference in categorization performance between sentences that biased the listener toward the /d/ interpretation of the target word and sentences that biased the listener toward the /t/ interpretation of the target word. In contrast, the current studies included a neutral context baseline that allowed statistical analyses to determine if the

context effect was greater than zero. An examination of Figure 2-2 or Figure 3-1 reveals that the difference between the two context conditions (depicted in red and blue) is always larger than or equal to the differences between the individual context conditions and the neutral context condition. Thus, the current studies demonstrate a more nuanced approach to calculating context effect by including an individual's performance in a neutral sentence context. This difference in how the context effect was calculated might contribute to differences in the significance of reported context effects across studies.

Wingfield et al. (1994) found that older listeners used preceding context in a similar manner and to a similar degree as younger listeners, but used subsequent context to a lesser degree than younger listeners. The current studies found that the effect of subsequent context was not significantly different between older and younger participants with AH. Wingfield et al. based their experimental design on Pickett and Pollack (1963) who extracted single words from the middle of fluent utterances and had listeners rate the word's intelligibility. With no change in the target word's acoustics, listeners consistently rated the target word as more intelligible with increasing numbers of surrounding words than when it was presented in isolation (Pickett & Pollack, 1963). The current studies used spectral degradation, background noise, and a phonetic continuum as stimuli rather than words that were extracted from conversational speech (as in Wingfield et al., 1994). Also, the stimuli in Wingfield et al.'s study did not control the proximity of semantically related words in the context sentence to the target word (e.g., "antlers" for "deer"). Thus, it is possible that the windowing effect found in young listeners by Connine et al. (1991) affected the older

listeners in Wingfield et al. (1994) more than the younger listeners. In contrast to both studies' stimuli, the predictive context sentences in the current studies always contained words that were semantically related to the target word within a few syllables of the target word. These differences in experimental tasks might explain the differences in findings for the effect of subsequent context as a function of age.

The relatively large context effect for words at the beginning of a sentence might reflect an adaptive strategy that is useful in conversational speech, especially for listeners with CIs. One feature of conversational speech is that one person often starts talking before the preceding talker has finished their sentence. This overlap in voices could easily cause a listener to miss the first words of the new talker. Even without temporal overlap of talkers, listeners with CIs have been shown to have difficulty discriminating VOT-contrasting speech sounds that occur soon after a non-predictive carrier phrase (e.g., Tinnemore et al., 2024; Xie et al., 2022). These studies provide converging evidence that listeners with CIs may be especially susceptible to missing sentence-initial words and sounds, and thus commonly rely on subsequent context to recover the intended message. It is also possible that there are other factors that influence how well a person remembers the first word of a sentence. Short- and intermediate-term memory interact with an individual's inhibition ability to determine the strength of the recency or primacy effects they experience (Greene et al., 2000). Thus, a large context effect for sentence-initial words may be a reflection of these cognitive abilities.

In the current studies, some participants anecdotally reported strategies of defaulting to a "deer" response whenever they doubted the acoustic sound. There is no way of knowing whether this and/or a different strategy was used by participants in the previous studies as well. One potential way to separate this strategy from a "true" response would be to have the participants indicate any responses that were purely guesses. A researcher would then be able to analyze the responses marked as guesses to determine if the participant was defaulting to a certain response when uncertain, alternating back and forth between the response options, or choosing according to the predictive context. It is likely that there could be many responses marked as a "guess" from any given participant, which may obscure the subconscious bias of predictive context. However, this could still be an excellent way to identify strategies used in everyday situations when words are missed in a conversation.

5.1.2 Age and Cognition

The listener's age was predictive of the amount of context benefit experienced by adults with CIs, but not by adults with AH. Most previous studies have found that older listeners with AH showed greater effects of context compared to younger listeners (e.g., Cohen & Faulkner, 1983; Madden, 1988; Pichora-Fuller, 2008; Wingfield et al., 1991, 1994). These age-related differences are usually attributed to compensatory strategies deliberately employed by older listeners to offset sensory deficits (e.g., Cohen & Faulkner, 1983; Madden, 1988), or simply an enhanced ability to utilize the redundancies and supportive structures in language efficiently (e.g., Pichora-Fuller,

2008; Wingfield et al., 1991). Age interacted with SNR, VOT, and target word position in Study 1, such that older listeners with AH showed shallower categorization psychometric functions at more difficult SNRs, especially for sentence-final target words, compared to younger listeners.

The presence of an age effect in listeners with CIs in Study 2 could be due to any one or more of the many ways in which the younger listeners with CIs differed from the older listeners with CIs. As discussed in Chapter 3, the younger participants with CIs included those who lost their hearing prior to learning language. Given that younger participants with CIs also showed smaller context effects than older participants, it is possible that these factors are related. Participants who learned language and spent years communicating via acoustic hearing showed larger context effects than those who learned language via a CI. Perhaps the experience of communicating via acoustic hearing is necessary to experience large context effects. On the other hand, age was highly predictive of response accuracy in the third study, especially for listeners with CIs. Consistent results across studies indicate that age significantly reduces accuracy in participants with CIs, leading to poorer performance on a phoneme categorization task.

The statistical analyses found no statistically significant differences between the gaze trajectories of older compared to younger participants. It is possible that older and younger participants processed speech similarly in this task. It is also possible that more sophisticated statistical techniques could reveal a difference not found in the present studies. Given the current findings, an argument can be made that younger and

older participants from both listener groups used similar strategies to make decisions in this task. This would support the findings of those who have found no significant age differences in the use of context to understand degraded speech (e.g., Pichora-Fuller, 2008).

The measures of cognitive abilities were inconsistent in their predictive abilities across studies. In Studies 1 and 3, processing speed was predictive of a participant's context effect (Study 1) and recognition accuracy (Study 3). In Study 2, working memory was predictive of the size of the context effect experienced by participants with CIs. Inhibition ability significantly predicted response times in both listener groups across studies. The finding that processing speed and working memory ability influence the size of the context effect in the current studies is consistent with the prediction that greater working memory ability and faster processing speeds would result in smaller context effects, especially for target words in the sentence-initial position. The current studies lend support to the theory that working memory, processing speed, and/or inhibition abilities are predictive of speech understanding performance when speech is degraded (e.g., Rönnberg et al., 2010; Sommers & Danielson, 1999; Wingfield, 1996).

It is possible that there are other cognitive abilities that may have an impact on the effect of context for an individual in addition to those assessed in the current study. It is also possible that adults with CIs may use a specific set of cognitive abilities to understand speech in noise that are different from those used by adults with AH. One potential measure of general cognition and problem-solving ability, Raven's

Progressive Matrices (Raven & Court, 1938), has been shown to be highly correlated with speech perception outcomes in listeners with CIs (e.g., Amini et al., 2023; Mattingly et al., 2018; Moberly et al., 2019; Zhan et al., 2020). It may also be beneficial to investigate other types of memory, such as short-term memory, that may play a role in accurately perceiving degraded speech. More research needs to be done to determine which, if any, additional cognitive abilities: (a) directly impact the effect of context; (b) can be quantified reliably; and (c) can be targeted for improvement by training or other therapeutic approaches.

5.1.3 Background noise level

Increasing amounts of background noise were associated with longer response times in all three studies, but the effect of background noise on context effect varied. In Studies 1 and 2, context effects were generally larger with increasing amounts of background noise. However, the level of background noise at which each individual experienced their greatest context effect varied widely. For most participants, if they experienced their maximum context effect in the -10 dB SNR condition for target words in one sentence position, then they experienced their maximum context effect at the same SNR condition for target words in the other sentence position. This represents a common strategy when the target words were extremely difficult to discern from the background noise. For those who didn't use this strategy, the SNR of a person's greatest context effect was usually not the same for sentence-initial and sentence-final target words. Neither the age of the listener, nor any of the other cognitive or demographic

measures observed, were correlated with the level of background noise that evoked an individual's greatest context effect. This is somewhat consistent with the notion that the greatest context effects are experienced when speech is moderately degraded (e.g., Bhandari et al., 2021). In general, background noise negatively affects speech understanding for listeners with CIs to a greater degree than for listeners with AH (e.g., Jin et al., 2020; Nelson et al., 2003; Perea Pérez et al., 2023; Zaltz et al., 2020). Thus, the participants with CIs in the current studies were likely displaying compensatory strategies in more of the background noise conditions than participants with AH.

In Study 3, the -5 dB SNR level of background noise (i.e., higher difficulty) prompted participants from both listener groups to respond almost exclusively with the word predicted by the context. This led to higher accuracy in the congruent condition compared to the easier level of background noise, and poorer accuracy in the incongruent condition. Far from indicating that more difficult levels of background noise lead to higher performance, these results highlight the benefits and detriments of adopting a strong reliance on context. When the acoustic target matches the context, as it usually does in conversational speech, relying on context allows better speech understanding than could be accomplished without context. When the acoustic target does not match the context, relying on context leads to poorer speech understanding when the context is misleading. Given that most speech occurs in neutral or congruent contexts, a reliance on context is an efficient strategy for maximizing communication success.

5.1.4 The Truth about Context Effects

While the use of context is often beneficial, extreme reliance on semantic context can occasionally be harmful. Thus far, the discussion has focused on the benefit of context and how listeners can improve their communication success by relying on the surrounding semantic context. There are, however, potential detriments of the use of context. Extreme reliance on semantic context, with little or no weight on deciphering the acoustics of the target word, will occasionally result in misunderstandings. These misunderstandings are likely to result in larger communication breakdowns than if the listener had simply missed a word and asked the talker to repeat it, because the listener may not know that they misheard the target word. By reconstructing the missed word in a way that makes sense in the semantic context, it is possible that significant time may pass in the conversation, accompanied by growing confusion, before the breakdown in communication is identified.

Another drawback to a reliance on context that can have a large impact on relationships is in the area of humor. Some types of humor, such as puns or wordplay, depend on an accurate interpretation of the acoustic target word that can often conflict with the semantic context. If a listener misses the key word of the joke and uses the semantic context to fill in the most logical word, rather than the word that the talker actually used, they may feel isolated when others in the group laugh.

One potential solution to these problems would involve explicit counseling on how to recognize when miscommunication has occurred and how to work with

communication partners to recover. In the case of a missed joke, there are immediate clues to the fact that an error in understanding has occurred (e.g., others chuckling). Communication partners can be counseled to use physical gestures or facial expressions to indicate humor. Another, more technical solution, would be to improve CI processing strategies such that phoneme-level information is encoded and transmitted more clearly.

5.2 Differences between electric and acoustic hearing (CI vs. AH)

In the current studies, listeners with CIs showed context effects that were consistently larger than those shown by listeners with AH. In addition, while context effects decreased with increasing SNR for both listener groups, the listeners with CIs were less affected by SNR overall (see Figure 3-5). The current set of studies used a measured approach to determine the effects of context cues at a small scale – specifically, on the perception of a single phoneme to cue word identity. The differences in the context effect across listener groups in the current studies are likely associated with differences in listening strategies, rather than differences in the ability to take advantage of context cues.

Differences in context effects across groups can be obscured by floor or ceiling effects when baseline speech understanding is not equal (e.g., Dingemanse & Goedegebure, 2019; O'Neill et al., 2019). Some recent studies have explicitly matched baseline performance across groups for better comparisons by using vocoded speech (e.g., Bhandari et al., 2021; O'Neill et al., 2021; Tinnemore et al., 2020). While

vocoding provides a rough equivalence to CI-processed speech, individual factors such as neural survival, electrode-to-neural interface, and neural plasticity from auditory deprivation are not captured in vocoded processing. In addition, and perhaps most importantly, listening to vocoded stimuli in the laboratory setting cannot compare to the daily experience of listening to CI-processed sound as the only input. Thus, use of vocoding may cause participants with AH to demonstrate uses of context that are inconsistent with their typical use of context. As discussed in Study 3, the use of vocoding may still provide participants with AH more access to the acoustic signal than participants with CIs receive through their processors.

Another common confounding factor in studies comparing listeners with CIs to those with AH is the age of the listener groups. Post-lingually deafened listeners with CIs are usually 50 years old and older. Control groups with normal hearing are often recruited from university student populations. Thus, differences in performance across listener groups may have more to do with age than listening status. O'Neill et al. (2021) directly addressed these confounds. These authors developed a corpus of sentences with or without semantic context and presented lists of sentences in quiet and at 0, +5, and +10 dB SNR to listeners with CIs, age-matched listeners with AH, and a group of younger listeners with AH. O'Neill et al. found no significant differences in the context benefit between age-matched listeners with AH or CIs, nor between younger vs. older listeners with AH. The interpretation of these results was that all listeners are equally able to take advantage of semantic context cues.

While the findings of the current studies do not match those of O'Neill et al. (2021), the current studies do not completely contradict their conclusions. Unlike O'Neill et al.'s (2021) study, the current studies did not attempt to match baseline performance across listener groups. The listeners with AH were presented stimuli that were spectrally degraded to a level roughly equivalent the spectral resolution available to an average listener with a CI (Friesen et al., 2001). This may be one reason that the size of the context effects was not equal across listener groups (see Figure 3-5). In addition, while all listeners may be equally able to use context, as is suggested by O'Neill et al. (2021), the findings of the current studies suggest that they may choose not to. Listeners are widely variable in their use of context and the strategies they use to understand speech in background noise. It is possible that the findings of O'Neill et al. (2021) are more reflective of the fact that listeners can use context, while the current studies are more reflective of the different ways that listeners choose to use context. The current studies were not designed to identify the strategies in use, but future research could identify and analyze the different strategies that listeners from each group employ, and determine the influence of those strategies on the size of the context effect. One such strategy, sometimes called "wait-and-see," has been observed during real-time processing of single words (e.g., McMurray et al., 2017, 2019; F. X. Smith & McMurray, 2022).

The findings of the third study failed to statistically support the use of the "wait-and-see" strategy. Listeners with CIs appeared to delay choosing a target word compared to listeners with AH in both levels of background noise, but this delay was

not statistically significant. Older participants showed slower response times than younger listeners in Studies 1 and 2. It is possible that this age difference is a reflection of slowed processing, both of motor movements and of decision making, in older adults (e.g., Salthouse, 1996). If a listener is attempting to find the most efficient strategy, a delay in committing to an interpretation of an unclear word could be beneficial. There are costs associated with making an incorrect identification of a word and later needing to revise that identification (e.g., Scheffers & Coles, 2000). On the other hand, a delay could be harmful by compounding delays in an individual's speech perception process.

5.3 Conclusion

Together, the findings of the experiments described in this dissertation support the benefits of relying on sentence context when seeking to maximize speech understanding. The age-related differences that were observed across the three studies appeared to be more indicative of strategic choices rather than any deficit in ability. The ability to benefit from semantic context does not appear to diminish with age. While the reliance on context generally increases with worsening listening environments, the spectrally degraded speech conveyed by CI processors is sufficient to provide the cues needed to benefit from semantic context.

The current studies found that context effects were significantly greater for target words at the beginning of a sentence than for those at the end of a sentence. This is a novel finding that fills a gap in the literature. These studies, assessing listeners' perception of discrete acoustic cues in different semantic contexts, were designed to

demonstrate how location of the target word in the sentence might affect the size of the context effect. Without a clear understanding of how subsequent context can affect the interpretation of an early word in the sentence, it might be difficult to understand some of the challenges that listeners encounter in conversational speech. This is especially true for those who hear a degraded representation of speech or are trying to understand speech in background noise. The current studies demonstrated that listeners of all ages and both listener groups resort to relying on context in these situations.

The third study measured real-time speech processing for two levels of background noise by adults with CIs and with AH. Participants overwhelmingly chose the word predicted by the context when the target was obscured by the more-difficult level of background noise. This highlights the fact that listeners rely on semantic context to maximize speech understanding and minimize errors in challenging conditions. The gaze trajectories measured in the third study showed large effects of subsequent context sentences on looks to the acoustic target word compared to the non-target distractor that was predicted by the sentence when the context was incongruent.

Overall, the findings of these studies confirm the effect that context has on speech understanding. These studies further expand on the current knowledge about the impact of subsequent context on a target word by comparing participants with CIs to participants with AH. Experience with degraded speech – either through age or through hearing loss – may prompt listeners to form a reliance on context without explicit instruction. However, some listeners may not naturally adopt this strategy and might benefit from counseling to promote its use. Therefore, these studies represent

progress toward improving the communication success and quality of life of younger and older adults with and without CIs.

Bibliography

Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., & Steggles, P. (1999). Towards a better understanding of context and context-awareness. In H.-W. Gellersen (Ed.), *Handheld and Ubiquitous Computing* (Vol. 1707, pp. 304–307). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-48157-5_29

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464. https://doi.org/10.1016/j.tics.2004.08.011

Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1515), 285–299. https://doi.org/10.1098/rstb.2008.0253

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. *International Journal of Audiology*, *47*(sup2), S53–S71. https://doi.org/10.1080/14992020802301142

Alain, C., & Woods, D. L. (1999). Age-related changes in processing auditory stimuli during visual attention: Evidence for deficits in inhibitory control and sensory memory. *Psychology and Aging*, *14*(3), 507–519.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439. https://doi.org/10.1006/jmla.1997.2558

Altmann, G. (1988). Ambiguity, parsing strategies, and computational models. *Language and Cognitive Processes*, *3*(2), 73–97. https://doi.org/10.1080/01690968808402083

Amichetti, N. M., Atagi, E., Kong, Y.-Y., & Wingfield, A. (2018). Linguistic context versus semantic competition in word recognition by younger and older adults with cochlear implants. *Ear and Hearing*, *39*(1), 101–109. https://doi.org/10.1097/AUD.0000000000000469

Amini, A. E., Naples, J. G., Hwa, T., Larrow, D. C., Campbell, F. M., Qiu, M., Castellanos, I., & Moberly, A. C. (2023). Emerging Relations among Cognitive Constructs and Cochlear Implant Outcomes: A Systematic Review and Meta-Analysis. *Otolaryngology–Head and Neck Surgery*, *169*(4), 792–810. https://doi.org/10.1002/ohn.344

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging affects neural precision of speech encoding. *Journal of Neuroscience*, *32*(41), 14156–14164. https://doi.org/10.1523/JNEUROSCI.2176-12.2012

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

Bacon, S. P., & Healy, E. W. (2000). Effects of ipsilateral and contralateral precursors on the temporal effect in simultaneous masking with pure tones. *The Journal of the Acoustical Society of America*, *107*(3), 1589–1597. https://doi.org/10.1121/1.428443

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging*, *12*, 12–21. https://doi.org/10.1037/0882-7974.12.1.12

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bazire, M., & Brézillon, P. (2005). Understanding context before using it. In A. Dey, B. Kokinov, D. Leake, & R. Turner (Eds.), *Modeling and Using Context* (Vol. 3554, pp. 29–40). Springer Berlin Heidelberg. https://doi.org/10.1007/11508373_3

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition*, *8*(6), 493–512. https://doi.org/10.3758/BF03213769

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., & Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech,*

Language, and Hearing Research, *54*(1), 243–262. https://doi.org/10.1044/1092-4388(2010/09-0233)

Ben-David, B. M., Erel, H., Goy, H., & Schneider, B. A. (2015). "Older is always better": Age-related differences in vocabulary scores across 16 years. *Psychology and Aging*, *30*(4), 856–862. https://doi.org/10.1037/pag0000051

Bhandari, P., Demberg, V., & Kray, J. (2021). Semantic predictability facilitates comprehension of degraded speech in a graded manner. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.714485

Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, R., Fraysse, B., Gallégo, S., Govaerts, P. J., Green, K., Huber, A. M., Kleine-Punte, A., Maat, B., Marx, M., Mawman, D., Mosnier, I., O'Connor, A. F., … Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology and Neurotology*, *18*(1), 36–47. https://doi.org/10.1159/000343189

Blomquist, C., Newman, R. S., Huang, Y. T., & Edwards, J. (2021). Children with cochlear implants use semantic prediction to facilitate spoken word recognition. *Journal of Speech, Language, and Hearing Research*, *64*(5), 1636–1649. https://doi.org/10.1044/2021_JSLHR-20-00319

Boisvert, I., Reis, M., Au, A., Cowan, R., & Dowell, R. C. (2020). Cochlear implantation outcomes in adults: A scoping review. *PLOS ONE*, *15*(5), e0232421. https://doi.org/10.1371/journal.pone.0232421

Boothroyd, A., & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *The Journal of the Acoustical Society of America*, *84*(1), 101–114. https://doi.org/10.1121/1.396976

Borsky, S., Tuller, B., & Shapiro, L. P. (1998). "How to milk a coat:" The effects of semantic and acoustic information on phoneme categorization. *The Journal of the Acoustical Society of America*, *103*(5), 2670–2676. https://doi.org/10.1121/1.422787

Bosen, A. K., Sevich, V. A., & Cannon, S. A. (2021). Forward digit span and word familiarity do not correlate with differences in speech recognition in individuals with cochlear implants after accounting for auditory resolution. *Journal of Speech, Language, and Hearing Research*, *64*(8), 3330–3342. https://doi.org/10.1044/2021_JSLHR-20-00574

Bronkhorst, A. W., Brand, T., & Wagener, K. (2002). Evaluation of context effects in sentence recognition. *The Journal of the Acoustical Society of America*, *111*(6), 2874–2886. https://doi.org/10.1121/1.1458025

Burke, D. M., & Peters, L. (1986). Word associations in old age: Evidence for consistency in semantic encoding during adulthood. *Psychology and Aging*, *1*, 283–292. https://doi.org/10.1037/0882-7974.1.4.283

Calandruccio, L., & Smiljanic, R. (2012). New Sentence Recognition Materials Developed Using a Basic Non-Native English Lexicon. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1342–1355. https://doi.org/10.1044/1092-4388(2012/11-0260)

Canlon, B., Illing, R. B., & Walton, J. (2010). Cell biology and physiology of the aging central auditory pathway. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), *The Aging Auditory System* (pp. 39–74). Springer. https://doi.org/10.1007/978-1-4419-0993-0_3

Castiglione, A., Benatti, A., Girasoli, L., Caserta, E., Montino, S., Pagliaro, M., Bovo, R., & Martini, A. (2015). Cochlear implantation outcomes in older adults. *Hearing, Balance and Communication*, *13*(2), 86–88. https://doi.org/10.3109/13625187.2015.1030885

CHABA. (1988). Speech understanding and aging. *The Journal of the Acoustical Society of America*, *83*(3), 859–895. https://doi.org/10.1121/1.395965

Cohen, G., & Faulkner, D. (1983). Word recognition: Age differences in contextual facilitation effects. *British Journal of Psychology*, *74*, 239–251. https://doi.org/10.1111/j.2044-8295.1983.tb01860.x

Colby, S. E., & McMurray, B. (2023). Efficiency of spoken word recognition slows across the adult lifespan. *Cognition*, *240*, 105588. https://doi.org/10.1016/j.cognition.2023.105588

Connine, C. M. (1987). Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory and Language*, *26*(5), 527–538. https://doi.org/10.1016/0749-596X(87)90138-0

Connine, C. M., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language*, *30*, 234–250.

Connine, C. M., & Clifton, C. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 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. *Cognitive Psychology*, *6*, 84–107. https://doi.org/10.1016/0010-0285(74)90005-X

Cruickshanks, K. J., Zhan, W., & Zhong, W. (2010). Epidemiology of age-related hearing impairment. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), *The Aging Auditory System* (pp. 259–274). Springer. https://doi.org/10.1007/978-1-4419-0993-0_9

Dagerman, K. S., MacDonald, M. C., & Harm, M. W. (2006). Aging and the use of context in ambiguity resolution: Complex changes from simple slowing. *Cognitive Science*, *30*(2), 311–345. https://doi.org/10.1207/s15516709cog0000_46

Degeest, S., Keppler, H., & Corthals, P. (2015). The effect of age on listening effort. *Journal of Speech, Language, and Hearing Research*, *58*(5), 1592–1600. https://doi.org/10.1044/2015_JSLHR-H-14-0288

Dillon, M. T., Buss, E., Adunka, M. C., King, E. R., Pillsbury, H. C., III, Adunka, O. F., & Buchman, C. A. (2013). Long-term speech perception in elderly cochlear implant users. *JAMA Otolaryngology–Head & Neck Surgery*, *139*(3), 279–283. https://doi.org/10.1001/jamaoto.2013.1814

Dingemanse, J. G., & Goedegebure, A. (2019). The important role of contextual information in speech perception in cochlear implant users and its consequences in speech tests. *Trends in Hearing*, *23*, 2331216519838672. https://doi.org/10.1177/2331216519838672

Dryden, A., Allen, H. A., Henshaw, H., & Heinrich, A. (2017). The Association Between Cognitive Performance and Speech-in-Noise Perception for Adult Listeners: A Systematic Literature Review and Meta-Analysis. *Trends in Hearing*, *21*, 2331216517744675. https://doi.org/10.1177/2331216517744675

Dubno, J. R., Ahlstrom, J. B., & Horwitz, A. R. (2000). Use of context by young and aged adults with normal hearing. *The Journal of the Acoustical Society of America*, *107*(1), 538–546. https://doi.org/10.1121/1.428322

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America*, *76*(1), 87–96. https://doi.org/10.1121/1.391011

Farris-Trimble, A., McMurray, B., Cigrand, N., & Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 308–327. https://doi.org/10.1037/a0034353

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, *42*(2), 133–141. https://doi.org/10.1111/j.1469-8986.2005.00274.x

Fitzgibbons, P. J., & Gordon-Salant, S. (1996). Auditory temporal processing in elderly listeners. *Journal of the American Academy of Audiology*, *7*, 183–189.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press; WorldCat.org. http://www.gbv.de/dms/bowker/toc/9780262060844.pdf

Forbes, S., Dink, J., & Ferguson, B. (2023). *eyetrackingR* (Version 0.2.1) [Computer software]. <http://www.eyetracking-r.com/>

Forli, F., Lazzerini, F., Fortunato, S., Bruschini, L., & Berrettini, S. (2019). Cochlear implant in the elderly: Results in terms of speech perception and quality of life. *Audiology and Neurotology*, *24*(2), 77–83. https://doi.org/10.1159/000499176

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, *110*(2), 1150–1163. https://doi.org/10.1121/1.1381538

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110–125. https://doi.org/10.1037/0096-1523.6.1.110

Gianakas, S. P., & Winn, M. B. (2019). Lexical bias in word recognition by cochlear implant listeners. *The Journal of the Acoustical Society of America*, *146*(5), 3373–3383. https://doi.org/10.1121/1.5132938

Golden, C. J. (2002). *Stroop color and word test: A manual for clinical and experimental uses.* WorldCat.org.

Goossens, T., Vercammen, C., Wouters, J., & Wieringen, A. van. (2016). Aging affects neural synchronization to speech-related acoustic modulations. *Frontiers in Aging Neuroscience*, *8*. https://doi.org/10.3389/fnagi.2016.00133

Gordon-Salant, S., & Cole, S. S. (2016). Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear & Hearing*, *37*(5), 593–602. https://doi.org/10.1097/AUD.0000000000000316

Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech, Language, and Hearing Research*, *36*(6), 1276–1285. https://doi.org/10.1044/jshr.3606.1276

Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *Journal of Speech, Language, and Hearing Research*, *42*(2), 300–311. https://doi.org/10.1044/jslhr.4202.300

Gordon-Salant, S., Frisina, R. D., Fay, R. R., & Popper, A. N. (Eds.). (2010). *The aging auditory system* (Vol. 34). Springer Science & Business Media. https://link.springer.com/book/10.1007/978-1-4419-0993-0

Gordon-Salant, S., Shader, M. J., & Wingfield, A. (2020). Age-related changes in speech understanding: Peripheral versus cognitive influences. In K. S. Helfer, E. L. Bartlett, A. N. Popper, & R. R. Fay (Eds.), *Aging and Hearing* (Vol. 72, pp. 199–230). Springer International Publishing. https://doi.org/10.1007/978-3-030-49367-7_9

Gordon-Salant, S., Yeni-Komshian, G., & Fitzgibbons, P. (2008). The role of temporal cues in word identification by younger and older adults: Effects of sentence context. *The Journal of the Acoustical Society of America*, *124*(5), 3249–3260. https://doi.org/10.1121/1.2982409

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Barrett, J. (2006). Age-related differences in identification and discrimination of temporal cues in speech segments. *The Journal of the Acoustical Society of America*, *119*(4), 2455–2466. https://doi.org/10.1121/1.2171527

Goupell, M. J., Gaskins, C. R., Shader, M. J., Walter, E. P., Anderson, S., & Gordon-Salant, S. (2017). Age-related differences in the processing of temporal envelope and spectral cues in a speech segment. *Ear and Hearing*, *38*(6), e335–e342. https://doi.org/10.1097/AUD.0000000000000447

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. *Journal of Speech, Language, and Hearing Research*, *56*(6), 1715–1732. https://doi.org/10.1044/1092-4388(2013/12-0053)

Greene, A. J., Prepscius, C., & Levy, W. B. (2000). Primacy Versus Recency in a Quantitative Model: Activity Is the Critical Distinction. *Learning & Memory*, *7*(1), 48–57. https://doi.org/10.1101/lm.7.1.48

Grindrod, C. M., & Raizen, A. L. (2019). Age-related changes in processing speed modulate context use during idiomatic ambiguity resolution. *Aging, Neuropsychology, and Cognition*, *26*(6), 842–864. https://doi.org/10.1080/13825585.2018.1537437

Guerreiro, M. J. S., Murphy, D. R., & Van Gerven, P. W. M. (2013). Making sense of age-related distractibility: The critical role of sensory modality. *Acta Psychologica*, *142*(2), 184–194. https://doi.org/10.1016/j.actpsy.2012.11.007

Haensel, J., Ilgner, J., Chen, Y.-S., Thuermer, C., & Westhofen, M. (2005). Speech perception in elderly patients following cochlear implantation. *Acta Oto-Laryngologica*, *125*(12), 1272–1276. https://doi.org/10.1080/00016480510044214

Hannemann, R., Obleser, J., & Eulitz, C. (2007). Top-down knowledge supports the retrieval of lexical information from degraded speech. *Brain Research*, *1153*, 134–143. https://doi.org/10.1016/j.brainres.2007.03.069

Harel, -Arbeli Tami, Wingfield, A., Palgi, Y., & Ben, -David Boaz M. (2021). Age-related differences in the online processing of spoken semantic context and

the effect of semantic competition: Evidence from eye gaze. *Journal of Speech, Language, and Hearing Research*, *64*(2), 315–327. https://doi.org/10.1044/2020_JSLHR-20-00142

Harel-Arbeli, T., Wingfield, A., Palgi, Y., & 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. *Journal of Speech, Language, and Hearing Research*, *64*(2), 315–327. https://doi.org/10.1044/2020_JSLHR-20-00142

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 163–169. https://doi.org/10.1037/0278-7393.17.1.163

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 22, pp. 193–225). Academic Press. https://doi.org/10.1016/S0079-7421(08)60041-9

Helfer, K. S., Bartlett, E. L., Popper, A. N., & Fay, R. R. (Eds.). (2020). *Aging and Hearing* (Vol. 72). Cham: Springer International Publishing (Springer Handbook of Auditory Research). https://link.springer.com/book/10.1007/978-3-030-49367-7

Helfer, K. S., & Wilber, L. A. (1990). Hearing loss, aging, and speech perception in reverberation and noise. *Journal of Speech, Language, and Hearing Research*, *33*(1), 149–155. https://doi.org/10.1044/jshr.3301.149

Huang, Y. T., Newman, R. S., Catalano, A., & Goupell, M. J. (2017). Using prosody to infer discourse prominence in cochlear-implant users and normal-hearing listeners. *Cognition*, *166*, 184–200. https://doi.org/10.1016/j.cognition.2017.05.029

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Attention, Perception, & Psychophysics*, *75*(3), 508–524. https://doi.org/10.3758/s13414-012-0406-9

Humes, L. E., & Dubno, J. R. (2010). Factors affecting speech understanding in older adults. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), *The Aging Auditory System* (pp. 211–257). Springer. https://doi.org/10.1007/978-1-4419-0993-0_8

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

Hunter, C. R., & Humes, L. E. (2022). Predictive sentence context reduces listening effort in older adults with and without hearing loss and with high and low working memory capacity. *Ear and Hearing*, *43*(4), 1164–1177. https://doi.org/10.1097/AUD.0000000000001192

Ito, A., & Knoeferle, P. (2022). Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behavior Research Methods*, *55*(7), 3461–3493. https://doi.org/10.3758/s13428-022-01969-3

Jaekel, B. N. (2020). *Effects of interrupting noise and speech repair mechanisms in adult cochlear-implant users*. http://hdl.handle.net/1903/26441

Jaekel, B. N., Newman, R. S., & Goupell, M. J. (2018). Age effects on perceptual restoration of degraded interrupted sentences. *The Journal of the Acoustical Society of America*, *143*(1), 84–97. https://doi.org/10.1121/1.5016968

Jaekel, B. N., Weinstein, S., Newman, R. S., & Goupell, M. J. (2021). Access to semantic cues does not lead to perceptual restoration of interrupted speech in cochlear-implant users. *The Journal of the Acoustical Society of America*, *149*(3), 1488–1497. https://doi.org/10.1121/10.0003573

Jin, S.-H., Liu, C., & Sladen, D. P. (2020). The Effects of Aging on Speech Perception in Noise: Comparison between Normal-Hearing and Cochlear-Implant Listeners. *Journal of the American Academy of Audiology*, *25*, 656–665. https://doi.org/10.3766/jaaa.25.7.4

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. https://psycnet.apa.org/doi/10.1037/0033-295X.99.1.122

Kaandorp, M. W., Smits, C., Merkus, P., Festen, J. M., & Goverts, S. T. (2017). Lexical-access ability and cognitive predictors of speech recognition in noise

in adult cochlear implant users. *Trends in Hearing*, *21*, 2331216517743887. https://doi.org/10.1177/2331216517743887

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, *61*(5), 1337–1351. https://doi.org/10.1121/1.381436

Kapnoula, E. C., & McMurray, B. (2021). Idiosyncratic use of bottom-up and top-down information leads to differences in speech perception flexibility: Converging evidence from ERPs and eye-tracking. *Brain and Language*, *223*, 105031. https://doi.org/10.1016/j.bandl.2021.105031

Kavé, G., & Halamish, V. (2015). Doubly blessed: Older adults know more vocabulary and know better what they know. *Psychology and Aging*, *30*, 68–73. https://doi.org/10.1037/a0038669

Kemper, S. (1992). Language and aging. In *The handbook of aging and cognition.* (pp. 213–270). Lawrence Erlbaum Associates, Inc.

Kirikae, I. (1969). Auditory function in advanced age with reference to histological changes in the central auditory system. *International Audiology*, *8*(2–3), 221–230. https://doi.org/10.3109/05384916909079063

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: Cochlear nerve degeneration after "temporary" noise-induced hearing loss. *The Journal of Neuroscience*, *29*(45), 14077–14085. https://doi.org/10.1523/JNEUROSCI.2845-09.2009

Kujawa, S. G., & Liberman, M. C. (2015). Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss. *Hearing Research*, *330*, 191–199. https://doi.org/10.1016/j.heares.2015.02.009

Lemke, U., & Besser, J. (2016). Cognitive load and listening effort: Concepts and age-related considerations. *Ear and Hearing*, *37*, 77S. https://doi.org/10.1097/AUD.0000000000000304

Lenarz, M., Sönmez, H., Joseph, G., Büchner, A., & Lenarz, T. (2012). Long-term performance of cochlear implants in postlingually deafened adults. *Otolaryngology–Head and Neck Surgery*, *147*(1), 112–118. https://doi.org/10.1177/0194599812438041

Lin, F. R. (2011). Hearing loss and cognition among older adults in the united states. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *66A*(10), 1131–1136. https://doi.org/10.1093/gerona/glr115

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2010). Effects of semantic context and feedback on perceptual learning of speech processed through an acoustic simulation of a cochlear implant. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(1), 224. https://doi.org/10.1037/a0017609

Loizou, P. C. (1999). Introduction to cochlear implants. *IEEE Engineering in Medicine and Biology Magazine*, *18*(1), 32–42. IEEE Engineering in Medicine and Biology Magazine. https://doi.org/10.1109/51.740962

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36.

Macherey, O., & Carlyon, R. P. (2014). Cochlear implants. *Current Biology*, *24*(18), R878–R884. https://doi.org/10.1016/j.cub.2014.06.053

Madden, D. J. (1988). Adult age differences in the effects of sentence context and stimulus degradation during visual word recognition. *Psychology and Aging*, *3*(2), 167–172. https://doi.org/10.1037/0882-7974.3.2.167

Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., & Merchant, S. N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. *Journal of the Association for Research in Otolaryngology*, *12*(6), 711–717. https://doi.org/10.1007/s10162-011-0283-2

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1–71. https://doi.org/10.1016/0010-0277(80)90015-3

Martin, I. A., Goupell, M. J., & Huang, Y. T. (2022). Children's syntactic parsing and sentence comprehension with a degraded auditory signal. *The Journal of the Acoustical Society of America*, *151*(2), 699–711. https://doi.org/10.1121/10.0009271

Mattingly, J. K., Castellanos, I., & Moberly, A. C. (2018). Nonverbal reasoning as a contributor to sentence recognition outcomes in adults with cochlear implants. *Otology & Neurotology : Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *39*(10), e956–e963. https://doi.org/10.1097/MAO.0000000000001998

Mattys, S. L., & Scharenborg, O. (2014). Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors. *Psychology and Aging*, *29*(1), 150–162. https://doi.org/10.1037/a0035387

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

McMurray, B., Ellis, T. P., & Apfelbaum, K. S. (2019). How do you deal with uncertainty? Cochlear Implant users differ in the dynamics of lexical processing of non-canonical inputs. *Ear and Hearing*, *40*(4), 961–980. https://doi.org/10.1097/AUD.0000000000000681

McMurray, B., Farris-Trimble, A., & Rigler, H. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition*, *169*, 147–164. https://doi.org/10.1016/j.cognition.2017.08.013

McRackan, T., Bauschard, M., Hatch, J., Franko-Tobin, E., Droghini, H. R., Velozo, C. A., Nguyen, S. A., & Dubno, J. R. (2018). Meta-analysis of cochlear implantation outcomes evaluated with general health-related patient-reported outcome measures. *Otology & Neurotology : Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *39*(1), 29–36. https://doi.org/10.1097/MAO.0000000000001620

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, *41*(5), 329–335. https://doi.org/10.1037/h0062491

Moberly, A. C., Lewis, J. H., Vasil, K. J., Ray, C., & Tamati, T. N. (2021). Bottom-up signal quality impacts the role of top-down cognitive-linguistic processing during speech recognition by adults with cochlear implants. *Otology & Neurotology*, *42*(10S), S33. https://doi.org/10.1097/MAO.0000000000003377

Moberly, A. C., Mattingly, J. K., & Castellanos, I. (2019). How does nonverbal reasoning affect sentence recognition in adults with cochlear implants and normal-hearing peers? *Audiology & Neuro-Otology*, *24*(3), 127–138. https://doi.org/10.1159/000500699

Moberly, A. C., & Reed, J. (2019). Making sense of sentences: Top-down processing of speech by adult cochlear implant users. *Journal of Speech, Language, and Hearing Research : JSLHR*, *62*(8), 2895–2905. https://doi.org/10.1044/2019_JSLHR-H-18-0472

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165. https://doi.org/10.1037/h0027366

Moss, H. E., & Marslen-Wilson, W. D. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1254–1276. https://doi.org/10.1037/0278-7393.19.6.1254

Murr, A. T., Canfarotta, M. W., O'Connell, B. P., Buss, E., King, E. R., Bucker, A. L., Dillon, S. A., Rooth, M. A., Dedmon, M. M., Brown, K. D., & Dillon, M. T. (2021). Speech recognition as a function of age and listening experience in adult cochlear implant users. *The Laryngoscope*, *131*(9), 2106–2111. https://doi.org/10.1002/lary.29663

Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-level information and high-level perception: The case of speech in noise. *PLOS Biology*, *6*(5), e126. https://doi.org/10.1371/journal.pbio.0060126

Nelken, I., & Ahissar, M. (2006). High-level and low-level processing in the auditory system: The role of primary auditory cortex. *Dynamics of Speech Production and Perception, 374,* 343–353.

Nelson, P. B., Jin, S.-H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *113*(2), 961–968. https://doi.org/10.1121/1.1531983

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America*, *87*(6), 2705–2715. https://doi.org/10.1121/1.399061

Obleser, J., & Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cerebral Cortex*, *20*(3), 633–640. https://doi.org/10.1093/cercor/bhp128

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

Obleser, J., Wise, R. J. S., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience*, *27*(9), 2283–2289. https://doi.org/10.1523/JNEUROSCI.4663-06.2007

O'Neill, E. R., Kreft, H. A., & Oxenham, A. J. (2019). Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *The Journal of the Acoustical Society of America*, *146*(1), 195–210. https://doi.org/10.1121/1.5116009

O'Neill, E. R., Parke, M. N., Kreft, H. A., & Oxenham, A. J. (2020). Development and validation of sentences without semantic context to complement the basic english lexicon sentences. *Journal of Speech, Language, and Hearing Research*, *63*(11), 3847–3854. https://doi.org/10.1044/2020_JSLHR-20-00174

O'Neill, E. R., Parke, M. N., Kreft, H. A., & Oxenham, A. J. (2021). Role of semantic context and talker variability in speech perception of cochlear-implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *149*(2), 1224–1239. https://doi.org/10.1121/10.0003532

Otte, J., Schuknecht, H. F., & Kerr, A. G. (1978). Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. *The Laryngoscope*, *88*(8), 1231–1246. https://doi.org/10.1288/00005537-197808000-00004

Parthasarathy, A., Bartlett, E. L., & Kujawa, S. G. (2019). Age-related changes in neural coding of envelope cues: Peripheral declines and central compensation. *Neuroscience*, *407*, 21–31. https://doi.org/10.1016/j.neuroscience.2018.12.007

Parthasarathy, A., Herrmann, B., & Bartlett, E. L. (2019). Aging alters envelope representations of speech-like sounds in the inferior colliculus. *Neurobiology of Aging*, *73*, 30–40. https://doi.org/10.1016/j.neurobiolaging.2018.08.023

Patro, C., & Mendel, L. L. (2016). Role of contextual cues on the perception of spectrally reduced interrupted speech. *The Journal of the Acoustical Society of America*, *140*(2), 1336–1345. https://doi.org/10.1121/1.4961450

Patro, C., & Mendel, L. L. (2018). Gated word recognition by postlingually deafened adults with cochlear implants: Influence of semantic context. *Journal of Speech, Language, and Hearing Research*, *61*(1), 145–158. https://doi.org/10.1044/2017_JSLHR-H-17-0141

Patro, C., & Mendel, L. L. (2020). Semantic influences on the perception of degraded speech by individuals with cochlear implants. *The Journal of the Acoustical Society of America*, *147*(3), 1778–1789. https://doi.org/10.1121/10.0000934

Perea Pérez, F., Hartley, D. E. H., Kitterick, P. T., Zekveld, A. A., Naylor, G., & Wiggins, I. M. (2023). Listening efficiency in adult cochlear-implant users compared with normally-hearing controls at ecologically relevant signal-to-noise ratios. *Frontiers in Human Neuroscience*, *17*. https://doi.org/10.3389/fnhum.2023.1214485

Phillips, N. A. (2016). The implications of cognitive aging for listening and the framework for understanding effortful listening (fuel). *Ear and Hearing*, *37*, 44S. https://doi.org/10.1097/AUD.0000000000000309

Pichora-Fuller, M. K. (2003). Processing speed and timing in aging adults: Psychoacoustics, speech perception, and comprehension. *International Journal of Audiology*, *42*(sup1), 59–67. https://doi.org/10.3109/14992020309074625

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, *47*(sup2), S72–S82. 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. *The Journal of the Acoustical Society of America*, *97*(1), 593–608. https://doi.org/10.1121/1.412282

Pickett, J. M., & Pollack, I. (1963). Intelligibility of excerpts from fluent speech: Effects of rate of utterance and duration of excerpt. *Language and Speech, 6*(3), 151–164. https://doi.org/10.1177/002383096300600304

Presacco, A., Simon, J. Z., & Anderson, S. (2016). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *Journal of Neurophysiology, 116*(5), 2346–2355. https://doi.org/10.1152/jn.00372.2016

Presacco, A., Simon, J. Z., & Anderson, S. (2019). Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLOS ONE, 14*(3), e0213899. https://doi.org/10.1371/journal.pone.0213899

R Core Team. (2024). *R: a language and environment for statistical computing* (Version 4.3.3) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rasmussen, K. M. B., West, N. C., Bille, M., Sandvej, M. G., & Cayé-Thomasen, P. (2022). Cochlear implantation improves both speech perception and patient-reported outcomes: A prospective follow-up study of treatment benefits among adult cochlear implant recipients. *Journal of Clinical Medicine, 11*(8), Article 8. https://doi.org/10.3390/jcm11082257

Raven, J. C., & Court, J. H. (1938). *Raven's progressive matrices*. Western Psychological Services.

Revill, K. P., & Spieler, D. H. (2012). The effect of lexical frequency on spoken word recognition in young and older listeners. *Psychology and Aging, 27*(1), 80–87. https://doi.org/10.1037/a0024113

Rönnberg, J., Rudner, M., Lunner, T., & Zekveld, A. A. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise and Health*, *12*(49), 263–269. https://doi.org/10.4103/1463-1741.70505

Rudner, M., Rönnberg, J., & Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *Journal of the American Academy of Audiology*, *22*(3), 156–167. https://doi.org/10.3766/jaaa.22.3.4

Rush, B. K., Barch, D. M., & Braver, T. S. (2006). Accounting for cognitive aging: Context processing, inhibition or processing speed? *Aging, Neuropsychology, and Cognition*, *13*(3–4), 588–610. https://doi.org/10.1080/13825580600680703

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*(1), 35–54. https://doi.org/10.1016/S0301-0511(00)00052-1

Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00557

Scheffers, M., & Coles, M. (2000). Scheffers MK, Coles MG. Performance monitoring in a confusing world: Error related brain activity, judgements of response accuracy, and types of errors. J Exp Psychol Hum Percept Perform 26: 141-151. *Journal of Experimental Psychology. Human Perception and Performance*, *26*, 141–151. https://doi.org/10.1037/0096-1523.26.1.141

Schmiedt, R. A. (2010). The physiology of cochlear presbycusis. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), *The Aging Auditory System* (pp. 9–38). Springer. https://doi.org/10.1007/978-1-4419-0993-0_2

Schubotz, L., Holler, J., Drijvers, L., & Özyürek, A. (2021). Aging and working memory modulate the ability to benefit from visible speech and iconic gestures during speech-in-noise comprehension. *Psychological Research*, *85*(5), 1997–2011. https://doi.org/10.1007/s00426-020-01363-8

Schuknecht, H. F., & Gacek, M. R. (1993). Cochlear pathology in presbycusis. *Annals of Otology, Rhinology & Laryngology*, *102*(1_suppl), 1–16. https://doi.org/10.1177/00034894931020S101

Shader, M. J., Gordon-Salant, S., & 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 in Hearing*, *24*, 2331216520936160. https://doi.org/10.1177/2331216520936160

Shader, M. J., Yancey, C. M., Gordon-Salant, S., & Goupell, M. J. (2020). Spectral-temporal trade-off in vocoded sentence recognition: Effects of age, hearing thresholds, and working memory. *Ear and Hearing*, *41*(5), 1226–1235. https://doi.org/10.1097/AUD.0000000000000840

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *The Journal of the Acoustical Society of America*, *123*(1), 489–499. https://doi.org/10.1121/1.2783762

Shorey, A. E., & Stilp, C. E. (2023). Short-term, not long-term, average spectra of preceding sentences bias consonant categorization. *The Journal of the Acoustical Society of America*, *153*(4), 2426. https://doi.org/10.1121/10.0017862

Signoret, C., Andersen, L. M., Dahlström, Ö., Blomberg, R., Lundqvist, D., Rudner, M., & Rönnberg, J. (2020). The influence of form- and meaning-based predictions on cortical speech processing under challenging listening conditions: A meg study. *Frontiers in Neuroscience*, *14*. https://www.frontiersin.org/articles/10.3389/fnins.2020.573254

Smith, F., & McMurray, B. (2018). Lexical access in the face of degraded speech: The effects of cognitive adaptation. *The Journal of the Acoustical Society of America*, *144*(3), 1800–1800. https://doi.org/10.1121/1.5067944

Smith, F. X., & McMurray, B. (2022). Lexical Access Changes Based on Listener Needs: Real-Time Word Recognition in Continuous Speech in Cochlear Implant Users. *Ear and Hearing*, *43*(5), 1487. https://doi.org/10.1097/AUD.0000000000001203

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, *32*(25), 8443–8453. https://doi.org/10.1523/JNEUROSCI.5069-11.2012

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. *Psychology and Aging*, *14*(3), 458–472. https://doi.org/10.1037/0882-7974.14.3.458

Speranza, F., Daneman, M., & Schneider, B. A. (2000). How aging affects the reading of words in noisy backgrounds. *Psychology and Aging*, *15*(2), 253–258. https://doi.org/10.1037/0882-7974.15.2.253

Stenbäck, V., Marsja, E., Hällgren, M., Lyxell, B., & Larsby, B. (2021). The contribution of age, working memory capacity, and inhibitory control on speech recognition in noise in young and older adult listeners. *Journal of Speech, Language, and Hearing Research*, *64*(11), 4513–4523. https://doi.org/10.1044/2021_JSLHR-20-00251

Stine, E. L., & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging*, *2*(3), 272–279. https://doi.org/10.1037/0882-7974.2.3.272

Stine-Morrow, E. A. L., Miller, L. M. S., & Nevin, J. A. (1999). The effects of context and feedback on age differences in spoken word recognition. *The Journals of Gerontology: Series B*, *54B*(2), P125–P134. https://doi.org/10.1093/geronb/54B.2.P125

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Svirsky, M. A., Teoh, S.-W., & Neuburger, H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of

age at cochlear implantation. *Audiology and Neurotology*, *9*(4), 224–233. https://doi.org/10.1159/000078392

Takahashi, G. A., & Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech, Language, and Hearing Research*, *35*(6), 1410–1421. https://doi.org/10.1044/jshr.3506.1410

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Tinnemore, A. R., Doyle, E., & Goupell, M. J. (2024). Temporal speech cue perception in listeners with cochlear implants depends on the time between those cues and previous sound energy. *JASA*, *156*(3). https://doi.org/10.1121/10.0029020

Tinnemore, A. R., Gordon-Salant, S., & Goupell, M. J. (2020). Audiovisual speech recognition with a cochlear implant and increased perceptual and cognitive demands. *Trends in Hearing*, *24*, 2331216520960601. https://doi.org/10.1177/2331216520960601

Tremblay, K. L., Piskosz, M., & Souza, P. (2002). Aging alters the neural representation of speech cues. *NeuroReport*, *13*(15), 1865–1870.

Tulsky, D. S., Carlozzi, N., Chiaravalloti, N. D., Beaumont, J. L., Kisala, P. A., Mungas, D., Conway, K., & Gershon, R. (2014). NIH Toolbox Cognition

Battery (NIHTB-CB): List Sorting Test to Measure Working Memory. *Journal of the International Neuropsychological Society*, *20*(6), 599–610. https://doi.org/10.1017/S135561771400040X

Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761–766. https://doi.org/10.1037/a0014802

Tye-Murray, N., Spencer, L., & Woodworth, G. G. (1995). Acquisition of speech by children who have prolonged cochlear implant experience. *Journal of Speech, Language, and Hearing Research*, *38*(2), 327–337. https://doi.org/10.1044/jshr.3802.327

Van Engen, K. J., Dey, A., Runge, N., Spehar, B., Sommers, M. S., & Peelle, J. E. (2020). Effects of age, word frequency, and noise on the time course of spoken word recognition. *Collabra. Psychology*, *6*(1), 17247. https://doi.org/10.1525/collabra.17247

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, *18*, 332–339. https://doi.org/10.1037/0882-7974.18.2.332

Voeten, C. C. (2023). *buildmer: Stepwise elimination and term reordering for mixed-effects regression.* (Version R package version 2.11) [R]. <https://CRAN.R-project.org/package=buildmer>

Wagner, A. E., Toffanin, P., & Başkent, D. (2016). The timing and effort of lexical access in natural and degraded speech. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00398

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. *The Journal of Neuroscience*, *32*(40), 14010–14021. https://doi.org/10.1523/JNEUROSCI.1528-12.2012

Willott, J. F. (1991). Central physiological correlates of ageing and presbycusis in mice. *Acta Oto-Laryngologica*, *111*(sup476), 153–156. https://doi.org/10.3109/00016489109127271

Willott, J. F. (1996). Anatomic and physiologic aging: A behavioral neuroscience perspective. *Journal of the American Academy of Audiology*, *7*(3), 141–151.

Wingfield, A. (1996). Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. *Journal of the American Academy of Audiology*, *7*(3), 175–182.

Wingfield, A., Aberdeen, J. S., & Stine, E. A. L. (1991). Word Onset Gating and Linguistic Context in Spoken Word Recognition by Young and Elderly Adults. *Journal of Gerontology*, *46*(3), P127–P129. https://doi.org/10.1093/geronj/46.3.P127

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

Wingfield, A., Peelle, J. E., & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and

older adults. *Aging, Neuropsychology, and Cognition*, *10*(4), 310–322. https://doi.org/10.1076/anec.10.4.310.28974

Wingfield, A., & Tun, P. A. (2001). Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal aging. *Seminars in Hearing*, *22*(3), 287–302. https://doi.org/10.1055/s-2001-15632

Winn, M. B. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, *20*, 2331216516669723. https://doi.org/10.1177/2331216516669723

Winn, M. B., & Litovsky, R. Y. (2015). Using speech sounds to test functional spectral resolution in listeners with cochlear implants. *The Journal of the Acoustical Society of America*, *137*(3), 1430–1442. https://doi.org/10.1121/1.4908308

Winn, M. B., & Moore, A. N. (2018). Pupillometry reveals that context benefit in speech perception can be disrupted by later-occurring sounds, especially in listeners with cochlear implants. *Trends in Hearing*, *22*, 2331216518808962. https://doi.org/10.1177/2331216518808962

Wotton, J. M., Elvebak, R. L., Moua, L. C., Heggem, N. M., Nelson, C. A., & Kirk, K. M. (2011). Congruent and incongruent semantic context influence vowel recognition. *Language and Speech*, *54*(3), 341–360. https://doi.org/10.1177/0023830911402476

Wouters, J., McDermott, H. J., & Francart, T. (2015). Sound coding in cochlear implants: From electric pulses to hearing. *IEEE Signal Processing Magazine*,

32(2), 67–80. IEEE Signal Processing Magazine. https://doi.org/10.1109/MSP.2014.2371671

Xie, Z., Anderson, S., & Goupell, M. J. (2022). Stimulus context affects the phonemic categorization of temporally based word contrasts in adult cochlear-implant users. *The Journal of the Acoustical Society of America*, *151*(3), 2149–2158. https://doi.org/10.1121/10.0009838

Xie, Z., Gaskins, C. R., Shader, M. J., Gordon-Salant, S., Anderson, S., & Goupell, M. J. (2019). Age-related temporal processing deficits in word segments in adult cochlear-implant users. *Trends in Hearing*, *23*, 233121651988668. https://doi.org/10.1177/2331216519886688

Yang, Z., & Cosetti, M. (2016). Safety and outcomes of cochlear implantation in the elderly: A review of recent literature. *Journal of Otology*, *11*(1), 1–6. https://doi.org/10.1016/j.joto.2016.03.004

Zaltz, Y., Bugannim, Y., Zechoval, D., Kishon-Rabin, L., & Perez, R. (2020). Listening in Noise Remains a Significant Challenge for Cochlear Implant Users: Evidence from Early Deafened and Those with Progressive Hearing Loss Compared to Peers with Normal Hearing. *Journal of Clinical Medicine*, *9*(5), Article 5. https://doi.org/10.3390/jcm9051381

Zekveld, A. A., Rudner, M., Johnsrude, I. S., & Rönnberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America*, *134*(3), 2225–2234. https://doi.org/10.1121/1.4817926

Zeng, F.-G. (2022). Celebrating the one millionth cochlear implant. *JASA Express Letters*, *2*(7), 077201. https://doi.org/10.1121/10.0012825

Zhan, K. Y., Lewis, J. H., Vasil, K. J., Tamati, T. N., Harris, M. S., Pisoni, D. B., Kronenberger, W. G., Ray, C., & Moberly, A. C. (2020). Cognitive functions in adults receiving cochlear implants: Predictors of speech recognition and changes after implantation. *Otology & Neurotology*, *41*(3), e322. https://doi.org/10.1097/MAO.0000000000002544