Sensory and Cognitive Factors Underlying Self-Perceived Listening Difficulties in Adults with Normal Hearing Thresholds

by

Jacie R. McHaney

B.S. Psychology, University of Texas at Austin, 2015

Submitted to the Graduate Faculty of the

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Jacie R. McHaney

It was defended on

March 17, 2023

and approved by

Aravindakshan Parthasarathy, Ph.D.; Assistant Professor, Department of Communication Science and Disorders, University of Pittsburgh

Amanda Hampton Wray, Ph.D., CCC-SLP, Assistant Professor, Department of Communication Science and Disorders, University of Pittsburgh

G. Nike Gnanateja, Ph.D., Assistant Professor, Department of Communication Science and Disorders, University of Wisconsin-Madison

Dissertation Director: Bharath Chandrasekaran, Ph.D., F-ASHA, Professor, Department of Communication Science and Disorders, University of Pittsburgh

Sensory and Cognitive Factors Underlying Individual Variability in Self-Perceived Listening Difficulties in Adults with Normal Hearing Thresholds

Jacie R. McHaney, Ph.D.

University of Pittsburgh, 2023

One in ten adults seeking help at audiology clinics present with a primary complaint of listening difficulties yet have normal hearing thresholds. These adults with complaints of listening difficulties particularly struggle with speech perception in noise, which requires a combination of bottom-up and top-down processes. Prior research has largely focused on dysfunctions in bottom-up processing, yet it remains unclear how top-down processes, like decisional and linguistic processes, may contribute to self-perceived listening difficulties (SPLDs) in adults with normal hearing thresholds. Therefore, there is a critical need to understand the extent to which top-down decisional and linguistic processes during speech perception underlie SPLDs in order for clinicians to provide targeted interventions for SPLDs. In this dissertation, I examined SPLDs in adults ages 18-53 with normal hearing thresholds across two studies. In Study 1, I examined decisional processes that support speech categorization in a phoneme in noise categorization task using a computational modeling approach. I found that the rate at which listeners accumulate critical sensory evidence to make a categorization decision increases when there is less interference from background noise. Moreover, individuals with fewer SPLDs benefit from more supportive listening contexts to a greater extent than those with more SPLDs. In study 2, I examined the extent to which listeners with more SPLDs relied on different types of linguistic information to aid speech comprehension using more naturalistic stimuli in an electroencephalography experiment. I found that speech comprehension performance did not differ based on SPLDs. However, listeners with

more SPLDs had increased representations of sentence-level information during listening, relative to those with fewer SPLDs. These findings suggest that listeners with more SPLDs may rely on sentence-level information as a compensatory strategy to aid comprehension performance. Taken together, the findings from this dissertation demonstrate that listeners with more SPLDs have different approaches to listening. These results encourage further investigation into SPLDs in adults with normal hearing thresholds.

1. Table of Contents

List of Tables

List of Figures

Preface

Science is funny because in the pursuit for answers, we end up with more questions. During my pursuit for answers over the last four years, I have received invaluable opportunities that I will forever be grateful for. This dissertation research would not have been possible without the support from the mentors and colleagues who have helped me grow as a scientist, and who provided me with unwavering encouragement, constructive feedback, and instruction. In the paragraphs that follow, I acknowledge all those who have supported my PhD pursuit and research goals.

My PhD mentors were integral to my professional development and success over the last few years. Dr. Bharath Chandrasekaran provided the best mentorship, support, and guidance possible. Thank you for always setting the bar high. Additionally, Dr. Aravind Parthasarathy contributed significantly to my growth as an independent researcher. Thank you for your generosity and guidance—and for making scientific discussions fun and engaging. Dr. Nike Gnanateja also provided critical mentorship throughout my training. I owe my electroencephalography analytical skills to you! I look forward to considering you all my fellow colleagues in the years to come.

I am incredibly grateful for the funding that supported this dissertation research and the funding that supported me throughout my training. This dissertation work was supported by the National Institutes of Health (NIH) Ruth L. Kirschstein National Research Service Award Individual Predoctoral Fellowship to Promote Diversity in Health-Related Research (F31DC020085-01A1) and the University of Pittsburgh's Clinical and Translational Science Institute's Quantitative Methodologies Program through the NIH grant UL1TR001857. Additionally, the Auditory and Vestibular Neuroscience Training Program through NIH grant

T32DC011499 supported early work in my PhD training that led to the culmination of this dissertation research. This work would also not have been possible without the support from the University of Pittsburgh Center for Research Computing, RRID:SCR_022735, through the resources provided. Specifically, this dissertation used the HTC cluster, which is supported by NIH award number S10OD028483.

These last few years have been some of the hardest, yet most rewarding of my life. During my PhD, I endured some major life events: I lost one of my sisters to a terrible disease, and I became a mom. My success and ability to persevere through all of this would not have been possible without the everlasting support from those dearest to me. The biggest thank you is owed to my husband, Andrew. Thank you to the end of time for always being there for me through the highs and lows in life. You handled this emotional rollercoaster with grace, and I truly would not have made it this far without you. Thank you to my son, Malcolm, for being the brightest joy in my life, the best part of my day, always. Becoming a mom while a PhD student was no easy feat, but I would not have it any other way. You somehow made me work harder and lit a new fire within me to achieve a PhD for you. Thank you to my parents, Chris and Nancy, my sister, Jennifer, my brother-in-law, Rodney, and my niece and nephew, Jay and Olivia, for always being so proud of me and accepting that I needed to move away to achieve my goals. Living so far from you all has not been easy, but it has made the time we do spend together that much sweeter. To my sister, Jamie, who passed away in 2021: Thank you for always being my cheerleader and bragging about me to anyone who would listen, even though you had no idea what I do. It always makes me laugh when I think of you saying, “I have no idea what she does, but it’s something with brains.” I know you would be very proud.

On a less serious note, I cannot forget to acknowledge my friends and hobbies that helped me remain level-headed through these past four years. I am incredibly lucky to have been surrounded by a large cohort of fellow CSD PhD students who were supportive, funny, and friendly. Specifically, I am grateful to have shared this PhD journey with Rob Cavanaugh, Emily Goldberg, and Brett Welch. Thank you for your advice, the countless coworking hours, and for making the work-life balance possible! I also feel very fortunate to have pursued my PhD at the University of Pittsburgh. Moving across the country from Texas opened my world to the domain of winter sports that I never was able to experience before. Skiing became my passion and favorite hobby. A day spent skiing was the best way for me to reset, refresh, and regroup to tackle my research with new energy.

Finally, thank you to the many hard-working graduate students, research assistants, and research staff across the SoundBrain and Translational Auditory Neuroscience Laboratories: Leslie Zhen, Maggie Zink, Olivia Flemm, Jennifer Klara, Claire Mitchell, Sarah Anthony, Kim Yurasits, Megan Hallihan, Katie Bergstrom, Shaina Wasileski, Miaofang Hu, and Jasmine Cardino. Science would not be possible without the hard work that research personnel provide!

1.0 Background

Processing speech in noisy listening situations is a complex skill that most individuals encounter on a near-daily basis. Advancing age and hearing loss often lead to difficulties with speech perception in noise (Dinino et al., 2021; Helfer & Wilber, 1990; Peelle et al., 2011). However, even adults without overt hearing loss can struggle with understanding speech in noisy environments. When a client presents to an audiology clinic with a primary complaint of these hearing difficulties, they are typically worked through the standard audiometric battery to assess peripheral hearing and speech reception thresholds. However, approximately ten percent of the clients who undergo this battery are provided a 'clean bill' of hearing health (Hind et al., 2011; Parthasarathy et al., 2020; Pryce & Wainwright, 2008; Spankovich et al., 2018; Tremblay et al., 2015). It can be very upsetting for a client to learn that there is a disconnect between their self-perceived listening difficulties (SPLDs) and their hearing thresholds. Listening difficulties can impact the client's personal relationships, overall well-being, personal productivity, and can lead to social isolation (Bainbridge & Wallhagen, 2014; Huddle et al., 2017; Lin & Albert, 2014; Pryce & Wainwright, 2008). Therefore, it is critical to understand the underlying physiology of SPLDs in adults with clinically normal hearing thresholds. A better understanding of the processes that drive SPLDs will help to develop a scalable, clinically viable behavioral test for SPLDs that can complement the current audiometric battery and lead to individualized, rehabilitative approaches for SPLDs.

Individuals with SPLDs tend to particularly struggle with speech perception in noisy listening situations. Speech perception in noise involves a combination of bottom-up sensory encoding and top-down cognitive resources for accurate perception. However, speech perception tests that are regularly used in audiology clinics do not capture decisional and linguistic factors that significantly contribute to speech processing and hearing abilities. The goal of this dissertation was to examine the extent to which individual differences in sensory decision-making processes and neural tracking of acoustic and linguistic information in speech underlies SPLDs in adults with normal hearing.

1.1 Measuring Self-Perceived Listening Difficulties

One important hurdle to investigating SPLDs is how to adequately quantify the listener's perceived difficulty. Clients often are self-aware of their hearing difficulties, which is typically noted in the client's file. A few self-assessment scales have been developed to quantify the level of general hearing difficulties, including the Hearing Handicap Inventory for Adults (HHIA; Newman et al., 1990) and the Speech, Spatial, and Qualities of Hearing Scale (SSQ) (Gatehouse & Noble, 2004).

The HHIA is a self-assessment scale that describes a client's reaction to their hearing loss in their personal life (Newman et al., 1990). The HHIA is widely used in audiology clinics and has been shown to have significant correlations with pure-tone thresholds and word recognition ability, with high test-retest reliability (Newman et al., 1990, 1991; Saccone & Steiger, 2007). For each of the 25 questions in the HHIA, the client responds with a 'yes', 'sometimes', or 'no', where each value is given four, two, or zero points, respectively. There are thirteen questions that pertain to emotional reactions, such as "Does a hearing problem cause you to feel embarrassed to meet new people?", and twelve questions on a social/situational subscale (e.g., "Does a hearing problem cause you difficulty when visiting friends, relatives, or neighbors?"). While the questions pertaining to social/situational effects of hearing loss do quantify some aspects of SPLDs, the ‘yes’ / ‘sometimes’ / ‘no’ scoring system has some downfalls. Particularly, it forces a client to select ‘yes’ if the problem is present, but another client may also choose ‘yes’ to the same problem, even when the problem may be less severe for the second client. Thus, it may not adequately quantify SPLDs.

The SSQ is a self-report questionnaire that covers three domains: speech, spatial, and general hearing qualities (Gatehouse & Noble, 2004). Under the speech domain, listeners rate their ability to perceive speech under a variety of listening situations with a range of assumed difficulty. The questions encompass visibility of other talkers, number of competing talkers, and background noise. The spatial domain addresses components of hearing such as direction, distance, and movement judgements. Lastly, the hearing domain covers other qualities of hearing not addressed in the speech and spatial domains. These qualities of hearing include segregating sounds, clarity of sounds, and listening effort. There are 14 speech, 17 spatial, and 18 hearing questions. Each question is on a Likert rating scale from 0-10, with the left-hand end of the scale (0) reflecting complete inability or the complete absence of the quality. The right-hand end (10) represents complete ability or the complete presence of that quality. As such, high scores reflect greater ability. Each question also has a ‘Not Applicable’ response option.

The SSQ was developed as a questionnaire to capture aspects of listening abilities and capacities and to assess how those measures related to hearing handicap (Gatehouse & Noble, 2004). The authors of the SSQ developed the questionnaire to understand what is disabling about hearing impairments and how those disabilities determine the experiences of hearing handicaps. The original study found that spatial and hearing aspects of the SSQ were strongly associated with a measure of hearing handicap (Gatehouse & Noble, 2004). Follow-up studies have also demonstrated that the SSQ has very good test-retest reliability and validity (Demeester et al., 2012; Noble et al., 2013; Singh & Pichora-Fuller, 2010). However, the SSQ is not widely used in audiology clinics to measure hearing difficulties, which may be partly due to the length of the questionnaire. Several studies have developed shortened versions of the SSQ with the goal of optimizing its use in clinical settings, narrowing the questionnaire down to five questions (Demeester et al., 2012) and twelve questions (Noble et al., 2013). Even still, the SSQ is primarily used in research settings to evaluate the benefit of hearing aids and cochlear implants (Noble et al., 2008; Noble & Gatehouse, 2004, 2006; Singh & Pichora-Fuller, 2010; Summerfield et al., 2009; van Deun et al., 2010; Zhang et al., 2015) and has yet to be regularly deployed in audiology clinics.

Given the breadth of listening abilities that the SSQ captures, it should serve as an adequate metric of SPLDs. Additionally, the SSQ has an advantage in calculating four distinct scores: one score for each subdomain (i.e., speech, spatial, and hearing) and an overall score. The forthcoming studies in this dissertation leverage the SSQ to quantify SPLDs in adults with normal hearing thresholds.

1.2 The Current State of Evaluations for Self-Perceived Listening Difficulties

In the absence of elevated hearing thresholds, central auditory processing disorder (CAPD) is often attributed to SPLDs. CAPD is defined as a deficit in “the neural processing of auditory information in the central auditory nervous system not due to higher order language or cognition” (American Speech-Language-Hearing Association, n.d.). However, a CAPD diagnosis is rare relative to the prevalence of SPLDs in the presence of normal hearing threshold (Cancel et al., in review). The lack of CAPD diagnoses is partially because there are no standardized guidelines for referring an individual for CAPD testing.

Standard audiological tests assess peripheral hearing, physiology, and speech perception abilities. When a client presents to an audiology clinic with complaints of SPLDs, an audiogram to measure hearing thresholds is typically the first test they will receive in an audiological battery. For adults, this typically consists of presenting pure tones at different frequencies and finding the softest level (i.e., threshold) at which the pure tone can be detected. Hearing thresholds are considered clinically normal when thresholds are at or below 20 dB HL for all octave frequencies (Katz, 2015). If thresholds are greater than 20 dB HL, which indicates hearing loss, then additional testing is performed to evaluate the client's overall hearing profile. However, as noted previously, a significant portion of clients complain of hearing difficulties yet have normal hearing thresholds (Hind et al., 2011; Parthasarathy et al., 2020; Pryce & Wainwright, 2008; Spankovich et al., 2018; Tremblay et al., 2015). In an analysis of 47,009 client records from 2015 to 2020 from the University of Pittsburgh Medical Center Audiology Clinics, 7,203 had normal hearing thresholds in both ears, comprising ~15% of the client population (Figure 1). Of these 7,203 clients, 42% complained of hearing loss or hearing in noise difficulties (Cancel et al., in review). These data closely align with the proportions observed in several other studies (Hind et al., 2011; Parthasarathy et al., 2020; Pryce & Wainwright, 2008; Spankovich et al., 2018; Tremblay et al., 2015).

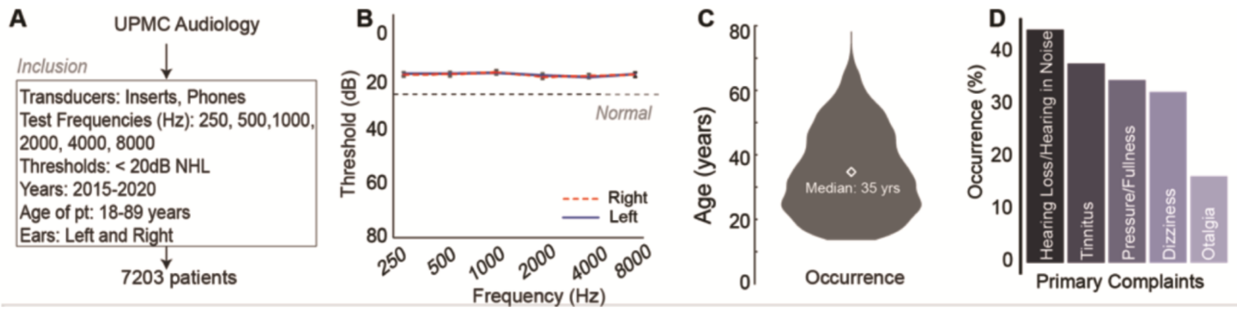

Figure 1: Analysis of University of Pittsburgh Medical Center Audiological Data. A) Client records from 2015-2020 from the Unviersity of Pittsburgh Medical Center Audiology were analyzed to examine the number of clients who had clinically-normal hearing thresholds in both ears. Out of 47,009 clients, 7,203 clients were identified as having normal hearing threshodls in both ears. B) Average audiogram thresholds for the 7,203 clients. C) Age ranges of the clients with normal hearing thresholds. The median age is denoted by the diamond (Mdn = 35 years). D) Primary complaints of the clients with normal hearing thresholds. Approximatly 42% of clients with normal hearing thresholds in both ears presented with complaints of hearing loss or hearing in noise difficulties. Figure adapted from Cancel et al. (in review).

Beyond an audiogram, additional tests can be performed to assess the peripheral and central auditory systems, including health of inner and outer hair cells in the cochlea, middle-ear muscle reflexes, and functioning of auditory structures in the brainstem. Human auditory communication is primarily comprised of continuous speech, which the aforementioned audiological tests do not precisely capture. To combat this issue, several speech perception tests have been developed and validated for use in audiology clinics (Table 1). These tests vary in the type of speech stimuli (i.e., words vs. sentences) and whether background noise is a component. The Speech Recognition Threshold (SRT) test is used to determine the minimum threshold level for hearing speech in quiet that an individual can correctly recognize 50% of the time (American Speech-Language-Hearing Association, 1987). The test often consists of 36 spondaic words, which are bisyllabic words that have equal stress across both syllables, presented in quiet. The SRT is typically obtained via a descending technique, wherein the list is presented 30-40 dB above the estimated SRT. Subsequent spondees after a correct repetition are presented 10 dB lower, and the hearing level is increased by 5 dB for every missed spondee. Testing ceases when five of the last six words are missed. The SRT is then calculated by estimating the 50% point on the psychometric function.

For measuring speech perception in noise, two tests are commonly used: 1) the Words in Noise (WIN) test (Wilson, 2003) and 2) the Quick Speech-in-Noise (QuickSIN) test (Killion et al., 2004). The WIN test consists of monosyllabic words masked in multi-talker babble at seven SNR levels from 0 dB to 24 dB in 4 dB steps. Six words are presented at each SNR level, and the client is instructed to repeat back the words heard. The output of the WIN test is a SNR level at which the client can correctly identify words 50% of the time. QuickSIN also measures speech perception in noise but uses sentences rather than words. The QuickSIN test consists of a list of six sentences masked in multi-talker babble at SNR levels starting at 25 dB. Each successive sentence in the list decreases in SNR level in 5 dB steps, with the final sentence presented at 0 dB SNR. The client is instructed to repeat the target sentence to the best of their ability, even if they can only recall a few words. There are five keywords per sentence for identification. The sum of keywords correctly identified at each SNR level is entered into an equation to calculate the SNR level at which the client can correctly identify sentences 50% of the time.

Table 1: Standard Speech Perception Tests Commonly Used in Audiology Clinics

Test	Stimulus	Signal-to-Noise Ratios	Noise Type	Outcome Measure
Speech Recognition Thresholds	Bisyllabic words	None	None	Minimum hearing level for speech for 50% correct identification
Words in Noise	Monosyllabic words	Seven signal-to-noise ratio levels from 24 to 0 dB in 4 dB steps	Multi-talker babble	Signal-to-noise ratio level at which a listener can understand words in background noises 50% of the time
Quick Speech-in Noise	Sentences	Six signal-to-noise ratio levels from 25 to 0 dB in 5 dB steps	Four-talker babble	Signal-to-noise ratio level at which a listener can understand speech in background 50% of the time

The SRT, WIN, and QuickSIN tests reviewed above each provide the level at which the client can accurately perceive speech in quiet or in noise half of the time. While these tests provide a measure of the client's speech perception abilities, they do not account for additional cognitive aspects that drive speech perception performance. For example, consider a situation where a client who presents to an audiology clinic with a primary complaint of speech perception in noise difficulties performs equally as well as a second client with no complaints of speech perception in noise difficulties on both the WIN and QuickSIN tests. The first client may have had to exert substantially more listening effort to reach the same level of performance as the second client. However, simply marking the keywords as correctly identified or not does not inform the audiologist about how much effort the listener had to put forth. Listeners similarly also tend to take longer to respond in difficult listening situations (Gatehouse & Gordon, 1990), which indicates that response times can provide insight into some the cognitive processes during listening. Therefore, measuring accuracies on clinical speech tests alone misses important aspects of a client's speech perception in noise abilities.

1.3 Sources of Speech Perception Difficulties

A combination of bottom-up sensory and top-down cognitive processes contributes to speech perception abilities (Bonte et al., 2006; Parthasarathy et al., 2020; Pichora-Fuller, 2008; Zekveld et al., 2006). Individual differences in speech perception can arise at any point along the auditory pathway to the cortex and require a multifaceted approach to understand. The incoming speech stream must first be represented in the ascending auditory pathway with high fidelity. Hearing loss caused by aging or noise exposure can lead to damage to the hair cells and to the synapses between inner hair cells and the auditory nerve fibers (Furman et al., 2013; Kujawa & Liberman, 2009; Valero et al., 2017), both of which impact the fidelity with which the auditory signal is encoded.

Once the encoded auditory signal travels along the ascending auditory pathway and reaches the auditory the cortex, the signal is mapped onto established cortical representations in the superior temporal gyrus (Benson et al., 2001; Chandrasekaran et al., 2010; Scott et al., 2000; Scott & Johnsrude, 2003). Once the auditory signal has reached the cortex, the time and frequency information are mapped onto established neural representations for that sound. When a speech signal is acoustically degraded, whether it be from background noise, hearing difficulties, or language barriers, it can be difficult to map the sound onto the correct neural representation for perception (Peelle, 2018). Moreover, additional brain regions beyond the auditory cortex have been shown to be recruited to assist with speech perception (Du et al., 2016; Peelle et al., 2010;

Vaden et al., 2015; Wong et al., 2009). For instance, activation in the superior temporal gyrus has been shown to be coupled with activity in the dorsal prefrontal cortex in the presence of noise and was associated with speech perception in noise performance (Wong et al., 2009). The recruitment of these extra-auditory regions is hypothesized to serve as a compensatory mechanism to aid speech perception in challenging listening situations and have primarily been demonstrated in older adults (Du et al., 2016; Wong et al., 2009). As such, the extent to which listeners with SPLDs employ similar compensatory mechanisms to aid speech perception remains unclear.

Top-down cognitive processes, like working memory and selective attention, also play critical roles in speech processing under challenging listening conditions (Gordon-Salant & Fitzgibbons, 1997; Pichora-Fuller et al., 1995) and are hypothesized to be associated with the compensatory recruitment that has been observed in prior speech perception in noise research (Du et al., 2016). During listening, selective attention is used to actively focus on the target speech, and working memory is used to hold onto the incoming speech while repairing the degraded signal (Shinn-Cunningham, 2008). Individual differences in selective attention (Holt et al., 2018; Oberfeld & Klöckner-Nowotny, 2016; Shinn-Cunningham & Best, 2008; Tierney et al., 2019) and working memory (Gatehouse & Gordon, 1990; Ng & Rönnberg, 2020; Rönnberg et al., 2010, 2013; Souza & Arehart, 2015; Xie et al., 2015) can contribute to speech perception in noise performance. For instance, listeners with poorer selective attention abilities tend to perform worse in multi-talker listening scenarios because of difficulties attending to the target speaker and ignoring irrelevant information from other speakers (Shinn-Cunningham, 2017). Additionally, listeners with lower working memory also perform worse in noisy listening situations because of reduced capacity to hold onto speech for repairing and recognition (Füllgrabe & Rosen, 2016; Rönnberg et al., 2010).

When a speech signal is degraded, the listener can use their stored linguistic knowledge to make inferences about the missing information and to make predictions for upcoming speech (N. Ding et al., 2016; Hickok & Poeppel, 2007; Hunter & Humes, 2022; Pichora-Fuller, 2008). Predicting upcoming speech based on stored linguistic knowledge is a critical process that has significant implications for speech comprehension. For example, prior supportive sentence-level context tends to benefit speech comprehension in listeners with hearing impairments more than listeners with normal hearing (Hunter & Humes, 2022; Pichora-Fuller, 2008). This suggests that when the speech signal is acoustically degraded, listeners rely on knowledge at larger linguistic units (i.e., words and phrases) for speech comprehension. Furthermore, the use of stored linguistic knowledge assumes the listener can correctly and rapidly access higher-order information, like words and meaning. Differences in the use of supportive linguistic context has been observed in listeners with hearing impairments as well as older adults with normal hearing (Hunter & Humes, 2022; Milburn et al., 2021; Pichora-Fuller, 2008). However, it remains unclear the extent to which reliance on linguistic knowledge to aid speech comprehension may differ in listeners with SPLDs who have normal hearing thresholds.

There are also several other sources of individual differences in speech perception performance that must be noted. First, depressive symptoms have been shown to negatively affect speech perception performance in single-talker masking conditions (Chandrasekaran et al., 2015; Xie, Zinszer, et al., 2019; Yi et al., 2019). These studies hypothesize that depression alters the cognitive symptoms that support speech perception in noise performance, making it difficult to attend to the target speaker. Individual differences in musical experience have also been shown to impact speech perception in noise performance. Specifically, individuals with more musical experience show an advantage in multi-talker listening scenarios as opposed to those with less experience (Parbery-Clark et al., 2009, 2011; Slater et al., 2015; Tierney et al., 2019). However, the presence of a musician-advantage in speech perception in noise remains under debate because several studies have been unable to replicate it (Boebinger et al., 2015; Madsen et al., 2017, 2019; Ruggles et al., 2014; Yeend et al., 2017). Furthermore, overall language experience (Xie et al., 2014) and genetic factors that impact cognitive processes (Xie et al., 2015) have also been linked with individual differences in speech perception. In summary, there are many sources of individual differences in top-down cognitive processes that can impact speech perception in noise abilities in adults with normal hearing thresholds.

Taken together, individual differences in SPLDs may arise from a combination of bottom-up sensory or top-down cognitive processes. Prior research on SPLDs in adults with normal hearing thresholds has primarily focused on bottom-up sensory related components (Spankovich et al., 2018; Tremblay et al., 2015). To this end, there is a critical gap in our understanding of SPLDs because much remains unknown about the extent to which top-down processes, such as decisional or linguistic processing, may impact SPLDs.

1.4 Integrating Accuracies and Response Times to Understand Speech Perception

The process of mapping phonemes from an incoming speech stream onto neural representations requires categorization decisions. Individuals with more SPLDs may differ in the decisional processes underlying sound categorization as opposed to those with fewer SPLDs. However, the examination of accuracies alone from speech perception tasks fails to capture the decisional processes involved in categorization. One straightforward option to better understand the decision-making component of speech perception abilities is to integrate the use of response times. Response times show a trade-off with accuracies; response times lengthen when accuracy decreases, but faster response times can also come at the cost of poorer accuracy (Binder et al., 2004). Therefore, response times are considered an appropriate behavioral correlate of the decision-making component of auditory processing.

The combination of accuracies and response times may be key to understanding perceptual decision-making processes during listening. Perceptual decision-making models, like drift-diffusion models (DDM), may provide a promising approach to understanding these perceptual processes during listening in noisy situations. DDMs are neurobiologically driven models that use accuracies and response times to understand the underlying processes of perceptual decision-making (Ratcliff, 1978; Ratcliff et al., 2016). Theoretically, DDMs assume that, when a stimulus is presented, evidence is noisily accumulated from the stimulus towards each decision alternative, and a decision is made when a particular evidence threshold has been reached (Figure 2)(Ratcliff, 1978; Ratcliff et al., 2016). *Evidence accumulation rate* and *decision threshold* are two primary parameters output from a DDM. Evidence accumulation rate reflects how efficient the listener is at extracting information from the stimulus to make a decision, while decision threshold reflects how much information the listener needs to accumulate before a decision is reached (Ratcliff, 1978; Ratcliff et al., 2016). The decision threshold is also referred to as response caution, which is the tradeoff between speed and accuracy of decision-making (Bogacz et al., 2010). Both parameters have been shown to be associated with distinct neural activity.

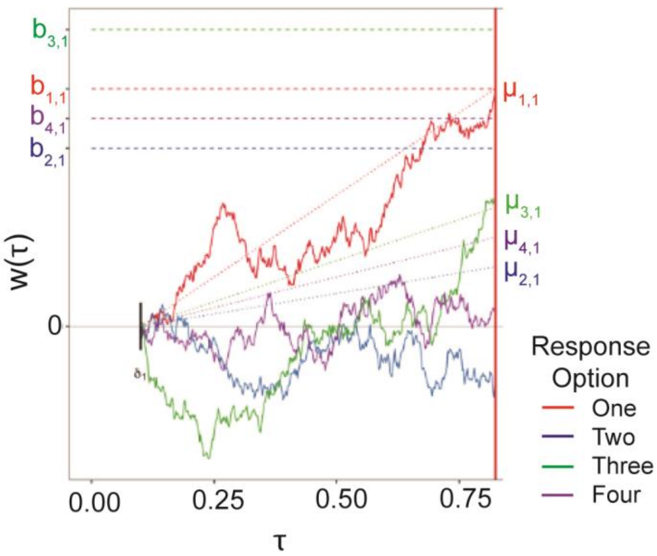

Figure 2: Drift-diffusion model for perceptual decision-making with four response options. When a stimulus is presented and initially encoded (vertical black line), evidence for each response option begins to accumulate (μ). A decision (d) is made when an evidence accumulator ($\mu_{d,s}$) reaches its corresponding decision threshold ($b_{d,s}$). Here, the stimulus (s) for response option 1 was presented and correctly identified. W(τ) reflects the decision process on the y-axis and relative time (τ) from stimulus presentation is on the x-axis. This figure is adapted from Paulon et al. (2020).

DDMs are typically applied to accuracies and response times from categorization tasks involving at least two decision-alternatives and have been applied both in the auditory (Roark et al., 2021; Vaden et al., 2022) and visual domains (Kayser et al., 2010; Ratcliff et al., 2001, 2006b, 2006a). In these categorization tasks, a stimulus is presented, and the participant is instructed to categorize the stimulus as quickly and as accurately as possibly by a button press on a keyboard or a mouse click. While DDMs have been applied in the auditory domain, little research has examined perceptual decision-making during speech perception in noise. Vaden et al. (2022) applied a similar diffusion model to accuracies and response time data from a words-in-noise task to understand the decisional processes underlying word recognition. They found that signal-to-noise ratio (SNR) levels modulated both evidence accumulation rates and decision thresholds. Vaden and colleagues (2022) also demonstrated an association between decision thresholds and activity in the cingulo-opercular network prior to stimulus presentation. Moreover, prior work suggests that pre-stimulus cingulo-opercular network activity is predictive of word recognition in noise performance (Vaden et al., 2013, 2016). Not only do these results demonstrate that background noise impacts decisional processes during speech perception in noise, but it also establishes a neurobiological connection between decisional processes and speech perception in noise performance. Therefore, this begs the question as to whether evidence accumulation rates and decision thresholds during speech perception in noise may differ as a function of SPLDs.

1.5 Evaluating Speech Perception in Naturalistic Listening Situations

Traditionally, speech perception has been measured using very controlled experimental designs that use shortened stimuli comprised of isolated syllables and words or sentences. However, isolated syllables, words, and sentences are not representative of everyday listening situations (Hamilton & Huth, 2018). More recently, research investigating speech perception has converted to using more naturalistic stimuli in the form of continuous speech. Continuous speech in the form of narrated passages and stories allow us to examine speech processing in more naturalistic listening situations (Aiken & Picton, 2008; Broderick et al., 2018; Hamilton & Huth, 2018; Huth et al., 2016; Moerel et al., 2012; Xie, Reetzke, et al., 2019). Additionally, continuous speech stimuli can be used to investigate the neural mechanisms that drive speech perception use neurophysiological methods like electroencephalography (EEG) (Crosse et al., 2016; Di Liberto et al., 2015; Etard & Reichenbach, 2019; Horton et al., 2013; Lesenfants, Vanthornhout, Verschueren, & Francart, 2019; Lesenfants, Vanthornhout, Verschueren, Decruy, et al., 2019; McHaney et al., 2021; Petersen et al., 2016; Reetzke et al., 2021; Xie et al., 2023) or magnetoencephalography (Brodbeck et al., 2020; Brodbeck, Presacco, et al., 2018; Decruy et al., 2019; Ding et al., 2014; Ding & Simon, 2012, 2013; Presacco et al., 2016).

Neural tracking of the acoustic speech envelope has been shown to be associated with speech comprehension in a variety of listening conditions (Brodbeck et al., 2020; Ding et al., 2014; Ding & Simon, 2012; Lesenfants, Vanthornhout, Verschueren, Decruy, et al., 2019; McHaney et al., 2021; Reetzke et al., 2021; Vanthornhout et al., 2018). However, studies have begun to demonstrate that neural tracking is driven by the acoustic *and* linguistic information in continuous speech (Brodbeck et al., 2022; Brodbeck, Hong, et al., 2018; Hamilton et al., 2018; Xie et al., 2023). For instance, in multi-talker listening situations, acoustic information from the attended and unattended speakers were represented in the neural response, but only lexical information from the attended speaker was present in the response (Brodbeck, Hong, et al., 2018). Studies have also shown that linguistic information at different contextual levels—phoneme, word, and sentence—significantly contributed to the neural responses evoked by continuous speech (Brodbeck et al., 2022; Xie et al., 2023). Furthermore, neural representations of linguistic information during speech processing may impact speech comprehension (Xie et al., 2023). However, the extent to which individuals with SPLDs differentially leverage acoustic and linguistic information in continuous speech remains unknown.

1.6 Specific Aims

This dissertation is comprised of two specific aims across two studies. Aim 1 (Chapter 2) examined the role of perceptual decision-making in SPLDs, while Aim 2 (Chapter 3) examined the extent to which the use of acoustic and linguistic information in speech differed based on SPLDs. The realization of these aims will provide insight into potential differences in sensory decision-making and use of acoustic and linguistic information during listening, which may underlie SPLDs.

1.6.1 Aim 1

The primary goal of Aim 1 (Study 1 in Chapter 2) was to examine the extent to which evidence accumulation rate and decision thresholds from DDMs explain individual differences in SPLDs. Adults between the ages of 18-53 with thresholds ≤25 dB at octave frequencies .25-8 kHz participated in Study 1. These participants completed a phoneme in noise categorization task, which used synthetically generated (Klatt, 1980) /ba/, /da/, and /ga/ phonemes, which primarily differed in their second formant transition (Johnson et al., 2008). The phonemes were presented in the context of a syllable in quiet and in speech-shaped noise (SSN) at various signal-to-noise ratio (SNR) levels. I hypothesized that individuals with more SPLDs would have lower evidence accumulation rates and decision thresholds than those with fewer SPLDs. Study 1 provided insight into how neurobiologically driven processes of perceptual decision-making contribute to individual differences in SPLDs.

1.6.2 Aim 2

Aim 2 (Study 2 in Chapter 3) takes a neural tracking approach to understand whether listeners with more SPLDs utilize different types of acoustic and linguistic information during listening than those with fewer SPLDs. Here, EEG was recorded while participants listened to continuous speech from an audiobook under two listening conditions: 1) in quiet and 2) masked in speech-shaped noise (SSN) at -2 dB SNR. Multivariate temporal response function encoding models were then fit to predict EEG responses to acoustic and linguistic speech features at various representation levels, including the acoustic envelope and information-theoretic models of sublexical-, word-form-, and sentence-level representations in the audiobook speech stimuli in both listening conditions. These acoustic and linguistic models provide insight into how listeners leverage acoustic and linguistic information in speech to assist with speech comprehension and the extent to which this information is used differently as a function of SPLDs.

2.0 Study 1: Decision-Making Processes Underlie an Aspect of Self-Perceived Listening Difficulties in Adults with Normal Hearing

2.1 Introduction

Processing speech in noisy listening situations is a complex task that we encounter on a near daily basis. For instance, when having a conversation with a friend in a crowded room, you need to be able to focus on your friend's speech while ignoring the irrelevant background noise. This task can be difficult, but we often perform it remarkably well. Successful speech perception in noise requires both bottom-up sensory encoding and top-down cognitive processes (Lam et al., 2017; Parthasarathy et al., 2020; Pichora-Fuller et al., 2016; Rönnberg et al., 2013). Hearing loss due to age and noise exposure can contribute to speech perception difficulties (Dinino et al., 2021; Helfer & Wilber, 1990; Peelle et al., 2011). However, individual variability in speech perception in noise still exists even when listeners have clinically normal hearing thresholds (Lam et al., 2017; Meister et al., 2018; Parbery-Clark et al., 2011; Ross et al., 2007; Xie et al., 2014, 2015). Approximately 10% of clients at audiology clinics who complain of listening difficulties have clinically normal hearing thresholds (Hind et al., 2011; Parthasarathy et al., 2020; Pryce & Wainwright, 2008; Spankovich et al., 2018; Tremblay et al., 2015). The reasons for these self-perceived listening difficulties (SPLDs) in the absence of overt hearing loss remain unclear.

Understanding the factors driving SPLDs in adults with normal hearing thresholds is clinically relevant because listening difficulties have been shown to be associated with increased rates of social isolation (Lin & Albert, 2014), decreased productivity (Huddle et al., 2017), and impacts to overall well-being (Bainbridge & Wallhagen, 2014; Pryce & Wainwright, 2008). Prior research has examined a myriad of deficits in bottom-up processing as contributing to SPLDs. These bottom-up deficits include auditory neuropathy (Starr et al., 1996), central auditory processing disorder (Jerger & Musiek, 2000), central presbycusis (Welsh et al., 1985), and idiopathic discriminatory dysfunction (Rappaport et al., 1993). Recently, anatomical studies in animal models have demonstrated a loss of synapses between the inner hair cells in the cochlea and the auditory nerve (Kujawa & Liberman, 2009; Parthasarathy & Kujawa, 2018; Sergeyenko et al., 2013). This pathology has been referred to as cochlear synaptopathy (Kujawa & Liberman, 2015) and has been observed in both noise-traumatized (Furman et al., 2013) and aging rodent models (Parthasarathy & Kujawa, 2018). The potential loss of cochlear synapses in humans with normal hearing thresholds is hypothesized to contribute to speech perception difficulties in noisy listening conditions (Bharadwaj, Verhulst, et al., 2014; Grant et al., 2022).

While age-related and noise-induced cochlear synaptopathy have been well documented in animal models (Kujawa & Liberman, 2009; Parthasarathy & Kujawa, 2018; Sergeyenko et al., 2013; Shaheen et al., 2015) and have also been demonstrated in post-mortem temporal bone studies in humans (Makary et al., 2011; Wu et al., 2019), there have been considerable inconsistencies in identifying a marker for the presence of cochlear synaptopathy in humans in vivo (Bramhall et al., 2019). For instance, some studies have found reduced auditory brainstem response (ABR) wave I amplitudes in humans with high noise exposure who have normal hearing thresholds (Bramhall et al., 2017; Liberman et al., 2016b; Stamper & Johnson, 2015a, 2015b; Valderrama et al., 2018), which aligns with cochlear synaptopathy findings in rodents (Kujawa & Liberman, 2009; Sergeyenko et al., 2013). However, other studies observed no effect of noise exposure on ABR wave I amplitudes in humans (Fulbright et al., 2017; Grinn et al., 2017; Guest et al., 2017; Prendergast et al., 2017, 2018). Moreover, some research has demonstrated a small relationship between reduced envelope following responses and noise-exposure in adults with normal hearing thresholds (Bharadwaj et al., 2015; Bharadwaj, Verhulst, et al., 2014; Paul et al., 2017; Verhulst et al., 2018), but other studies have failed to replicate this relationship (Guest et al., 2017; Prendergast et al., 2017). Therefore, a clear link between deficits in bottom-up processing and SPLDs in adults with normal hearing thresholds has yet to be established.

Although there has been extensive research on bottom-up processes, much less is understood about the role of top-down processes in SPLDs. Prior research has identified at least three neural networks beyond the auditory cortex that are active during speech perception (Du et al., 2016; Peelle & Wingfield, 2016; Shinn-Cunningham, 2017; Vaden et al., 2013). For instance, when a listener engages top-down selective attention to focus on spatial auditory cues, frontal and parietal regions show increased activation as opposed to when the listener selectively attends to spectro-temporal acoustic cues (Bharadwaj, Lee, et al., 2014; Hill & Miller, 2010; Lee et al., 2012). Older adults have also been shown to upregulate activity in frontal articulatory regions, which is hypothesized to serve as a compensatory mechanism to aid speech perception in noisy listening conditions (Du et al., 2016). Moreover, the cingulo-opercular network (CON) has been proposed as a critical system for processing degraded speech (Peelle & Wingfield, 2016). Activity in the CON prior to a word being presented is associated with word recognition accuracy on that trial (Vaden et al., 2013). This association between CON activity and word recognition suggests that the CON may serve as a performance-monitoring and adaptive control mechanism for word recognition in noise. Collectively, activity in any one of these neural networks could partially account for individual differences in speech perception.

There have been few attempts to link potential neural network dysfunctions to individual differences in speech perception in noise. One exception is a study by Vaden and colleagues

(2022), which used a Shifted-Wald diffusion model to examine the association between decisional processes during a word recognition in noise task and CON activity. This study found that elevated CON activity was associated with higher decision criteria. Higher decision criteria are associated with more cautious responding, which allows an individual to collect more evidence before making a decision. In the context of word recognition, the CON may serve as an adaptive control to adjust decision criteria to optimize word recognition performance (Vaden et al., 2022). This begs the question as to whether individual differences in top-down processes, such as CON adaptive control, can be attributed to SPLDs in adults with normal hearing thresholds.

In the current study, I applied drift-diffusion models (DDM) to accuracies and response times from a phoneme in noise categorization task to understand individual differences in decisional processes as a function of SPLDs. DDMs assume that the brain accumulates sensory evidence to make a categorical decision (Ratcliff et al., 2016). This sensory evidence accumulation process is reflected by an increase in local neuronal firing rates associated with each category alternative. When neuronal firing rates associated with a category option crosses a particular threshold, a category decision is made (Brody & Hanks, 2016; Gold & Shadlen, 2007). The two parameters from the DDM that are of interest to the current study are evidence accumulation rate and decision threshold. Each category has its own decision threshold, where larger thresholds reflect more cautious responding because more evidence must accumulate before a category decision can be made (Bogacz et al., 2010). Here, I examined the extent to which evidence accumulation rate and decision thresholds may differ based on SPLDs. Prior research has demonstrated that better word recognition performance is associated with higher decision thresholds (Vaden et al., 2022). Therefore, I hypothesized that adults with fewer SPLDs would have higher decision thresholds and be more cautious responders than those with more SPLDs.

Additionally, I hypothesized that listeners with fewer SPLDs would be more efficient evidence accumulators, which would indicate that they are better able to extract critical information from a stimulus to make a category decision.

2.2 Methods

2.2.1 Participants

A total of 77 native English-speaking participants between the ages of 18-53 years (Figure 3A; *M* = 27.58, *SD* = 11.22) were recruited for this study from the greater Pittsburgh, PA community. All participants had normal otoscopy and hearing thresholds ≤25 dB for frequencies 250-4000 Hz in octave steps and ≤30 dB at 8000 Hz. A subset of participants (n = 41) was a part of a larger study examining speech perception in noise in aging and underwent several additional testing procedures not covered in the scope of the present study. Participants received either monetary compensation or research credit for their participation. This research protocol was approved by the Institutional Review Board at the University of Pittsburgh.

2.2.2 Behavioral Measures

2.2.2.1 Speech, Spatial, and Qualities of Hearing Scale

All participants completed the Speech, Spatial, and Qualities of Hearing Scale (SSQ) (Gatehouse & Noble, 2004). The SSQ is a self-report measure of hearing abilities across several domains, including speech hearing, spatial hearing, and other qualities of hearing. Fourteen questions in the speech hearing section assess realistic speech contexts such as competing sounds, differences in background conditions, and visibility of other talkers (e.g., Can you have a conversation with someone when another person is speaking whose voice is the same pitch as the person you're talking to?). The spatial hearing questions (n = 17) assess directional and distance aspects of hearing (e.g., In the street, can you tell how far away someone is, from the sound of their voice or footsteps?). The final section of the SSQ has eighteen questions that address other hearing qualities, such as clarity, recognition, and listening effort (e.g., Do you find it easy to recognize different people you know by the sound of each one's voice?). All questions are on a Likert-scale from 0-10.

Domain scores for each section of the SSQ were calculated for each participant by averaging the responses in each section. The domain scores were then averaged into a composite SSQ score (Figure 3B), which was used as a metric of SPLD. Higher SSQ scores reflect fewer SPLD, while lower SSQ scores reflect more SPLD.

2.2.2.2 Audiometry

The hearing sensitivity for all participants was evaluated using pure tone audiometry. Air conduction hearing thresholds were measured via insert earphones for octave frequencies 250 through 8000 Hz. A modified Hughson and Westlake method (Carhart & Jerger, 1959) was used to estimate the threshold at each frequency. Pure tone thresholds for all participants are shown in Figure 3C.

2.2.3 Phoneme in Noise Categorization

2.2.3.1 Stimuli

Participants categorized synthetically generated /ba/, /da/, and /ga/ stimuli (Johnson et al., 2008). The three stimuli were synthesized at a sampling rate of 20 kHz using a Klatt cascade/parallel formant synthesizer (Klatt, 1980). Each stimulus was 170 ms long, with a voicing onset (100 Hz F_0) at 10 ms. Each stimulus had an identical and constant fundamental frequency (F_0) throughout. The stimuli had the same formant transition durations (50 ms), a linearly rising first formant (400-720 Hz), and flat fourth (3300 Hz), fifth (3750 Hz), and sixth (4900 Hz) formants. The second (F_2) and third (F_3) formant starting positions differ between stimuli but remain constant after their transition endpoints (F_2: 1240 Hz; F_3: 2500 Hz). F_2 rises from 900 Hz for /ba/, falls from 1700 Hz for /da/, and falls from 3000 Hz for /ga/. F_3 rises from 2400 Hz for /ba/, falls from 2580 Hz for /da/, and falls from 3100 Hz for /ga/ (Figure 3D).

The stimuli were presented in quiet as well as in SSN. The SSN masker was derived from the three phoneme stimuli and all stimuli from the Quick Speech in Noise (QuickSIN) test lists 1-8 (Killion et al., 2004). The QuickSIN test is a standardized test for measuring speech perception in noise abilities in individuals with hearing loss. The stimuli from QuickSIN were chosen to add more speech samples for the SSN masker. The SSN masker was created using the following steps. First, the fast Fourier Transform (FFT) was derived from all stimuli files combined. The FFT was then manipulated such that the phases of the spectral components were randomized. The modified FFT was then converted back into the time domain using an inverse Fourier Transform. Thus, the spectrum from the resulting SSN closely matched the spectrum of the original stimuli files.

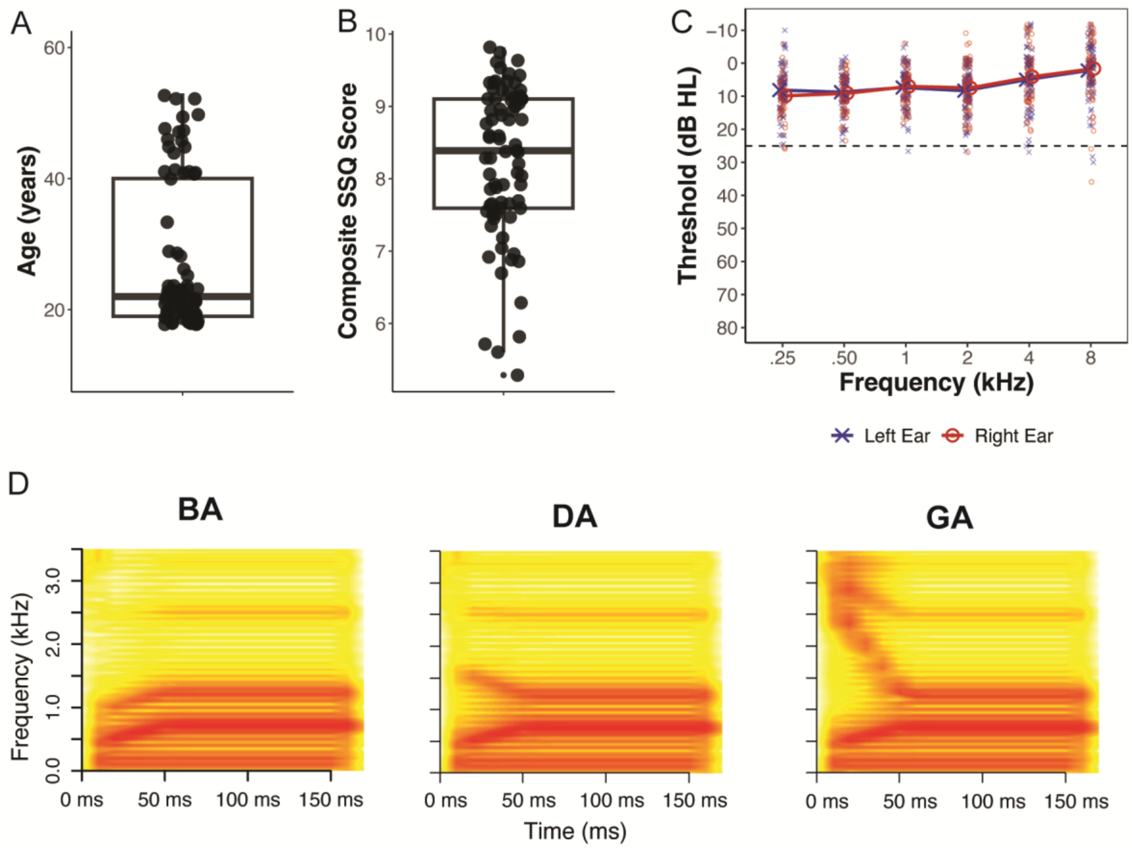

Figure 3: Study 1 Behavioral Measures and Stimuli. A) Distribution of ages of study participants in years. B) Distribution of self-perceived listening difficulties (SPLDs) as measured by the composite Speech, Spatial, and Qualities of Hearing Scale (SSQ). Here, larger composite SSQ scores reflect fewer SPLDs, while smaller SSQ scores reflect more SPLDs. C) Audiogram thresholds for all study participants. The solid lines denote average thresholds. Individual participant thresholds for right ear are denoted by circles and left ear by an X. All participants were required to have thresholds ≤25 dB in both ears at frequencies 250-4000 Hz to participate in the study. D) Spectrogram of the stimuli from the phoneme in noise categorization task.

2.2.3.2 Categorization Task Procedure

Participants categorized consonant-vowel stop syllable phonemes (/ba/, /da/, /ga/) in quiet or masked in speech-shaped noise (SSN) at SNR levels: +8, -2, -6, and -9 dB. The task began with four practice trials per phoneme in quiet (4 trials x 3 phonemes = 12 trials). Participants were instructed to categorize the sound as a 'BA', 'DA', or 'GA' by pressing the appropriate button on the keyboard. Feedback in the form of 'Correct!' or 'Incorrect.' was provided on each practice trial. After completing the practice trials, participants moved onto the main categorization task. This task consisted of five blocks of 60 trials in noise (3 phonemes X 4 SNRs X 5 trials) and 15 trials in quiet, for a total of 75 trials per block. Participants were provided 1500 ms to make their keyboard response. Unlike the practice trials, participants did not receive feedback on their categorization decision during the main task. Stimuli were presented through Sennheiser Headphones at a comfortable listening level. Participants who were a part of the larger speech perception in noise study in aging (n = 41) completed the task in MATLAB 2020b (Mathworks Inc., Natick, Massachusetts). The remaining 36 participants who were not a part of the larger study completed the task using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020).

2.2.4 Drift-Diffusion Modeling

The DDMs developed by Paulon and colleagues (2020) were applied to the accuracies and response time data from the phoneme in noise categorization task using the *lddmm* package, version 0.1.0 (Paulon & Sarkar, 2021) in R, version 4.1.3 (R Core Team, 2022). For every combination of phoneme *s* and decision response *d*, the DDM fits an evidence accumulation rate parameter $\mu_{d,s}$ and a decision threshold parameter $b_{d,s}$. which is the tradeoff between speed of decision making and accuracy. An offset parameter δ_s is also fit for each phoneme, which represents the time taken by all actions that are not directly relevant to the decision-making component. These include actions such as stimulus encoding and the physical motor response for the button press. The DDM allows for $\mu_{d,s}$ and $b_{d,s}$ to differ between individuals, which captures the variability across participants. The bottom one percent of trials based on response times were removed (Roark et al., 2021), however, the top one percent were kept because participants were constrained to a 1500 ms response period.

The DDM analyses use a Bayesian framework that relies on posterior samples from a Markov chain Monte Carlo (MCMC) algorithm for estimation and inference. The algorithm was run for 6000 iterations with the first 2000 iterations discarded as burn-in. To reduce autocorrelation, the remaining 4000 iterations were thinned in intervals of five. The DDM was performed on all trials from the task, but only trials that were categorized correctly were used for further statistical analyses outlined in the Statistical Analyses section below (Roark et al., 2021). Posterior means for correct responses are reported as point estimates with 95% pointwise credible intervals to assess uncertainty.

2.2.5 Statistical Analyses

All data were analyzed using R, version 4.1.3 (R Core Team, 2022) and visualized using the *ggplot2* package, version 3.3.6 (Wickham, 2016). First, pairwise t-tests were performed using the package *rstatix*, version 0.7.0 (Kassambara, 2021) to assess whether accuracies and response times in the quiet condition were significantly different than performance at +8 dB. If performance in the quiet condition was similar to SNR 8, then the remaining analyses would focus solely on the noise condition (SNRs +8, -2, -6, and -9 dB) in order to examine the effects of decision-making in noisy listening environments of varying complexity.

Separate linear mixed effects models were estimated to examine accuracies and response times from the phoneme in noise categorization task using the package *lme4*, version 1.1-30 (Bates et al., 2015). The *lmerTest* package, version 3.1-3 (Kuznetsova et al., 2017) was used to estimate *p*-values with Swatterthwaite's method. At each SNR level, the average accuracy and response time was calculated for each subject, which served as the outcome measure of the linear mixed effects models. Each model had fixed effects of SNR and SSQ score. The random effect structure began with a random slope of SSQ score per subject and a random slope of SSQ per age. The final random effect structure was determined based on the maximal model that converged and provided a non-singular fit. Similar models were fit with outcome measures of evidence accumulation rates and decision thresholds. Multiple pairwise comparisons at each SNR level were performed using the *emmeans* and *emtrends* functions in the *emmeans* package, version 1.7.3 (Lenth, 2022). Benjamini-Hochberg adjusted *p*-values were reported to control for the false discovery rate (Benjamini & Hochberg, 1995).

2.3 Results

2.3.1 Phoneme in Noise Categorization Performance

Mean accuracies and response times between quiet and SNR 8 dB on the phoneme in noise categorization task were compared to determine whether the quiet condition should be included in the remainder of the analyses. A pairwise t-test comparing average accuracies in quiet ($M = .889$, $SD = .129$) and SNR 8 ($M = .882$, $SD = .127$) revealed no significant differences, $t(76) = 0.709$, $p = .481$, $d = 0.081$, 95% CI [-0.024, 0.012]. For response times, a pairwise t-test revealed no significant differences between quiet (*M* = 693.070 ms, *SD* = 176.521ms) and SNR 8 (*M* = 708.924 ms, *SD* = 184.453 ms), $t(76) = 1.977$, $p = .052$, $d = 0.167$, 95% CI [-0.116, 31.825]. As such, all quiet trials were removed from the remainder of the analyses in order to focus on decision-making processes in noisy listening environments.

Linear mixed effects models were fit to analyze changes in accuracies and response times across SNR levels as a function of SPLDs, as measured by the SSQ (Figure 4). Multiple comparisons were performed to assess differences between each SNR level. Benjamini-Hochberg adjusted *p*-values are reported. For mean accuracies, the maximal random effect structure that resulted in non-singular fit included a random slope of SSQ per subject and a random intercept of age. Results from the linear mixed effects model on accuracies (Table 2) revealed that there were significant effects of SNR level, such that accuracies increased for higher SNR levels (ps < .001). Additionally, there was a significant interaction of SSQ and the difference in accuracies between SNR levels -9 and -2 dB (p = .031). This interaction indicates that the increase in accuracy from SNR -9 to SNR -2 was significantly larger for listeners with fewer SPLDs. However, there were no significant effects of SSQ score (ps > .05), which suggests that accuracies at each individual SNR level did not significantly differ based on SPLDs. Collectively, these results indicated that listeners with fewer SPLDs may have received greater benefit from decreasing background noise but overall accuracy at each SNR level was not impacted by SPLDs.

Regarding the linear mixed effects model on response times, the maximal random effect structure consisted of a single random intercept of subject. All other random effect structures resulted in a singular fit, or the model did not converge. This linear mixed effects model (Table 3) revealed that there were significant effects of SNR level (*p*s < .01), indicating that higher SNR levels had shorter response times than lower SNR levels. However, there were no significant effects of SSQ score (*p*s $> .05$), which suggests that response times at each SNR level did not differ as a function of SPLDs. Additionally, no significant interactions of SNR level and SSQ score were observed (*p*s $> .05$). These null interactions indicate that the change in response times between SNR levels were not significantly different as a function of SPLDs. Taken together, accuracies and response times during phoneme in noise categorization differed as a function of SNR level but were largely not impacted by SPLDs.

Table 2: Linear Mixed Effects Model Results with Multiple Comparisons for Mean Accuracies. Note: SNR = Signal-to-noise Ratio; SSQ = Speech, Spatial, and Qualities of Hearing Scale; Benjamini-Hochberg adjusted *p*-values are reported; ******p* < .001; **p* ≤ .05.

Fixed Effect	Estimate	*SE*	*t* value	*p* value
SNR 8 vs. SNR -2	.081	.015	5.578	< .001***
SNR 8 vs. SNR -6	.329	.015	22.591	< .001***
SNR 8 vs. SNR -9	.471	.015	32.379	< .001***
SNR -2 vs. SNR -6	.248	.015	17.012	< .001***
SNR -2 vs. SNR -9	.390	.015	26.801	< .001***
SNR -6 vs. SNR -9	.143	.015	9.788	< .001***
SSQ score at SNR 8	-.113	.072	-1.563	.219
SSQ score at SNR -2	-.172	.072	-2.374	.080
SSQ score at SNR -6	-.102	.072	-1.405	.219
SSQ score at SNR -9	.007	.072	0.097	.923
SNR 8 vs. SNR -2 X SSQ Score	-.013	.014	-0.928	.425
SNR 8 vs. SNR -6 X SSQ Score	.003	.014	0.181	.857
SNR 8 vs. SNR -9 X SSQ Score	.027	.014	1.899	.174
SNR -2 vs. SNR -6 X SSQ Score	.015	.014	1.109	.403
SNR -2 vs. SNR -9 X SSQ Score	.039	.014	2.827	.031*
SNR -6 vs. SNR -9 X SSQ Score	.024	.014	1.718	.174

Table 3: Linear Mixed Effects Model Results with Multiple Comparisons for Mean Response Times. Note: SNR = Signal-to-noise Ratio; SSQ = Speech, Spatial, and Qualities of Hearing Scale; Benjamini-Hochberg adjusted *p*-values are reported; ***$p < .001$; **$p < .01$.

Fixed Effect	Estimate	*SE*	*t* value	*p* value
SNR 8 vs. SNR -2	-33.314	11.214	-2.971	.003**
SNR 8 vs. SNR -6	-86.130	11.214	-7.681	< .001***
SNR 8 vs. SNR -9	-119.399	11.214	-10.647	< .001***
SNR -2 vs. SNR -6	-52.816	11.214	-4.710	< .001***
SNR -2 vs. SNR -9	-86.084	11.214	-7.677	< .001***
SNR -6 vs. SNR -9	-33.268	11.214	-2.967	.003**
SSQ score at SNR 8	-175.620	104.553	-1.680	.097
SSQ score at SNR -2	-192.541	104.553	-1.842	.092
SSQ score at SNR -6	-232.349	104.553	-2.222	.058
SSQ score at SNR -9	-244.408	104.553	-2.338	.058
SNR 8 vs. SNR -2 X SSQ Score	-3.735	10.751	-0.347	.805
SNR 8 vs. SNR -6 X SSQ Score	-12.523	10.751	-1.165	.576
SNR 8 vs. SNR -9 X SSQ Score	-15.185	10.751	-1.412	.576
SNR -2 vs. SNR -6 X SSQ Score	-8.788	10.751	-0.817	.622
SNR -2 vs. SNR -9 X SSQ Score	-11.450	10.751	-1.065	.576
SNR -6 vs. SNR -9 X SSQ Score	-2.662	10.751	-0.248	.805

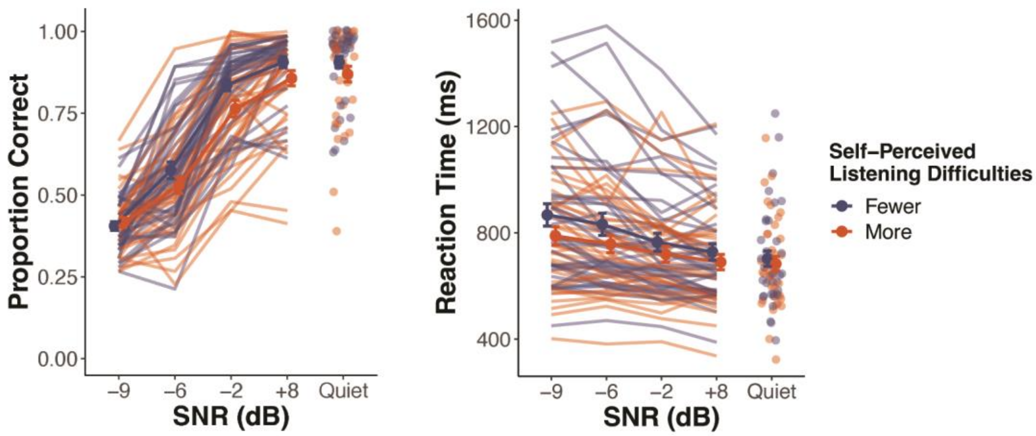

Figure 4: Accuracies and Response Times from the Phoneme in Noise Categorization Task. Left: Accuracies measured in proportion correct at each signal-to-noise ratio (SNR) level and in quiet for individuals with more and fewer self-perceived listening difficulties (SPLDs). Accuracies significantly increased for higher SNR levels. Accuracies did not differ at each SNR level based on SPLDs, although the increase in accuracy from SNR -9 dB to Right: Reaction times measured in ms for each SNR level and in quiet. Reaction times significantly decrease for easier SNR levels. For visualization purposes only, SPLDs are based on a median split of Speech, Spatial, and Qualities of Hearing Scale (SSQ) scores. In both panels, the solid points and thick lines denote averages, and each individual lighter line or point (for Quiet) denotes a single participant's data. Error bars denote standard error of the mean.

2.3.2 Decision Processes During Phoneme in Noise Categorization

Linear mixed effects models were separately fit for evidence accumulation rate and decision threshold parameters derived from the DDMs to understand the effect of SNR level and SPLDs on decision processes during phoneme in noise categorization (Figure 5). Multiple comparisons were performed to assess differences between each SNR level. Benjamini-Hochberg adjusted *p*-values are reported. The maximal random effect structure for the linear mixed effects model on evidence accumulation rates that resulted in a non-singular fit included random intercepts for subject and age. The results from this model (Table 4) revealed that evidence accumulation rates significantly increased with higher SNR levels (ps < .001). There were also significant effects of SSQ at SNR 8 (p = .02), SNR -2 (p = .02), and SNR -6 (p = .05). However, the effect of SSQ was not significant at -9 dB SNR. These results indicated that listeners with fewer SPLDs had higher evidence accumulation rates at all SNR levels except for the lowest SNR level when interference from background noise was greatest. Lastly, there were significant interactions of SSQ with the increase in evidence accumulation rate between SNR levels -9 and -2 (p = .046), as well as the increase in evidence accumulation rate between SNR -9 and 8 (p = .046). These significant interactions suggest that listeners with more SPLDs received less benefit from decreasing background noise than those with fewer SPLDs. As such, listeners with more SPLDs were not as efficient at extracting relevant information from the stimulus to make a decision compared to those with fewer SPLDs.

Table 4: Linear Mixed Effects Model Results with Multiple Comparisons for Evidence Accumuation Rate. Note: SNR = Signal-to-noise Ratio; SSQ = Speech, Spatial, and Qualities of Hearing Scale; Benjamini-Hochberg adjusted *p*-values are reported; ****p* < .001; **p* ≤ .05.

Fixed Effect	Estimate	*SE*	*t* value	*p* value
SNR 8 vs. SNR -2	0.245	0.030	8.118	< .001***
SNR 8 vs. SNR -6	0.960	0.030	31.877	< .001***
SNR 8 vs. SNR -9	1.536	0.030	51.004	< .001***
SNR -2 vs. SNR -6	0.716	0.030	23.759	< .001***
SNR -2 vs. SNR -9	1.292	0.030	42.886	< .001***
SNR -6 vs. SNR -9	0.576	0.030	19.126	< .001***
SSQ score at SNR 8	-0.857	0.316	-2.714	.020*
SSQ score at SNR -2	-0.832	0.316	-2.634	.020*
SSQ score at SNR -6	-0.667	0.316	-2.112	.050*
SSQ score at SNR -9	-0.512	0.316	-1.622	.109
SNR 8 vs. SNR -2 X SSQ Score	0.006	0.029	0.193	.848
SNR 8 vs. SNR -6 X SSQ Score	0.042	0.029	1.452	.286
SNR 8 vs. SNR -9 X SSQ Score	0.076	0.029	2.635	.046*
SNR -2 vs. SNR -6 X SSQ Score	0.036	0.029	1.260	.286
SNR -2 vs. SNR -9 X SSQ Score	0.071	0.029	2.443	.046*
SNR -6 vs. SNR -9 X SSQ Score	0.034	0.029	1.183	.286

As for decision thresholds, the maximal random effect structure for the linear mixed effects model contained a single random intercept for subject. All other random effect structures resulted in a model with a singular fit or that did not converge. The linear mixed effects model revealed significant effects of SNR level (*p*s < .05; Table 5). Specifically, decision thresholds tended to increase with higher SNR levels. However, SNR -2 was the only SNR level that had a higher decision threshold than SNR 8 (*p* = .003). There were also significant effects of SSQ score at each SNR level (*p*s < .05). This indicates that listeners with more SPLDs had lower decision thresholds than those with fewer SPLDs. Lower decision thresholds are an indicator of less cautious responding (Bogacz et al., 2010), suggesting that listeners with more SPLDs tended to favor speed over accuracy of decision making. However, there were no significant interactions between SSQ score and SNR levels, which indicates that the change in decision thresholds between SNR levels did not significantly differ as a function of SPLDs. Collectively, the findings from the DDM suggest that listeners with more SPLDs were less efficient at extracting critical information from a stimulus with decreasing background noise and were more impulsive decision-makers during the phoneme in noise categorization task.

Table 5: Linear Mixed Effects Model with Multiple Comparisons for Decision Thresholds. Note: SNR = Signal-to-noise Ratio; SSQ = Speech, Spatial, and Qualities of Hearing Scale; Benjamini-Hochberg adjusted *p*-values are reported; ****p* < .001; ***p* < .01 **p* ≤ .05.

Fixed Effect	Estimate	*SE*	*t* value	*p* value
SNR 8 vs. SNR -2	-0.025	0.008	-3.025	.003**
SNR 8 vs. SNR -6	0.021	0.008	2.525	.012*
SNR 8 vs. SNR -9	0.050	0.008	6.005	< .001***
SNR -2 vs. SNR -6	0.046	0.008	5.550	< .001***
SNR -2 vs. SNR -9	0.075	0.008	9.030	< .001
SNR -6 vs. SNR -9	0.029	0.008	3.480	.001**
SSQ score at SNR 8	-0.477	0.186	-2.566	.012*
SSQ score at SNR -2	-0.500	0.186	-2.685	.012*
SSQ score at SNR -6	-.508	0.186	-2.731	.012*
SSQ score at SNR -9	-0.517	0.186	-2.777	.012*
SNR 8 vs. SNR -2 X SSQ Score	-0.005	0.008	-0.616	.812
SNR 8 vs. SNR -6 X SSQ Score	-0.007	0.008	-0.854	.812
SNR 8 vs. SNR -9 X SSQ Score	-0.009	0.008	-1.093	.812
SNR -2 vs. SNR -6 X SSQ Score	-0.002	0.008	-0.238	.812
SNR -2 vs. SNR -9 X SSQ Score	-0.004	0.008	-0.477	.812
SNR -6 vs. SNR -9 X SSQ Score	-0.002	0.008	-0.239	.812

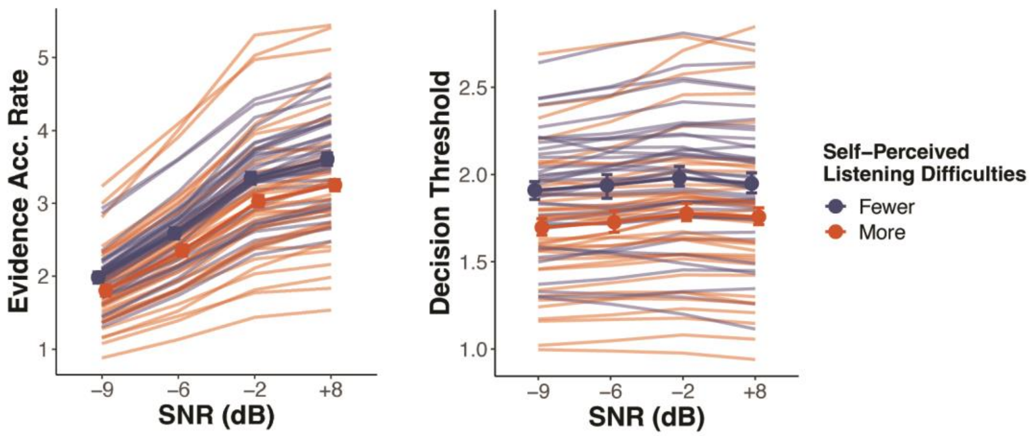

Figure 5: Evidence Accumulation Rates and Decision Thresholds. Left: Evidence accumulation rates at each signal-to-noise ratio (SNR) level split by self-perceived listening difficulties (SPLDs). Evidence accumulation rates significantly increased with higher SNR levels. Overall, individuals with more SPLDs had lower evidence accumulation rates with higher SNR levels. Right: Decision thresholds at each SNR level. Individuals with more SPLDs had lower decision thresholds across SNR levels. For visualization purposes only, SPLDs are based on a median split of Speech, Spatial, and Qualities of Hearing Scale (SSQ) scores. In both panels, the solid points and thick lines denote averages, and each individual lighter line denotes a single participant's data. Error bars denote 95% credible intervals.

2.4 Discussion

This study investigated the extent to which individual differences in evidence accumulation rates and decision thresholds during phoneme in noise categorization were related to SPLDs in adults with normal hearing thresholds. Participants with more SPLDs had slower evidence accumulation rates and lower decision thresholds than those with fewer SPLDs. Additionally, participants had similar accuracies, regardless of SPLDs. These results suggests that dysfunctional decisional processes that support speech in noise categorization may underlie an aspect of SPLDs.

2.4.1 Self-Perceived Listening Difficulties Impact Evidence Accumulation

While there were similar categorization accuracies across SPLDs, there were differences in the decisional processes that support categorization (i.e., evidence accumulation rates and decision thresholds). Evidence accumulation rates reflect the efficiency with which listeners extract critical information from the stimulus for categorization. Neural activity in frontoparietal regions has been shown to be associated with evidence accumulation rates (Gold & Shadlen, 2007; Heekeren et al., 2004; Mulder et al., 2014), as well as selective attention processes during speech perception in noise (Shinn-Cunningham, 2017). This suggests there may be an overlap in the neural regions that support the accumulation of critical evidence from an auditory stimulus for categorization in noisy listening situations and the cognitive processes for speech perception.

In the current study, evidence accumulation rates increased with higher SNR levels. At the most difficult SNR level, evidence accumulation rates were low, indicating that listeners were slower to extract information from the phoneme stimulus to make a category decision. At the easiest SNR level, evidence accumulation rates were significantly higher than at more difficult SNR levels. Higher evidence accumulation rates indicated that listeners were more efficient and faster to gather information from the stimulus for categorization when there was less interference from background noise.

There was also a clear differentiation in evidence accumulation rates as a function of SPLDs. Specifically, listeners with fewer SPLDs had a greater increase in evidence accumulation rates between the hardest and easiest SNR levels compared to those with more SPLDs. This difference in evidence accumulation rates based on SPLDs suggests that background noise interferes with the ability of listeners with more SPLDs to efficiently accrue information for the stimulus for categorization to a greater extent than those with fewer SPLDs. Thus, those with more SPLDs were inefficient evidence accumulators in the presence of background noise.

2.4.2 Listeners with More Self-Perceived Listening Difficulties are Impulsive Responders

A second component of decisional processes is the decision threshold. Decision thresholds reflect the amount of information a listener needs to accumulate before a category decision can be made. The decision threshold also reflects response-caution, which is a trade-off between speed and accuracy of categorization. Participants were instructed to categorize the sounds as quickly and as accurately as possible. Yet, the decision thresholds differed based on SPLDs, regardless of the SNR level. The difference in decision thresholds across all SNR levels suggests that the amount of information listeners required to accumulate before making a categorization decision was independent of the level of background noise. Specifically, listeners with more SPLDs had lower decision thresholds than those with fewer SPLDs at all SNR levels. Lower decision thresholds reflect more impulsive responses and the prioritization of speed over the accuracy of the categorization decision. Therefore, listeners with more SPLDs were willing to make faster and riskier responses that had a higher likelihood of being inaccurate.

2.4.3 Neural Implications for Self-Perceived Listening Difficulties

The results from this study demonstrate that the top-down decisional processes that support speech sound categorization in noise are disrupted in listeners with more SPLDs. Prior research indicates that word in noise recognition is associated with decision thresholds and activity in the CON (Vaden et al., 2022). Increased activity in the CON prior to correct word in noise recognition is associated with higher decision thresholds, and thus, more cautious responding. As such, individuals with higher decision thresholds may benefit from more pre-stimulus CON activity. In the context of the current study, listeners with fewer SPLDs may have higher pre-stimulus CON activity, which may help to serve as an adaptive control mechanism in noisy listening situations to promote speech perception in noise.

Furthermore, the neural basis of information accumulation suggests that groups of neurons are tuned to different decision alternatives and their firing rates increase based on the integration of information from sensory input neurons (Bogacz et al., 2010; Gold & Shadlen, 2007; Heekeren et al., 2008). A decision is made once the firing rates among a group of these integrator neurons reaches the given firing rate decision threshold (Roitman & Shadlen, 2002). For listeners with more SPLDs, these findings suggest that there may be a smaller distance between baseline firing rates of integrator neurons and their decision threshold, resulting in risky and impulsive decision-making. Taken together, these findings could indicate that listeners with more SPLDs have a smaller dynamic range of activity in the CON. Additional studies are needed that specifically investigate individual differences in baseline CON activity and during speech in noise processing to understand how that may impact SPLDs in adults with normal hearing thresholds.

Activity in the CON may not be the only neural basis of SPLDs. It is also plausible that listeners with more SPLDs are impulsive responders from a combination of poorer stimulus encoding quality and less efficient extraction of information from the stimulus. Future research is necessary to rule out temporal processing deficits in listeners with more SPLDs that may contribute to deficits in the decisional processes that support speech in noise categorization.

2.4.4 Study Limitations

This study provides evidence for a link between individual differences in top-down decisional processes and SPLDs in adults with normal hearing thresholds. While all participants had pure tone air conduction hearing thresholds within normal range, there were a lack of measures to characterize middle- and inner-ear functioning. Thus, deficits in bottom-up processes on SPLDs cannot be completely ruled out. Future studies should collect a full audiometric profile, as well as examine the extent to which bottom-up processes, such as temporal processing, may be linked with the decisional processes that support speech categorization.

2.4.5 Conclusion

Complaints of listening difficulties are prevalent among adults with normal hearing thresholds. Often, these SPLDs are considered to be driven by bottom-up processing deficits. Prior research into SPLDs has focused almost exclusively on potential deficits in bottom-up processes, with little consideration into potential dysfunctions in top-down cognitive processes. The results from the current study demonstrate that top-down decisional processes that support speech categorization were disrupted in adults with more SPLDs. Listeners with more SPLDs were less efficient evidence accumulators and were more impulsive responders in noisy listening situations. Moreover, these deficits in the decisional processes supporting categorization may be associated with activity in the CON. These findings encourage further research into top-down cognitive processes in adults with normal hearing thresholds who have SPLDs.

3.0 Study 2: Neural Tracking of Linguistic Information in Continuous Speech Differs Based on Self-Perceived Listening Difficulties

3.1 Introduction

Communicating in noisy environments is a difficult task that our brains perform exceptionally well. In a crowded restaurant, there are multiple conversations happening around you. Each of these speech streams, including the speech from your friend across the table, gets encoded in the auditory system. The competing speech streams have overlapping acoustic and linguistic properties with your friend's speech, yet you are able to ignore the irrelevant speech to focus exclusively on your friend's voice. While most listeners can perform this task with relative ease, some individuals particularly struggle with speech perception in noise. Speech comprehension difficulties can derive from factors such as age-related hearing loss (Helfer & Wilber, 1990; Pichora-Fuller et al., 1995) and noise exposure (Hope et al., 2013; Liberman et al., 2016b). However, speech comprehension difficulties can still occur when hearing thresholds remain intact (Chandrasekaran et al., 2015; Lam et al., 2017; Xie et al., 2014, 2015; Xie, Zinszer, et al., 2019). Approximately ten percent of clients who present to audiology clinics with complaints of hearing difficulties have no abnormal indicators of hearing health (Hind et al., 2011; Parthasarathy et al., 2020; Pryce & Wainwright, 2008; Spankovich et al., 2018; Tremblay et al., 2015). Much prior research has focused on the roles of pitch tracking (Anderson et al., 2010), audiovisual cues (Smayda et al., 2016; Xie et al., 2014), and cognitive resources in speech comprehension performance (Rönnberg et al., 2010; Zekveld et al., 2011). However, the cause of self-perceived listening difficulties (SPLDs) in adults with normal hearing thresholds remains unclear.

During speech processing, the incoming speech signal is first broken down into the time and frequency representations at the level of the cochlea (Moore, 2008; Shamma, 1985). The time and frequency information then travels along the ascending auditory pathway to the auditory cortex, where the listener must decode this information into meaningful speech (Poeppel et al., 2008). When background noise is present, it can be more difficult to reconstruct the auditory signal into meaningful speech because the noise can interfere with the perception of individual phonemes and words (Committee on Hearing Bioacoustics and Biomechanics, 1988). In cases where background noise leads to gaps in the perceived speech stream, the listener must use stored linguistic knowledge and memories to correctly reconstruct the speech signal based on contextual information in the speech (Pichora-Fuller, 2008). According to the information theory of communication (Shannon, 1948), listeners predict upcoming sounds based on prior context and their stored linguistic knowledge, which is known as conditional probability. The conditional probability of the next phoneme or word is continuously updated based on the prior perceived input. Thus, conditional probability requires a listener to successfully access their stored linguistic knowledge (Gwilliams & Davis, 2022).

From conditional probability, two other measures can be derived that relate to how listeners extract linguistic content from speech: 1) entropy and 2) surprisal. Entropy is a probability measure of the amount of uncertainty for an upcoming sound, word, or other linguistic component. If entropy is high, the level of uncertainty is high. At the phoneme level, this suggests that based on a sequence of phonemes, the listener is uncertain what the next phoneme may be. On the other hand, surprisal is a measure of the overall information that is gained, i.e., how unpredictable is an event given the preceding context (Shannon, 1948). For example, surprisal is high when a listener predicts a particular phoneme is upcoming next based on the preceding phonemes, but a different phoneme is spoken instead. More information is gained about the linguistic input when an unexpected phoneme is spoken than if the predicted phoneme had been presented, resulting in higher surprisal. Measuring a listener's surprisal and entropy measures would inform on their access to higher-level linguistic knowledge during listening.

Recent methodological advancements have made it possible to measure conditional probability during listening using neural tracking of continuous speech. Neural tracking is a metric of the coherence of cortical responses to the ongoing fluctuations in the acoustic speech envelope (Lalor & Foxe, 2010; Obleser & Kayser, 2019). Moreover, neural tracking of the acoustic speech envelope has been shown to also be modulated by linguistic features of the attended speech, such as phonemes, words, and sentences (Brodbeck et al., 2022; Brodbeck & Simon, 2020; Xie et al., 2023). Greater neural tracking of phonemes, words, and sentences reflects enhanced surprisal and/or entropy measures. Therefore, neural tracking of linguistic information can serve as a metric of a listener's reliance on higher-level linguistic knowledge during speech processing. The extent to which neural tracking of linguistic information at different contextual levels (i.e., phonemes, words, sentences) in continuous speech may differ based on SPLDs remains unclear.

In the current study, I examined the extent to which adults with normal hearing thresholds who report SPLDs differ in their reliance on linguistic knowledge during listening, as measured by neural tracking of linguistic information in speech in quiet or masked in speech-shaped noise (SSN). The Speech, Spatial, and Qualities of Hearing (SSQ) questionnaire, a self-report of hearing abilities, was used as a metric of SPLDs because the SSQ has consistently been shown to accurately reflect an individual's perceived listening abilities and capacities (Demeester et al., 2012; Noble et al., 2008, 2013; Noble & Gatehouse, 2004, 2006; Singh & Pichora-Fuller, 2010). I first hypothesized that speech comprehension performance would be worse in SSN than in quiet and that listeners with more SPLDs would have poorer comprehension overall compared to those with fewer SPLDs. Prior research has demonstrated reduced neural tracking of the acoustic envelope when speech is masked in noise (McHaney et al., 2021). Additionally, neural tracking of linguistic information declines` when attention is diverted away from the target speech (Xie et al., 2023). Therefore, for both acoustic and linguistic features, I predicted that neural tracking would be greater in quiet than in SSN.

Lastly, I predicted that neural tracking of linguistic information would not differ based on SPLDs when listening to speech in quiet, but that differences would be present when listening to speech masked in SSN. Listeners often complain of difficulties listening in noisy environments (Dinino et al., 2021; Füllgrabe & Rosen, 2016; Mai et al., 2018; Pichora-Fuller et al., 1995). Therefore, I predicted that neural tracking of linguistic information would be greater in SSN for listeners with more SPLDs relative to those with fewer SPLDs. However, the extent to which neural tracking of linguistic information at different contextual levels differs based on SPLDs is unclear. On the one hand, listeners with more SPLDs may have greater neural tracking for each contextual level. Alternatively, listeners with more SPLDs may only have enhanced neural tracking at specific contextual levels of linguistic information, which would suggest that there may be greater reliance on specific linguistic information during listening. Overall, the findings from this study will help to inform our understanding on the sources of SPLDs in adults with normal hearing.

3.2 Methods

3.2.1 Participants

A total of 63 adults between the ages of 18-52 years (*M* = 26.714, *SD* = 10.726; Figure 6A) were recruited for this study from the greater Pittsburgh, PA community. Participants were self-reported native English speakers. All participants were required to have hearing thresholds ≤25 dB for frequencies 250-4000 Hz in octave steps and normal otoscopy. Monetary compensation or research course credit was provided for participation. This research protocol was approved by the Institutional Review Board at the University of Pittsburgh.

3.2.2 Behavioral Measures

3.2.2.1 Audiometry

The hearing sensitivity for all participants was evaluated using pure tone audiometry. Air conduction hearing thresholds were measured via insert earphones for octave frequencies 250 through 8000 Hz. A modified Hughson and Westlake method (Carhart & Jerger, 1959) was used to estimate the threshold at each frequency. Individual hearing thresholds are displayed in Figure 6B.

3.2.2.2 Speech, Spatial, and Qualities of Hearing

All participants completed the Speech, Spatial, and Qualities of Hearing Scale (SSQ) (Gatehouse & Noble, 2004). The SSQ is a self-report measure of hearing abilities across several domains, including speech hearing, spatial hearing, and other qualities of hearing. Fourteen questions in the speech hearing section assess realistic speech contexts such as competing sounds, differences in background conditions, and visibility of other talkers (e.g., Can you have a conversation with someone when another person is speaking whose voice is the same pitch as the person you're talking to?). The spatial hearing questions (n = 17) assess directional and distance aspects of hearing (e.g., In the street, can you tell how far away someone is, from the sound of their voice or footsteps?). The final section of the SSQ has eighteen questions that address other hearing qualities, such as clarity, recognition, and listening effort (e.g., Do you find it easy to recognize different people you know by the sound of each one's voice?). All questions are on a Likert-scale from 0-10.

Domain scores for each section of the SSQ were calculated for each participant by averaging the responses in each section. The domain scores were then averaged into a composite SSQ score (Figure 6C), which was used as a metric of SPLD. Participant were median split into groups of more or fewer SPLDs (*Mdn* = 8.47). Higher SSQ scores reflect fewer SPLDs, while lower SSQ scores reflect more SPLDs.

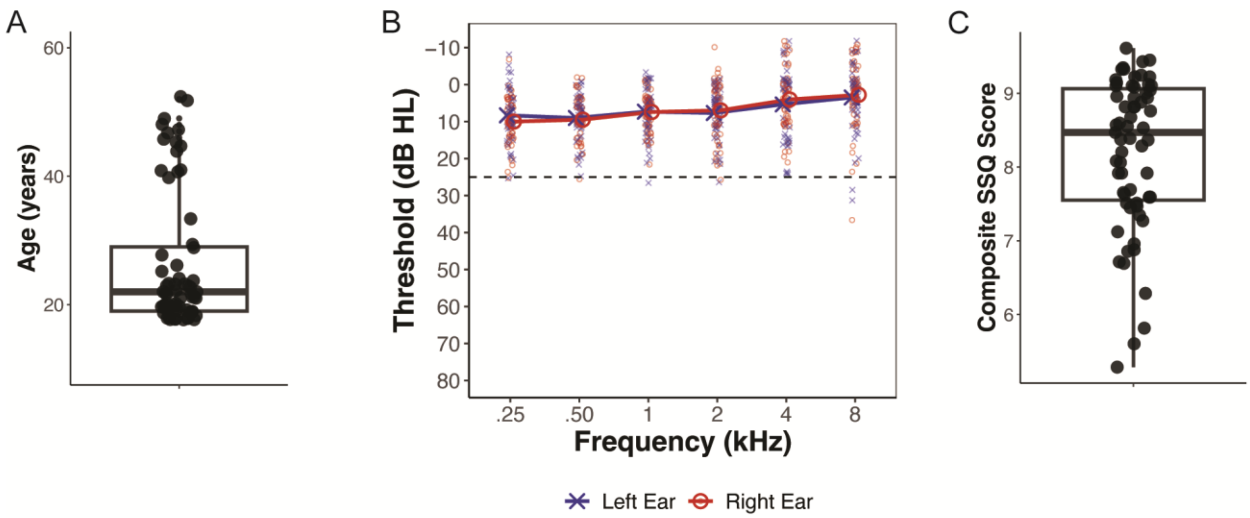

Figure 6: Participant information. A) The age distribution of pariticpants (18-52 years). B) Audiogram thresholds for all paritcipants. All thresholds were ≤ 25 dB for octave frequencies .25-4 kHz. C) Distribution of comoposite Speech, Spatial, and Qualitiess of Hearing Scale (SSQ). Higher scores reflect fewer self-perceived listening difficultiess (SPLDs, and lower composite SSQ scores reflect more SPLDs.

3.2.3 Continuous Speech Stimuli

Participants listened to continuous speech from the public domain audiobook *Alice's Adventures in Wonderland* (Carroll, 1865). The story was read in American English by a male speaker and sampled at a frequency of 22.05 kHz. Long speaker pauses were manually truncated to a maximum of 500 ms, and the resulting speech from the audiobook was divided into segments approximately 60 seconds in duration each (min: 57 s, max: 65 s). Segments began and ended with complete sentences. The mean syllable rate was ~6 Hz. Participants listened to the story in quiet, masked in speech-shaped noise (SSN) at -2 dB SNR, and masked in time-reversed speech at -2 dB SNR. In the quiet condition, participants listened to the normal audiobook story with no background noise. The SSN condition consisted of the audiobook story masked in SSN. The SSN masker was derived from all sentence stimuli from QuickSIN test lists 1-8 using the following steps. First, all stimuli files were combined, and a fast Fourier Transform (FFT) was computed. Then, the phases of the spectral components from the FFT were randomized. An inverse Fourier Transform was then performed to convert the modified FFT back into the time domain. The third time-reversed speech condition consisted of the same speech as in the quiet condition but overlapped with time-reversed speech (noise) at -2 dB SNR from a different portion of the story that was unheard by the participants in this experiment. The focus of the current EEG study was to examine acoustic and linguistic features in response to the target speech during noisy listening situations. The masker in the time-reversed speech condition also contained some aspects of linguistic information. The analyses in this study do not include the reversed-speech condition because linguistic predictors for the reversed-speech masker could not be created. Therefore, it cannot be guaranteed that linguistic information from the masker was not represented in the analysis to the target speech. Thus, the analyses in the current study focused exclusively on the quiet and SSN listening conditions.

The order of conditions was counterbalanced across participants. Fifteen segments of the story were presented in each condition in chronological order, regardless of condition order, such that the storyline was preserved without repetitions or discontinuities. Participants were presented with two multiple-choice comprehension questions with four answer choices after each segment assessing speech comprehension of the preceding segment. Participants were provided feedback on their performance in the format of “You got X correct out of 2 in this trial.” The average accuracy in each condition per participant was used for further analyses. Stimuli and comprehension questions were presented using E-Prime 3.0 (Psychology Software Tools, Pittsburgh, PA).

3.2.4 EEG Acquisition and Processing

Electrophysiological responses to continuous speech were amplified and digitized with BrainVision actiCHAMP amplifier and collected using BrainVision PyCorder 1.0.7 (Brain Products, Gilching, Germany) with 64-channel actiCAP active electrodes (Brain Products) secured in an elastic cap (EasyCap; http://www.easycap.de/). Electrodes were placed on the scalp according to the International 10-20 system (Klem et al., 1999), and a common ground was placed at the AFz electrode site. Electrode impedance was less than 20 kΩ for all channels. Responses were recorded at a sampling rate of 25 kHz.

The EEG data were preprocessed using EEGLAB 14.1.2 (Delorme & Makeig, 2004) in MATLAB (MathWorks Inc., Natick, Massachusetts, USA) with the following steps. The raw EEG data were downsampled to 128 Hz to improve computational efficiency. The down-sampled EEG was then filtered using minimum-phase causal windowed sinc FIR filters. The high pass filter cut-off frequency was set at 1 Hz with a filter order of 846, while the low pass filter cut-off frequency was set at 15 Hz with a filter order of 212. The filtered EEG was then re-referenced to the average of the two mastoid channels (Di Liberto et al., 2015, 2018; O'Sullivan et al., 2014). Re-referenced channels with electrical activity greater/lower than 3 standard deviations of the surrounding channels were rejected, and their data were interpolated based on the activity in the surrounding channels using spherical spline interpolation. Artifacts in the EEG data were suppressed using artifact subspace reconstruction (ASR) (Mullen et al., 2015). Visually identified clean sections (~one minute) of the data were input as the calibration data for the ASR. The ASR cleaned data were then separated into epochs from -5 seconds to 70 seconds (re: trial onset), yielding 15 epochs in each condition. Independent component analysis (ICA) was performed on the epoched data to remove eye-movement and muscle artifacts. ICA was performed using the infomax algorithm and was adjusted to extract only 50 components from the data to account for the reduced rank following channel interpolation and referencing. The independent components to be removed were visually identified based on time-course, topography, and spectrum. Components corresponding to ocular and muscular activity were removed. The clean EEG was reconstructed from the remaining independent components.

3.2.5 Estimation of Neural Tracking

Neural tracking of the speech envelope was estimated using the multivariate temporal response function (mTRF) approach (Crosse et al., 2016; Ding & Simon, 2012). I first estimated several acoustic and linguistic models (described in detail in subsequent sections) to predict the EEG signal using multiple time-delayed regressions. Each model implemented a model-specific hypothesis for predicting brain activity from predefined predictor variables. To evaluate the extent to which acoustic and linguistic models were associated with predictive power of the EEG, I tested the predictive power of different combinations of predictor variables. Thus, each predictor variable reflects the extent to which EEG responses were modulated by an acoustic or linguistic property of the speech signal (Brodbeck et al., 2022; Brodbeck, Hong, et al., 2018; Brodbeck & Simon, 2020). This approach allowed us to assess the extent of neural representations of the acoustic and linguistic properties of the target speech in quiet and in SSN. Examples of each acoustic and linguistic predictor variable are displayed in Figure 7.

Forward encoding mTRF models were fit using the boosting algorithm in the Eelbrain toolbox (Brodbeck et al., 2021) to predict EEG responses from the acoustic and linguistic predictors. A *z*-transformed Pearson's correlation between the predicted and actual EEG responses served as a metric for how accurately the mTRF models could predict EEG responses to novel trials from the same listening condition. Here, higher prediction accuracy corresponds to greater neural tracking of the given predictor. mTRFs were estimated separately for each participant and condition using a 5-fold cross-validation strategy. The EEG trials were divided into five test sets. EEG responses in each test set were predicted from the average of four mTRF models that were estimated from the remaining four test sets, wherein three of the four test sets served as training data and the fourth test set served as the validation data. A 50 ms Hamming basis function was used to reduce sparseness with a stimulus-EEG lag from -100 to 500 ms and to generate the mTRFs. The mTRFs were estimated for all acoustic and linguistic predictors with coordinate descent to minimize the ℓ_1 error. The predicted responses from the five test sets were concatenated together before deriving the prediction accuracy measure. The mTRFs were then averaged across all five test sets for analysis purposes.

To estimate neural tracking of a given acoustic or linguistic predictor, the change in prediction accuracy was calculated when the predictor(s) of interest was(ere) removed from the Full model that included all acoustic and linguistic predictors. This change in prediction accuracy (Δz) reflects the variability in EEG responses that can be uniquely attributed to the predictor(s) of interest and cannot be attributed to any other predictors. This method is appropriate for analyzing predictors of interest because properties of naturalistic speech tend to be correlated in time. The Full model differed in each condition. In the SSN listening condition, the Full model was comprised of all linguistic predictors, plus the acoustic predictors to the target speech, SSN masker, and the mixture of the target speech and SSN masker. The Full model in the quiet condition consisted of only the acoustic predictors to the target speech and the linguistic predictors.

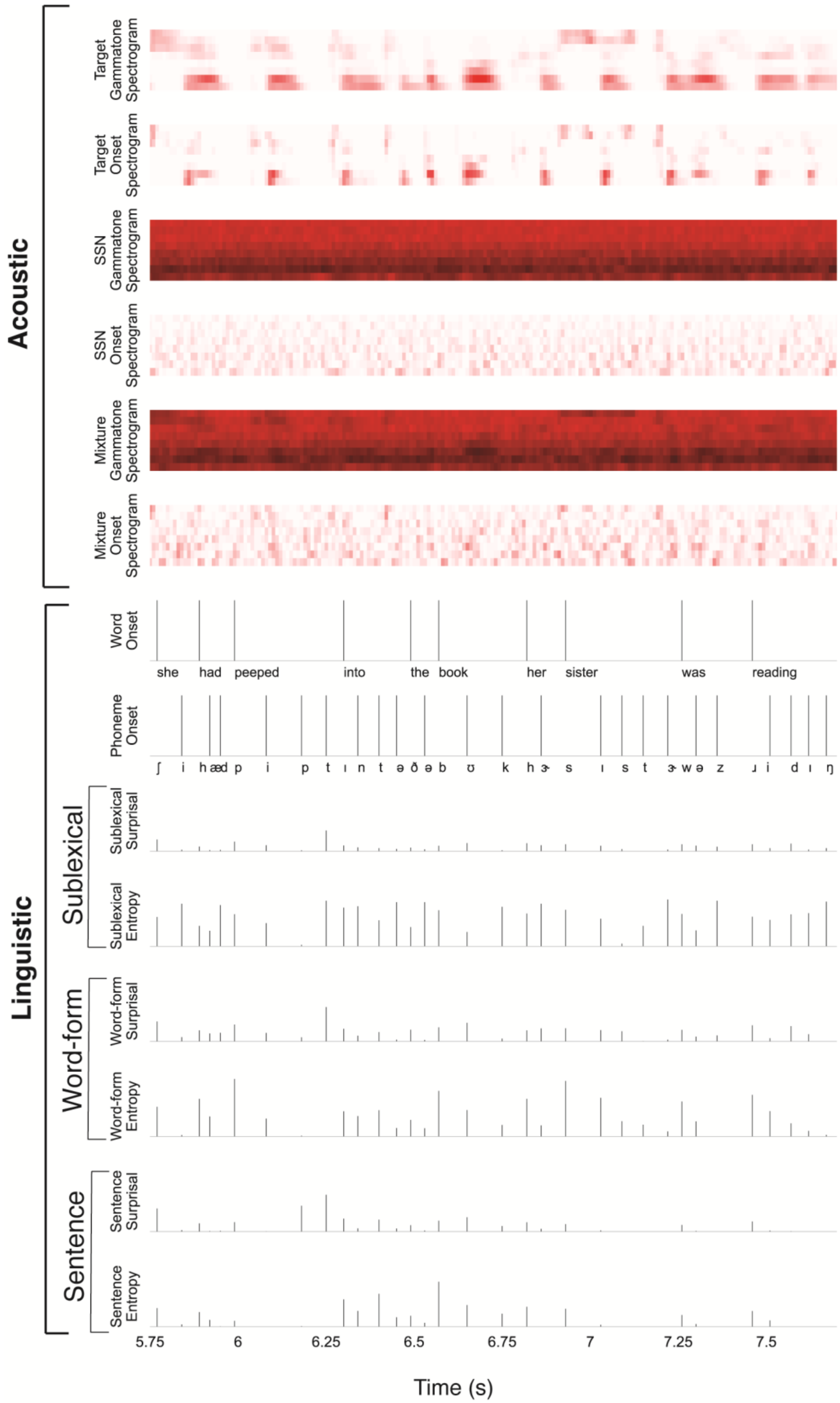

Figure 7: Stimulus Features for Acoustic and Linguistic Predictors.

3.2.6 Acoustic Model

The acoustic model was designed to assess the EEG in response to acoustic spectro-temporal features in the stimuli. Acoustic predictors were created separately for the target speech, the SSN masker, and the mixture of the target speech and SSN masker. The acoustic predictors were derived from 256-band gammatone-based spectrograms of the given stimuli, with cut-off frequencies of .02 to 5 kHz. The 256-band spectrograms were down sampled to 1 kHz and scaled with an exponent of 0.6. A *spectrogram* predictor was then created by summing the 256-band spectrograms in eight logarithmically spaced frequency bands. An *onset spectrogram* predictor was also defined to control for representations of acoustic edges using an auditory edge detection model (Fishbach et al., 2001). The auditory edge detection model was applied to each frequency band from the 256-band spectrogram, using the parameters as outlined in Brodbeck et al. (2020). Eight logarithmically spaced frequency bands were created from the sum of the onsets across the 256-band spectrograms to generate 8-band onset spectrogram predictors.

3.2.7 Linguistic Models

Information-theoretic models were used to assess linguistic processing. Information theoretic models assume that listeners maintain predictive models of speech that can be associated with brain activity through measures of *surprisal* and *entropy* (Brodbeck, Hong, et al., 2018). As such, predictive models were defined across phoneme sequences for the target speech using a Montreal Forced Aligner (McAuliffe et al., 2017). Three information-theoretic models were created: 1) sublexical model, 2) word-form model, and 3) sentence model. Each model consisted of a time-series of pulses at each phoneme onset. A word onset predictor was also included to control for neural responses at the onset of each word-initial phoneme. Lastly, a phoneme onset predictor was included to control for neural responses to the onsets of all other phonemes.

3.2.7.1 Sublexical Model

The sublexical model assumes that listeners use local context to predict upcoming phonemes based on prior sounds. A 5-gram model (Heafield, 2011) was trained on phoneme sequences from the SUBTLEX-US corpus (Keuleers & Brysbaert, 2010). This 5-gram model functions such that a probability distribution is computed for each phoneme in the audiobook target stimuli that is based on the four preceding phonemes. Sublexical surprisal and sublexical entropy predictors were calculated based on the probability distribution. Sublexical surprisal captures the surprisal of a given phoneme at a given position and measures the amount of new information provided by a stimulus. Sublexical entropy reflects the uncertainty of what an upcoming phoneme will be. An entropy response indicates that listeners predict future input before encountering that phoneme (Pickering & Gambi, 2018).

3.2.7.2 Word-form Model

The word-form model assumes that listeners use a probabilistic model to recognize a word that is currently being heard based on all the information since the start of the current word. Notably, the word-form model does not account for any informational context prior to the start of the current word. The model uses the cohort model of word recognition (Brodbeck, Hong, et al., 2018), which assigns each word in the stimuli a probability based on its count frequency in the SUBTLEX corpus (Keuleers & Brysbaert, 2010). The cohort model was applied to each word in the stimuli, which begins with the complete lexicon at the first phoneme position and removes words that are no longer compatible for each subsequent phoneme. The resulting probability distribution provides phoneme surprisal and cohort entropy predictors. Phoneme surprisal was the inverse log of the posterior probability of a given phoneme based on the preceding phonemes. Cohort entropy is the amount of information over all the words in the cohort.

3.2.7.3 Sentence Model

The sentence model assumes that listeners use a wider, global context to modulate their phoneme-by-phoneme expectations and functions similarly to the word-form model. However, the sentence model differs from the word-form model in that the expectation for each word changes based on sentence context. A lexical 5-gram model (Heafield, 2011) was trained on the SUBTLEX-US corpus (Keuleers & Brysbaert, 2010) to create probabilities for each word in the lexicon based on the preceding four words. The resulting probability distributions provide predictors of phoneme surprisal and cohort entropy. For the sentence model, phoneme surprisal was the inverse log of the posterior probability of a given word based on the preceding words, and cohort entropy is the amount of information over all words in the cohort.

3.2.8 Statistical Analyses

Statistical tests were conducted using R, version 4.1.3 (R Core Team, 2022), unless otherwise noted. The following R packages were used: *ggplot2,* version 3.3.6 (Wickham, 2016) and *rstatix* version 0.7.0 (Kassambara, 2021).

First, associations between age and comprehension scores, prediction accuracy, and SSQ scores were assessed to examine whether participant ages were related to speech comprehension and neural tracking. If age showed a significant relationship with comprehension scores, prediction accuracy, or SSQ scores, then age was included as a covariate in the analyses. Two-way repeated measures analysis of variances (ANOVA) were performed with an alpha level of .05 to examine the effect of listening condition and SPLD group on speech comprehension scores and neural processing of acoustic and linguistic information. Tukey post hoc tests were performed on any significant effects or interactions.

Neural processing of acoustic and linguistic information was calculated using mass-univariate analysis by comparing the predictive accuracy (z) between the Full model and a model missing the predictors of interest. For each acoustic and linguistic model, I first averaged the predictive accuracy across both the quiet and SSN listening conditions at individual electrodes. Then, I used a mass-univariate, single-sample t-test to test whether the averaged difference in prediction accuracy (Δz) between the Full model and the model missing the predictors of interest was statistically greater than zero. The mass-univariate is a cluster-based permutation test implemented in the *Eelbrain* package that uses a t-value equivalent to uncorrected $p \leq .05$ as the cluster-forming threshold. Clusters were based on the identification of meaningful effects across groups of adjacent electrodes that showed the same effect (Maris & Oostenveld, 2007). A corrected p-value was then computed for each cluster based on the cluster-mass statistician null distribution from 10,000 permutations (Maris & Oostenveld, 2007). The largest t-value from the cluster (t_{max}) is reported as an estimate of effect size (Brodbeck, Hong, et al., 2018).

To understand the effects of SPLD and listening condition on neural tracking of acoustic information, prediction accuracy for the acoustic model to the *target* speaker was analyzed using a two-way repeated measures ANOVA with a within-subjects independent variable of listening condition (quiet, SSN) and a between-subjects independent variable of SPLD group (more, fewer). For each listening condition, the acoustic model to the target speaker was calculated as the difference between the Full model and a model with all predictors except for the gammatone spectrogram to the target speaker and the gammatone onset spectrogram to the target speaker. The ANOVA was fit to examine the extent to which prediction accuracy of acoustic information from the target speaker differed as a function of SPLDs and listening condition. If the acoustic model showed significant effects or interaction, a follow-up ANOVA was used to examine the extent to which individual predictors, gammatone spectrogram and gammatone onsets spectrogram, contributed to the significant results. Then, a two-way repeated measures ANOVA was used to examine the extent to which prediction accuracy of linguistic features differed as a function of linguistic model (sublexical, word, sentence), listening condition, and SPLDs.

Finally, I examined the effect of SPLDs on the mTRFs for any predictors that showed a significant effect of SPLDs on prediction accuracy. The global field power of the mTRF was calculated across the corresponding region of interest from the prediction accuracy analysis. The global field power of the mTRF was then compared between listeners with more and fewer SPLDs using an ANOVA. The mTRF analyses were implemented in the Eelbrain package with default parameters except for the analysis time window, which was modified to 0 to 500 ms.

3.3 Results

3.3.1 Speech comprehension performance is not affected by SPLDs

First, comprehension question performance was analyzed to examine whether performance differed in quiet and SSN and whether listeners with more SPLDs performed worse than those with fewer SPLDs (Figure 8). The two-way repeated measures ANOVA revealed that there was a significant effect of condition (F(1,58)= 16.237, p < .001, η_p^2 = .219), which indicates that overall speech comprehension performance was greater in quiet (M = .873 SE = .082) compared to the SSN (M = .817, SE = .123) listening condition. However, there was no significant effect of SPLD (F(1,58)= 2.387, p = .128, η_p^2 = .040), nor interaction of condition and SPLD (F(1,58)= 0.858, p = .358, η_p^2 = .015). These null results indicate that speech comprehension performance was not impacted by SPLDs.

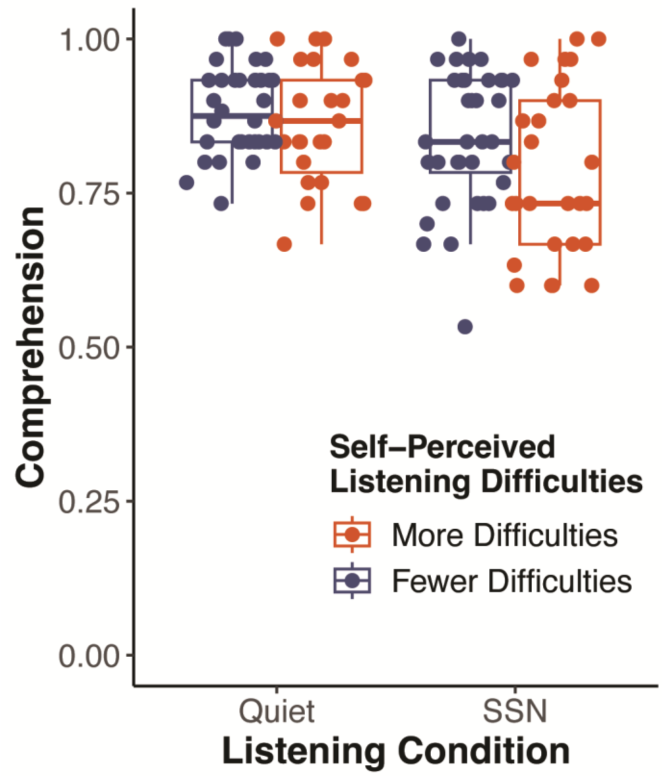

Figure 8: Comprehension Question Performance. Participants were median split into more self-perceived listening difficulties (SPLDs) or fewer SPLDs (Mdn = 8.469). Comprehension question performance in each listening condition and split by more and fewer SPLDs is displayed. Individual data points reflect average performance for each particpant. Performance was significantly poorer overall in the SSN listening condition, but speech comprehension performance did not differ as a function of SPLDs.

3.3.2 Neural tracking of acoustic information is affected by listening condition but not SPLDs

The acoustic predictors of the target speech significantly contributed to model prediction beyond the linguistic predictors in a cluster that spread across all 61 electrodes (t_{max} = 12.082, p < .001; Figure 9A). This indicates that there was robust neural tracking of the acoustic information in speech. Results of the two-way repeated measures ANOVA revealed that prediction accuracy for the acoustic predictors was modulated by listening condition ($F(1,58)$= 98.735, $p < .001$, η_p^2 = .630). Prediction accuracy was significantly lower in the SSN listening condition (M = .004, SD = .002) than the quiet listening condition (M = .014, SD = .007; Figure 9B). Prediction accuracy did not differ by SPLDs ($F(1,58)$= 0.017, p = .898, $\eta_p^2 < .001$). The interaction between listening condition and SPLDs was also not significant ($F(1,58)$= 0.179, p = .674, η_p^2 = .003). These results suggest that neural tracking of acoustic information is reduced in the presence of background noise.

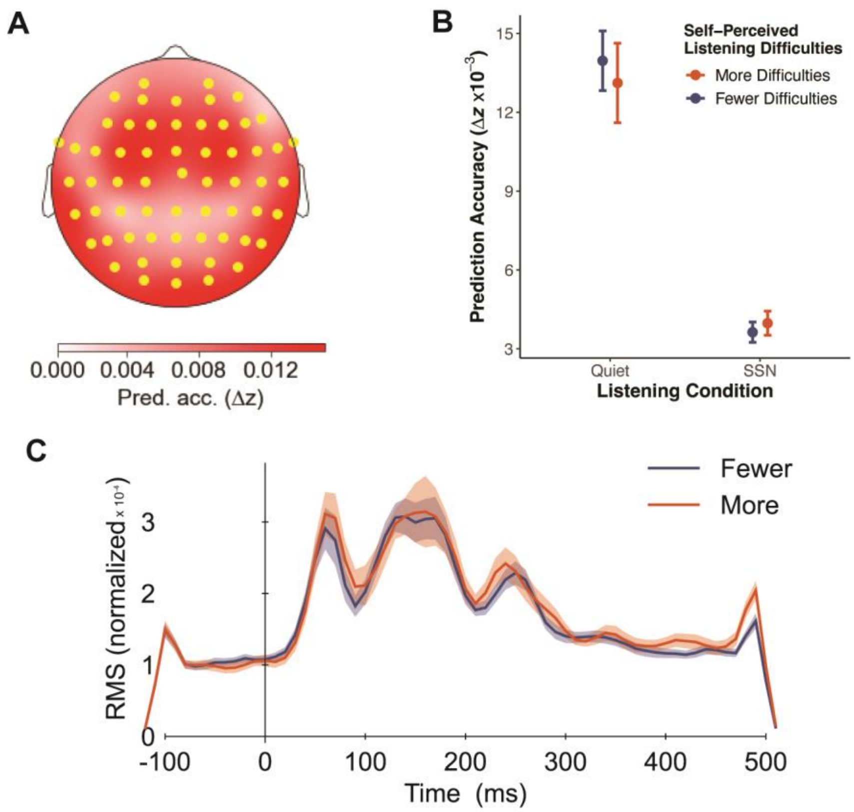

Figure 9: Neural Tracking of Acoustic Information Across Listening Conditions. A) Prediction accuracy to acoustic predictors (gammatone spectrogram and gammatone spectrogram onsets) were significantly above zero in a cluster of electrodes highlighted in yellow. B) Average prediction accuracy in quiet and speech shaped noise (SSN) for listeners with more and fewer self-perceived listening difficulties (SPLDs) for acoustic predictors (i.e., combined gammatone spectrogram and gammatone spectrogram onsets). Error bars denote standard error of the mean. Prediction accuracy decreases in SSN but does not differ by SPLDs. C) Global field power of multivariate temporal response functions (mTRFs) by SPLDs. Shaded areas denote standard error of the mean.

Next, I assessed the extent to which specific acoustic predictors of the target speech contributed to the condition effect. Both acoustic predictors significantly contributed to overall model prediction (gammatone spectrogram: 59 clusters; t_{max} = 6.309, p < .001; gammatone spectrogram onsets: 61 clusters; t_{max} = 10.218, p < .001; Figure 10A). A two-way repeated measures ANOVA revealed a significant interaction of acoustic predictor and listening condition (F(1,59)= 36.704, p < .001, η_p^2 = .384). Post hoc analysis showed that prediction accuracy for gammatone spectrogram onsets (M = .005, SD = .003) was significantly greater than gammatone spectrogram (M = .001, SD = .001) in the quiet condition (p < .001). Additionally, prediction accuracy for gammatone spectrogram onsets significantly decreased in SSN (p < .001). However, prediction accuracy of gammatone spectrogram (M = .001, SD = .001) and gammatone spectrogram onsets (M = .001, SD = .001) did not differ within the SSN listening condition (p = .827).

There were also significant effects of acoustic predictor (F(1,59)= 47.408, p < .001, η_p^2 = .446) and listening condition (F(1,59)= 73.097, p < .001, η_p^2 = .553). These significant effects indicate that prediction accuracy was overall greater for gammatone spectrogram onsets compared to gammatone spectrogram, regardless of listening condition (gammatone spectrogram onsets: M = .003, SD = .003; gammatone spectrogram: M = .001, SD = .001). Moreover, the significant effect of listening condition indicates that overall neural tracking was higher in the quiet condition (M = .003, SD = .003) compared to the SSN condition (M = .001, SD = .001). Together, these significant findings indicate that prediction accuracy for acoustic information in the target speaker is driven by gammatone spectrogram onsets in quiet, but this effect is dampened in the presence of SSN (Figure 10B).

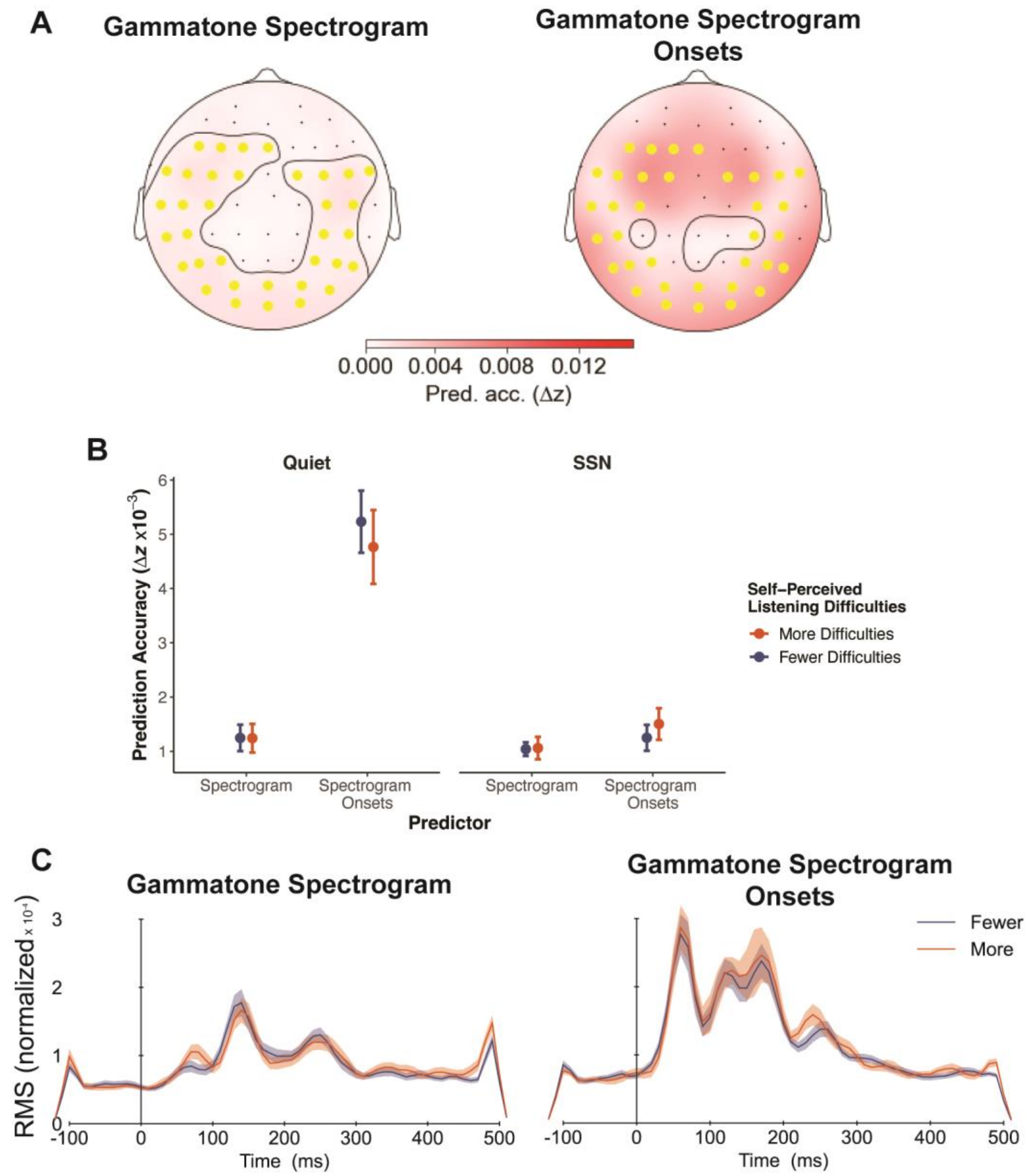

Figure 10: Neural Tracking of Acoustic Predictors. A) Increase in prediction accuracy due to gammatone spectrogram (left) and gammatone spectrogram onsets (right) was significantly above zero in clusters of electrodes highlighted in yellow. B) Prediction accuracy in quiet and speech shaped noise (SSN) listening condition for the gammatone spectrogram and gammatone spectrogram onsets. Prediction accuracy for gammatone spectrogram onsets significantly decreased in the SSN listening condition. However, prediction accuracy of each acoustic predictor did not differ based on self-perceived listening difficulties (SPLDs). Error bars denote standard error of the mean. C) Global field power of multivariate temporal response functions (mTRFs) by SPLDs. Shaded areas denote standard error of the mean.

In summary, neural tracking of acoustic information significantly dropped in the presence of background noise. However, listeners with more and fewer SPLDs did not differ in neural tracking of the acoustic information in continuous speech.

3.3.3 Neural tracking of linguistic information is not modulated by SPLDs

The linguistic predictors significantly contributed to overall model prediction, beyond the effect of acoustic predictors (55 clusters; $t_{max} = 6.404$, $p < .001$; Figure 11). This indicates that listeners tracked the linguistic information in the target speech. Age showed a significant Spearman's rank correlation with overall linguistic prediction accuracy ($\rho = -.330$, $p < .001$), which suggests that older participants had lower prediction accuracy of linguistic information. Thus, age was included as a covariate in the ANOVA. The results from the two-way repeated measures analysis of covariance (ANCOVA) revealed no significant main effect of SPLDs ($F(1,57) = 0.511$, $p = .478$, $\eta_p^2 = .009$) nor interaction of condition and SPLDs ($F(1,57) = 0.520$, $p = .474$, $\eta_p^2 = .009$). However, there was a marginal main effect of condition ($F(1,57) = 3.814$, $p = .056$, $\eta_p^2 = .063$). This marginally significant main effect suggests that there was a trend for greater prediction accuracy for linguistic information in quiet ($M = .002$, $SD = .003$), relative to the SSN listening condition ($M = .001$, $SD = .002$). A significant main effect of age was also observed ($F(1,57) = 9.600$, $p = .003$, $\eta_p^2 = .144$), which was expected given the significant correlation reported earlier. Collectively, these results suggest neural tracking of linguistic information overall decreases with age and is also reduced in the presence of background noise.

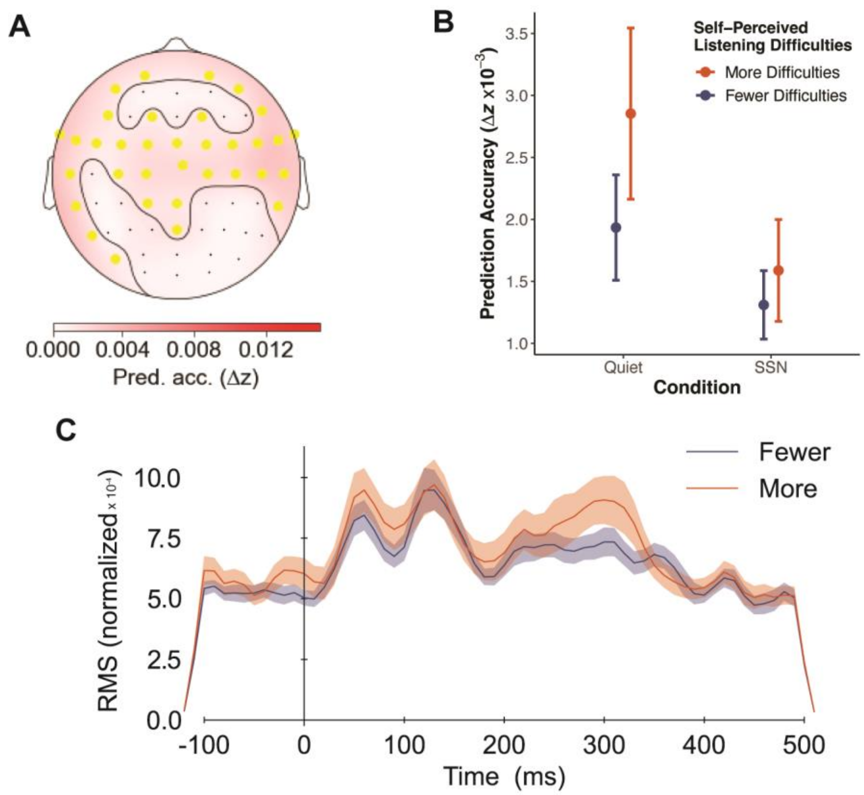

Figure 11: Neural Tracking of Linguistic Information. Prediction accuracy to linguistic information were significantly above zero in a cluster of electrodes highlighted in yellow. B) Average prediction accuracy in quiet and speech shaped noise (SSN) for listeners with more and fewer self-perceived listening difficulties (SPLDs) for linguistic predictors. Error bars denote standard error of the mean. Prediction accuracy did not differ by SPLDs nor listening condition. C) Global field power of multivariate temporal response functions (mTRFs) by SPLDs. Shaded areas denote standard error of the mean.

3.3.4 SPLDs affect neural tracking of sentence-level information

Lastly, I conducted an analysis to examine the extent to which prediction accuracy of contextual information (i.e., sublexical, word, sentence) differed by condition and SPLDs (Figure 12). Each level of contextual information significantly contributed to overall model prediction (sublexical: 50 clusters; t_{max} = 6.447, p < .001; word: 3 clusters; t_{max} = 3.617, p = .014; sentence: 8 clusters; t_{max} = 3.621, p = .013). Similar to the linguistic model, age showed a significant relationship with sublexical prediction accuracy (ρ = -.270, p = .003). First, a three-way repeated measures ANCOVA with a between-subjects independent variable of SPLD group, within-subjects independent variable of listening condition and linguistic predictor (sublexical, word-form, sentence-level), and a covariate of age was performed. However, the ANCOVA revealed no significant effect of age on prediction accuracy (F(1,57)= 2.666, p = .108, η_p^2 = .045). This indicates that age did not have a significant main effect on prediction accuracy at the sublexical, word, or sentence-level. Therefore, a two-way repeated measures ANOVA, without age, was performed. Results from the ANOVA revealed a significant main effect of predictor (F(2,116)= 7.267, p = .001, η_p^2 = .111). Post hoc analysis indicated that prediction accuracy was significantly higher at the sublexical level (M = .001, SD = .001) compared to the word level (M = .0004, SD = .001; p = .004). However, no significant main effects of SPLDs (F(1,58)= 1.194, p = .279, η_p^2 = .020) or condition (F(1,58)= 3.570, p = .064, η_p^2 = .058) were observed.

There was a significant interaction of predictor and SPLDs (F(2,116)= 3.758, p = .026, η_p^2 = .061, power = .954). Tukey’s post hoc analysis showed that listeners with more SPLDs had higher prediction accuracy for sentence-level information than listeners with fewer SPLDs (more SPLDs: M = .001, SD = .002; fewer SPLDs: M = .0003, SD = .001). No other significant interactions were present between SPLDs and condition (F(1,58)= 0.567, p = .454, η_p^2 = .010), predictor and condition ($F(2,116) = 2.182$, $p = .117$, $\eta_p^2 = .036$), or SPLDs, predictor, and condition ($F(2,116) = 0.015$, $p = .985$, $\eta_p^2 < .001$). Collectively, these results indicate that: 1) neural tracking of sublexical information was greater than word and sentence-level information, and 2) listeners with more SPLDs had enhanced neural tracking of sentence-level information compared to listeners with fewer SPLDs.

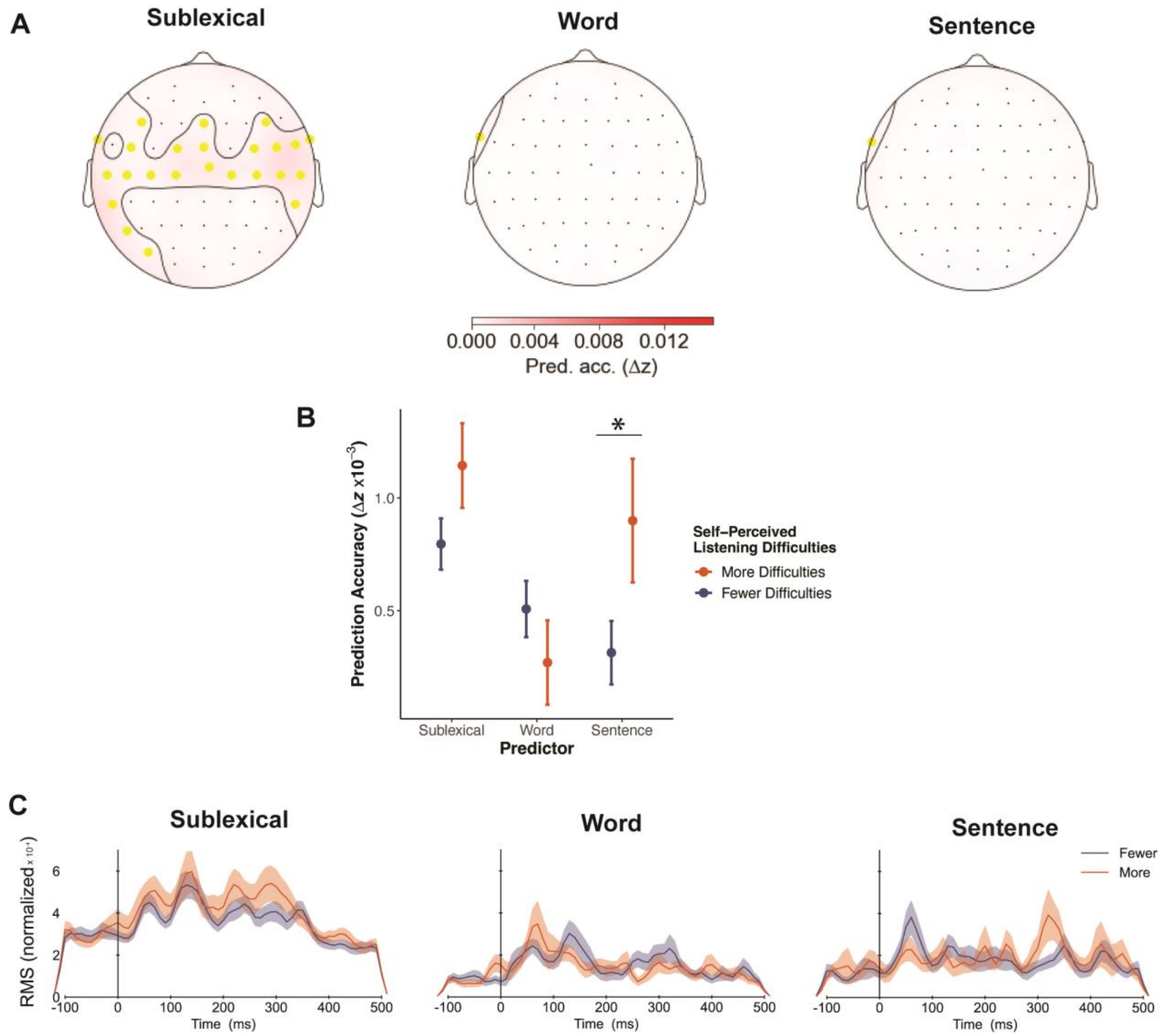

Figure 12: Neural Tracking of Linguistic Information Across Contextual Levels. A) Increase in prediction accuracy due to sublexical (left), word-form (middle), and sentence-level information (right) was significantly above zero in clusters of electrodes highlighted in yellow. B) Prediction accuracy in quiet and speech shaped noise (SSN) listening conditions for sublexical, word-form, and sentence-level information. Prediction accuracy for sentence-level information was significantly greater for listeners with more self-perceived listening difficulties (SPLDs) relative to those with fewer SPLDs (*p* = .026), regardless of the listening condition. Error bars denote standard error of the mean. C) Global field power of multivariate temporal response functions (mTRFs) by SPLDs. Shaded areas denote standard error of the mean. mTRFs did not differ based on SPLDs.

3.4 Discussion

This study investigated the extent to which neural tracking of acoustic and linguistic information in continuous speech differed based on SPLDs. I found that listeners with more SPLDs had enhanced neural tracking of sentence-level information than those with fewer SPLDs, regardless of the presence of background noise. Moreover, listeners with more SPLDs had comparable speech comprehension performance as those with fewer SPLDs. This suggests that listeners with more SPLDs may have increased reliance on sentence-level information as a compensatory mechanism to aid speech comprehension performance.

Neural tracking of acoustic information significantly reduced in the presence of noise (i.e., SSN), which was consistent with prior findings demonstrating reduced neural tracking of acoustics when speech is masked by noise (Brodbeck et al., 2020; McHaney et al., 2021). In both listening conditions, I specifically examined acoustics of the target speech, but also accounted for acoustics of the SSN masker alone, and the mixture of the SSN masker and target speech. This was done because prior research has demonstrated that acoustics of the masker and mixture are encoded in the auditory system even when they are unattended (Brodbeck et al., 2020). However, listeners with more SPLDs had comparable neural tracking of acoustics of the target speech in both quiet and SSN as listeners with fewer SPLDs. Moreover, neural tracking of individual acoustic predictors (i.e., gammatone spectrogram and gammatone spectrogram onsets) was similar in quiet and SSN across SPLDs. This indicated that neural tracking of acoustic information in speech is not a contributing factor to SPLDs.

Contrary to my prediction, neural tracking of linguistic information was only marginally reduced in SSN. Selective attention is necessary to focus on and attend to the target speaker when speech is masked in noise (Shinn-Cunningham, 2008; Shinn-Cunningham & Best, 2008). Prior research has demonstrated a reduction in neural tracking of linguistic information when attentional demands are high (Xie et al., 2023). Thus, I predicted that there would be a significant reduction in neural tracking of linguistic information in the SSN listening condition. However, several factors in the current study may have contributed to the lack of a significant reduction in neural tracking of linguistic information. First, participants in the study included both younger and middle-aged adults and participant age was a significant covariate to neural tracking of linguistic information. This indicated that neural tracking of linguistic information in continuous speech decreased with age. Aging effects on neural tracking were not the focus of the current study, but additional research is needed to understand the extent to which age impacts neural tracking of linguistic information in continuous speech. Furthermore, the use of SSN as the masker may have contributed to the insignificant reduction of neural tracking of linguistic information. SSN is a type of energetic masker, which primarily interferes with speech perception at the level of the peripheral auditory system (Brungart, 2001). Therefore, energetic masking does not require cognitive resources to overcome the detrimental effects of masking to the degree that informational masking does, such as with multi-talker background noise (Brungart, 2001; Kidd et al., 1998).

While SPLDs did not have an effect on overall neural tracking of overall linguistic information, SPLDs did have an effect on sublinguistic (sublexical, word-form, sentence-level) neural tracking. Specifically, those with more SPLDs had enhanced neural tracking of sentence-level contextual information compared to those with fewer SPLDs, regardless of the presence of noise. However, only one cluster of electrodes was identified as being significantly different from zero in the sentence model, so the findings should be interpreted with caution. Based on the findings from this study, neural tracking of sentence-level information during listening may play one of two roles for listeners with more SPLDs in speech comprehension, but first it is important to understand what neural tracking of sentence-level information means. The sentence model in the current study estimated the probability of the next phoneme based on the preceding four words, which means that the expectation of the next phoneme was modulated by broader sentence context. This suggests that listeners with more SPLDs utilize broader context clues from the combination of multiple preceding words to predict what the speaker will say next. Reliance on broader context likely requires greater working memory resources to hold onto larger linguistic units (i.e., words), while continuously updating predictions for upcoming sounds. Moreover, in the presence of background noise, working memory is also used to hold onto the perceived speech while filling in missing gaps in the speech stream caused by noise (Rönnberg et al., 2010). Thus, listeners with more SPLDs likely expend greater cognitive resources, both in quiet and in the presence of noise, to track sentence-level information when listening to continuous speech.

Speech comprehension performance was comparable across listening conditions for those with more and fewer SPLDs, which suggests that listeners with more SPLDs may leverage broader sentence-level context as a compensatory strategy to aid listening. However, comparable speech comprehension performance does not beget ease of listening. Diverting cognitive resources (e.g., working memory) to support listening makes listening more effortful (Pichora-Fuller et al., 2016). When listening effort is high, listeners report greater instances of fatigue and difficulties understanding speech (Pichora-Fuller et al., 2016). However, some evidence suggests that reliance on sentence-level context actually *reduces* listening effort (Hunter & Humes, 2022). Therefore, it remains unclear how neural tracking of sentence-level context impacts listening effort in listeners with SPLDs.

The findings from this study are similar to studies on aging, which found that older adults benefit from supportive context to predict upcoming words to a greater degree than younger adults

(Gordon-Salant & Fitzgibbons, 1997; Hunter & Humes, 2022; Milburn et al., 2021; Pichora-Fuller, 2008; Pichora-Fuller et al., 1995). It is presumed that older adults largely have more life experience dealing with difficult listening situations, therefore, they have adopted a sentence-level listening strategy to aid speech comprehension performance (Pichora-Fuller, 2008). In the context of the findings from the current study, however, it is unlikely that listeners with more SPLDs have implemented this sentence-level listening strategy due to more life experience given that this study focused on adults under the age of 52. Additionally, the prior studies in older adults included listeners with elevated hearing thresholds (Gordon-Salant & Fitzgibbons, 1997; Hunter & Humes, 2022; Pichora-Fuller et al., 1995), which could suggest that their reliance on sentence-level context was to compensate for degraded sensory input from hearing loss. However, all participants in the current study had clinically normal hearing thresholds. Thus, the listeners with more SPLDs likely did not have bottom-up sensory encoding deficiencies that led to using broader context to assist listening. Future research should focus on replicating the sentence-level findings from the current study to understand why listeners with SPLDs tend to rely on sentence-level contextual information and the extent to which it is driven by sensory encoding deficits.

3.4.1 Conclusion

Access to linguistic knowledge during listening is critical to speech comprehension performance (Gwilliams & Davis, 2022; Pichora-Fuller, 2008). This study provides novel insights into how individuals with SPLDs leverage linguistic information in continuous speech during listening. Specifically, listeners with more SPLDs rely on sentence-level contextual information to a greater degree than listeners with fewer SPLDs, which may serve as a listening strategy to aid speech comprehension performance. Sentence-level information leverages broader context to predict upcoming sounds, which may result in greater listening difficulties because more cognitive resources are likely used to hold onto the larger linguistic units (i.e., words, phrases). Future research should examine the extent to which listeners with SPLDs employ this listening strategy in response to deficiencies in bottom-up sensory encoding, or if the listening strategy directly contributes to their SPLDs.

4.0 General Discussion

Understanding speech in noisy listening environments becomes increasingly difficult with age (Helfer & Wilber, 1990; Peelle et al., 2010, 2011; Pichora-Fuller et al., 1995; Smayda et al., 2016). Factors such as age-related hearing loss (Helfer & Wilber, 1990; Liberman et al., 2016b; Peelle & Wingfield, 2016) and lifetime noise exposure (Hope et al., 2013; Le Prell, 2019; Skoe et al., 2019) can contribute to speech perception difficulties. However, individual differences in speech perception can persist even in listeners with normal hearing thresholds (Lam et al., 2017; Oberfeld & Klöckner-Nowotny, 2016; Parbery-Clark et al., 2011; Souza & Arehart, 2015; Tierney et al., 2019; Xie et al., 2014, 2015). One in ten clients who visit audiology clinics with complaints of listening difficulties have normal audiograms (Hind et al., 2011; Parthasarathy et al., 2020; Pryce & Wainwright, 2008; Spankovich et al., 2018; Tremblay et al., 2015). In the presence of normal hearing thresholds, it can be difficult to devise a helpful treatment plan for the client's SPLDs because the source driving their difficulties is unclear.

Speech processing involves a combination of bottom-up and top-down cognitive processes (Lam et al., 2017; Parthasarathy et al., 2020; Pichora-Fuller et al., 2016; Rönnberg et al., 2013). Prior research into SPLDs in the presence of normal hearing thresholds has focused on potential deficits in bottom-up auditory processing (Bharadwaj et al., 2015; Bharadwaj, Verhulst, et al., 2014; Bramhall et al., 2017; Liberman et al., 2016a; Paul et al., 2017; Stamper & Johnson, 2015a, 2015b; Valderrama et al., 2018; Verhulst et al., 2018). However, much less is known about the contributions of top-down processes into SPLDs. The goal of this dissertation was to examine the contributions of top-down decisional and linguistic processes that support speech perception in SPLDs in young and middle-aged listeners with normal hearing thresholds. Aim 1 of this dissertation study examined the extent to which the decisional processes that support speech in noise categorization differed in listeners with more SPLDs (Study 1). Aim 2 of this dissertation study investigated the extent to which SPLDs impacted the use of acoustic and linguistic information during continuous speech processing in quiet and noisy listening situations (Study 2).

Collectively, the results from these studies demonstrate that listeners with SPLDs have different approaches to listening. Study 1 revealed that listeners with more SPLDs were less efficient at extracting relevant information from a stimulus to make a categorization decision and received less benefit from decreasing background noise. Those with more SPLDs were also more impulsive responders, making riskier decisions that had a higher likelihood of being inaccurate. These findings have significant implications for understanding the neurobiology of top-down cognitive resources in SPLDs. Prior research indicates that decision thresholds are associated with pre-stimulus CON activity during speech perception in noise (Vaden et al., 2022). The link between decision thresholds and CON activity suggests that CON activity benefits speech perception in noise by raising decision thresholds for more cautious responding. In the context of the findings from Study 1, listeners with fewer SPLDs may have higher pre-stimulus CON activity which may allow them to make speech in noise categorization decisions more efficiently.

Additionally, Study 1 demonstrated that listeners with more SPLDs were more impulsive responders when categorizing phonemes in noise than those with fewer SPLDs, regardless of SNR level. Impulsive responding suggests that listeners with more SPLDs were prioritizing speed of categorization over accuracy of their response. This phenomenon is also known as the speed-accuracy tradeoff (Bogacz et al., 2010). At first, it may seem as if impulsive responding does not negatively affect speech in noise categorization given that overall accuracies on the task did not differ as a function of SPLDs. However, it’s important to consider the combination of evidence accumulation rates and decision thresholds when interpreting these findings. The lower decision threshold suggests that listeners with more SPLDs required less information to make a decision. Although listeners with more SPLDs needed less information for categorization, were not very efficient at extracting critical information from the speech stimulus to reach that lower threshold. Listeners wither fewer SPLDs had higher decision thresholds and faster evidence accumulation rates for easier SNR levels than those with more SPLDs. This indicates that listeners with fewer SPLDs were able to collect enough information to reach their higher threshold at a faster rate than listeners with more SPLDs did to reach their lower threshold. Overall, the findings from Study 1 demonstrate that listeners with more SPLDs have a deficit in the decisional processes that support speech in noise categorization.

Study 2 demonstrated that listeners with more SPLDs had enhanced neural tracking of sentence-level linguistic information in continuous speech compared to listeners with fewer SPLDs. Neural tracking of acoustic information in continuous speech did not differ as a function of SPLDs. Additionally, speech comprehension performance did not differ in quiet or in the SSN listening conditions based on SPLDs. These findings suggests that listeners with more SPLDs may leverage global sentence-level context as a strategy to aid speech comprehension. However, this 'strategy' may have cognitive consequences for the listener. Neural tracking of sentence-level information means that the listener must hold onto larger linguistic units (i.e., multiple words) in working memory for speech comprehension. Actively holding onto words in memory to update predictions for upcoming speech sounds would leave fewer cognitive resources available for other tasks.

Furthermore, increased representations of sentence-level linguistic information were evident in listeners with more SPLDs, regardless of the listening condition. This suggests that listeners with more SPLDs may leverage global context in all types of listening scenarios, which may not be beneficial. In a multi-talker listening scenario, there would be a higher probability for speech from a background talker to interfere with the words that are being held in working memory. Thus, the global context that is being used to update predictions for upcoming speech sounds would have a higher likelihood of being contaminated from background talkers. Thus, leveraging global context in the form of sentence-level information is likely a poor listening 'strategy' and could very well contribute to SPLDs.

When combining the data and findings from both studies, they provide evidence to suggest that individuals with more SPLDs process speech differently than those with fewer SPLDs. Listeners with more SPLDs received less benefit from decreasing background noise levels to extract critical information from an auditory stimulus for decision-making relative to listeners with fewer SPLDs. Additionally, listeners with more SPLDs may rely on sentence-level linguistic context during listening to aid speech comprehension. However, it remains unclear whether a causal relationship exists between evidence accumulation rates and enhanced neural tracking of sentence-level linguistic information. Do listeners upregulate neural tracking of sentence-level information because of poor sensory extraction of relevant information? Or does the cognitive load of tracking sentence-level information cause listeners to have poor sensory evidence accumulation? Future research should investigate the extent to which a shared, underlying mechanism exists between evidence accumulation rates and neural tracking of sentence-level information.

Theoretically, overlapping mechanisms between evidence accumulation rate and neural tracking of sentence-level information are possible. Extant research has demonstrated that evidence accumulation rate (Gold & Shadlen, 2002; Heekeren et al., 2004; Rolls et al., 2010;

Wang, 2002) and cognitive processes, such as working memory (D'Esposito, 2007), are associated with activity in the prefrontal cortex. If enhanced neural tracking of sentence-level information is in fact more taxing on cognitive resources, as posited in Study 2, then neural tracking of sentence-level information could be associated with increased prefrontal cortex activity during speech perception as well. The EEG data from Study 2 cannot be used to examine precise neural sources of sentence-level neural tracking due to poor source localization from the EEG (Michel & Murray, 2012). Prior research using functional magnetic resonance imaging has demonstrated that extra-auditory regions in the prefrontal cortex are recruited to assist with listening in challenging environments (Du et al., 2016; Peelle & Wingfield, 2016; Shinn-Cunningham, 2017; Wong et al., 2009). Thus, it is plausible that neural tracking of sentence-level information is associated with recruitment of similar frontal regions to aid listening. Carefully designed neuroimaging studies that assess neural tracking of linguistic information in continuous speech are needed to understand the extent to which neural tracking of linguistic information at different contextual levels and evidence accumulation share a common mechanism.

4.1 Limitations

While careful consideration was given to the designs of both studies and analyses in this dissertation, there are a few limitations to acknowledge. First, the range of SSQ scores across both studies were 5.2 to 9.9, where higher scores reflect fewer SPLDs, even though sample sizes in both studies were large (Study 1: n = 77; Study 2: n = 63). As such, it is unclear how evidence accumulation rates and neural tracking of sentence-level information may change with more extreme SPLDs (i.e., SSQ scores < 5). Extreme SPLDs may not have been observed because both studies in this dissertation focused on adults with normal hearing thresholds. A prior study in younger and middle-aged adults used SSQ scores below 7.25 as a cutoff for overall speech, spatial, and hearing difficulties, even though some of the adults had measurable hearing impairment (Demeester et al., 2012). In addition to this, the researchers that developed the SSQ observed a mean SSQ score of 5.6 in older adults with hearing impairments (Gatehouse & Noble, 2004). Therefore, the SSQ scores across both studies in this dissertation closely align with scores observed in prior studies and are likely accurate reflections of SPLDs in adults with normal hearing thresholds.

A second limitation of this dissertation surrounds the use of pure tone hearing thresholds as the sole exclusionary criteria for hearing acuity in both studies. Pure tone thresholds provide a quick and accurate measure of hearing sensitivity but do not provide a detailed view of hearing health. For instance, hearing thresholds can be within normal range, while other aspects of the auditory system show signs of impairments (Mepani et al., 2020; Mokrian et al., 2014). Therefore, it begs the question as to whether the participants in this dissertation may have exhibited other signs of hearing impairments that were not captured by the standard audiogram. A potential avenue for future research could examine the extent to which listeners with more SPLDs may have specific audiometric profiles across a battery of audiological tests. An in-depth audiological profile may help to explain the mechanisms underlying the decisional processes that support speech in noise categorization and enhanced representations of sentence-level information during listening.

Finally, the significant result from Study 2 demonstrating increased representations of sentence-level information in listeners with more SPLDs was based off of a single significant cluster of electrodes. The mTRF results in Study 2 identified eight clusters of electrodes that showed similar patterns of activity that contributed to the overall neural tracking of both acoustic and linguistic information. However, only one cluster out of the eight was identified as significantly contributing to overall neural tracking. As such, the results from Study 2 should be interpreted with caution. Although the finding was based off of a single cluster, the results are promising and encourage further investigation into individual differences in the use of local and global context to aid speech comprehension in adults with SPLDs.

4.2 Conclusions

The collective findings from this dissertation study contribute to our understanding of SPLDs in listeners with normal hearing thresholds. Study 1 provides novel insights into perceptual decision-making in adults with SPLDs and normal-hearing thresholds. Listeners with more SPLDs were less efficient evidence accumulators and more impulsive responders in noisy listening situations. Moreover, the findings from Study 2 provide neurophysiological evidence to indicate that listeners with more SPLDs may leverage global linguistic context as a potential listening strategy to aid speech comprehension performance. However, this strategy may be more taxing on cognitive resources, which may underlie an aspect of SPLDs. In summary, the findings from this dissertation study encourage further investigation into top-down cognitive processes in adults with normal hearing thresholds and SPLDs.

Bibliography

Aiken, S. J., & Picton, T. W. (2008). Human cortical responses to the speech envelope. *Ear and Hearing*, *29*(2), 139–157. https://doi.org/10.1097/AUD.0b013e31816453dc

American Speech-Language-Hearing Association. (n.d.). *Central Auditory Processing Disorder*. Practice Portal. Retrieved April 1, 2023, from https://www.asha.org/practice-portal/clinical-topics/central-auditory-processing-disorder/

American Speech-Language-Hearing Association. (1987). *Determining threshold level for speech*.

Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural Timing Is Linked to Speech Perception in Noise. *Journal of Neuroscience*, *30*(14), 4922–4926. https://doi.org/10.1523/JNEUROSCI.0107-10.2010

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Bainbridge, K. E., & Wallhagen, M. I. (2014). Hearing Loss in an Aging American Population: Extent, Impact, and Management. *Annual Review of Public Health*, *35*(1), 139–152. https://doi.org/10.1146/annurev-publhealth-032013-182510

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.1017/S0033291714001470

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/J.2517-6161.1995.TB02031.X

Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., & Liberman, A. M. (2001). Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain and Language*, *78*(3), 364–396. https://doi.org/10.1006/brln.2001.2484

Bharadwaj, H. M., Lee, A. K. C., & Shinn-Cunningham, B. G. (2014). Measuring auditory selective attention using frequency tagging. *Frontiers in Integrative Neuroscience*, *8*, 6. https://doi.org/10.3389/FNINT.2014.00006/BIBTEX

Bharadwaj, H. M., Masud, S., Mehraei, X. G., Verhulst, S., Barbara, X., & Shinn-Cunningham, G. (2015). Individual Differences Reveal Correlates of Hidden Hearing Deficits. *The Journal of Neuroscience*, *35*(5), 2161–2172. https://doi.org/10.1523/JNEUROSCI.3915-14.2015

Bharadwaj, H. M., Verhulst, S., Shaheen, L., Charles Liberman, M., & Shinn-Cunningham, B. G. (2014). Cochlear neuropathy and the coding of supra-threshold sound. *Frontiers in Systems Neuroscience*, *8*(FEB), 26. https://doi.org/10.3389/FNSYS.2014.00026/BIBTEX

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, *7*(3), 295–300. https://doi.org/10.1038/nn1198

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, *137*(1), 378. https://doi.org/10.1121/1.4904537

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends in Neurosciences*, *33*(1), 10–16. https://doi.org/10.1016/j.tins.2009.09.002

Bonte, M., Parviainen, T., Hytönen, K., & Salmelin, R. (2006). Time course of top-down and bottom-up influences on syllable processing in the auditory cortex. *Cerebral Cortex*, *16*(1), 115–123. https://doi.org/10.1093/cercor/bhi091

Bramhall, N. F., Beach, E. F., Epp, B., Le Prell, C. G., Lopez-Poveda, E. A., Plack, C. J., Schaette, R., Verhulst, S., & Canlon, B. (2019). The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing Research*, *377*, 88–103. https://doi.org/10.1016/j.heares.2019.02.016

Bramhall, N. F., Konrad-Martin, D., McMillan, G. P., & Griest, S. E. (2017). Auditory Brainstem Response Altered in Humans With Noise Exposure Despite Normal Outer Hair Cell Function. *Ear and Hearing*, *38*(1), e1. https://doi.org/10.1097/AUD.0000000000000370

Brodbeck, C., Bhattasali, S., Cruz Heredia, A. A. L., Resnik, P., Simon, J. Z., & Lau, E. (2022). Parallel processing in speech perception with local and global representations of linguistic context. *ELife*, *11*, 1–28. https://doi.org/10.7554/eLife.72056

Brodbeck, C., Das, P., Kulasingham, J. P., Bhattasali, S., Gaston, P., Resnik, P., & Simon, J. Z. (2021). Eelbrain: A Python toolkit for time-continuous analysis with temporal response functions. *BioRxiv*. https://doi.org/10.1101/2021.08.01.454687

Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Current Biology*, *28*(24), 3976–3983. https://doi.org/10.1016/J.CUB.2018.10.042

Brodbeck, C., Jiao, A., Hong, L. E., & Simon, J. Z. (2020). Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. *PLoS Biology*, *18*(10), e3000883. https://doi.org/10.1371/journal.pbio.3000883

Brodbeck, C., Presacco, A., Anderson, S., & Simon, J. Z. (2018). Over-Representation of Speech in Older Adults Originates from Early Response in Higher Order Auditory Cortex. *Acta Acustica United with Acustica*, *104*(5), 774–777. https://doi.org/10.3813/aaa.919221

Brodbeck, C., & Simon, J. Z. (2020). Continuous speech processing. *Current Opinion in Physiology, 18*, 25–31. https://doi.org/10.1016/j.cophys.2020.07.014

Broderick, M. P., Anderson, A. J., di Liberto, G. M., Crosse, M. J., & Lalor, E. C. (2018). Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. *Current Biology*, *28*(5), 803–809. https://doi.org/10.1016/j.cub.2018.01.080

Brody, C. D., & Hanks, T. D. (2016). Neural underpinnings of the evidence accumulator. *Current Opinion in Neurobiology*, *37*, 149–157. https://doi.org/10.1016/J.CONB.2016.01.003

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *109*(3), 1101–1109. https://doi.org/10.1121/1.1345696

Cancel, V. E., McHaney, J. R., Milne, V., Palmer, C., & Parthasarathy, A. (in review). *Speech in Noise Difficulties with Normal Audiograms: Insights from the Auditory Processing Disorder Battery*.

Carhart, R., & Jerger, J. F. (1959). Preferred Method For Clinical Determination Of Pure-Tone Thresholds. *Journal of Speech and Hearing Disorders*, *24*(4), 330–345. https://doi.org/10.1044/jshd.2404.330

Carroll, L. (1865). *Alice's Adventures in Wonderland*. https://librivox.org/alices-adventures-in-wonderland-by-lewis-carroll-5

Chandrasekaran, B., Sampath, P. D., & Wong, P. C. M. (2010). Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, *128*(1), 456–465. https://doi.org/10.1121/1.3445785

Chandrasekaran, B., van Engen, K., Xie, Z., Beevers, C. G., & Maddox, W. T. (2015). Influence of depressive symptoms on speech perception in adverse listening conditions. *Cognition and Emotion*, *29*(5), 900–909. https://doi.org/10.1080/02699931.2014.944106

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

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Frontiers in Human Neuroscience*, *10*(November), 1–14. https://doi.org/10.3389/fnhum.2016.00604

Decruy, L., Vanthornhout, J., & Francart, T. (2019). Evidence for enhanced neural tracking of the speech envelope un- derlying age-related speech-in-noise difficulties. *Journal of Neurophysiology*, *122*, 601–615.

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

Demeester, K., Topsakal, V., Hendrickx, J. J., Fransen, E., van Laer, L., van Camp, G., van de Heyning, P., & van Wieringen, A. (2012). Hearing disability measured by the speech, spatial, and qualities of hearing scale in clinically normal-hearing and hearing-impaired middle-aged persons, and disability screening by means of a reduced SSQ (the SSQ5). *Ear and Hearing*, *33*(5), 615–626. https://doi.org/10.1097/AUD.0b013e31824e0ba7

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 761–772. https://doi.org/10.1098/RSTB.2007.2086

Di Liberto, G. M., Crosse, M. J., & Lalor, E. C. (2018). Cortical Measures of Phoneme-Level Speech Encoding Correlate with the Perceived Clarity of Natural Speech. *Eneuro*, *5*(2). https://doi.org/10.1523/eneuro.0084-18.2018

Di Liberto, G. M., O'Sullivan, J. A., & Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Current Biology*, *25*(19), 2457–2465. https://doi.org/10.1016/j.cub.2015.08.030

Ding, N., Chatterjee, M., & Simon, J. Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage*, *88*, 41–46. https://doi.org/10.1016/j.neuroimage.2013.10.054

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158. https://doi.org/10.1038/nn.4186

Ding, N., & Simon, J. Z. (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, *107*(1), 78–89. https://doi.org/10.1152/jn.00297.2011

Ding, N., & Simon, J. Z. (2013). Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. *Journal of Neuroscience*, *33*(13), 5728–5735. https://doi.org/10.1523/JNEUROSCI.5297-12.2013

Dinino, M., Holt, L. L., & Shinn-Cunningham, B. G. (2021). Cutting Through the Noise: Noise-Induced Cochlear Synaptopathy and Individual Differences in Speech Understanding Among Listeners With Normal Audiograms. *Ear and Hearing*. https://doi.org/10.1097/AUD.0000000000001147

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature Communications*, *7*, 1–12. https://doi.org/10.1038/ncomms12241

Etard, O., & Reichenbach, T. (2019). Neural speech tracking in the theta and in the delta frequency band differentially encode clarity and comprehension of speech in noise. *Journal of Neuroscience*, *39*(29), 5750–5759. https://doi.org/10.14469/hpc/2232

Fishbach, A., Nelken, I., & Yeshurun, Y. (2001). Auditory Edge Detection: A Neural Model for Physiological and Psychoacoustical Responses to Amplitude Transients. *Journal of Neurophysiology*, *85*(6), 2303–2323. www.jn.org

Fulbright, A. N. C., Le Prell, C. G., Griffiths, S. K., & Lobarinas, E. (2017). Effects of Recreational Noise on Threshold and Suprathreshold Measures of Auditory Function. *Seminars in Hearing*, *38*(4), 298–318. https://doi.org/10.1055/S-0037-1606325/ID/JR00741-47

Füllgrabe, C., & Rosen, S. (2016). Investigating the role of working memory in speech-in-noise identification for listeners with normal hearing. In *Physiology, psychoacoustics and cognition in normal and impaired hearing* (pp. 29–36). Springer, Cham.

Furman, A. C., Kujawa, S. G., & Charles Liberman, M. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of Neurophysiology*, *110*(3), 577–586. https://doi.org/10.1152/JN.00164.2013

Gatehouse, S., & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of Audiology*, *24*(1), 63–68. https://doi.org/10.3109/03005369009077843

Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, *43*(2), 85–99. https://doi.org/10.1080/14992020400050014

Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the Brain. *Neuron*, *36*(2), 299–308. https://doi.org/10.1016/S0896-6273(02)00971-6

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. In *Annual Review of Neuroscience* (Vol. 30, pp. 535–574). https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gordon-Salant, S., & Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance. *Journal of Speech, Language, and Hearing Research*, *40*(2), 423–431. https://doi.org/10.1044/jslhr.4002.423

Grant, K. J., Parthasarathy, A., Vasilkov, V., Caswell-Midwinter, B., Freitas, M. E., De Gruttola, V., Polley, D. B., Liberman, M. C., & Maison, S. F. (2022). Predicting neural deficits in sensorineural hearing loss from word recognition scores. *Scientific Reports*, *12*, 8929. https://doi.org/10.1038/s41598-022-13023-5

Grinn, S. K., Wiseman, K. B., Baker, J. A., & Le Prell, C. G. (2017). Hidden hearing loss? No effect of common recreational noise exposure on cochlear nerve response amplitude in humans. *Frontiers in Neuroscience*, *11*, 465. https://doi.org/10.3389/FNINS.2017.00465/BIBTEX

Guest, H., Munro, K. J., Prendergast, G., Howe, S., & Plack, C. J. (2017). Tinnitus with a normal audiogram: Relation to noise exposure but no evidence for cochlear synaptopathy. *Hearing Research*, *344*, 265–274. https://doi.org/10.1016/J.HEARES.2016.12.002

Gwilliams, L., & Davis, M. H. (2022). Extracting language content from speech sounds: the information theoretic approach. In *Speech Perception* (pp. 113–139). Springer.

Hamilton, L. S., Edwards, E., & Chang, E. F. (2018). A Spatial Map of Onset and Sustained Responses to Speech in the Human Superior Temporal Gyrus. *Current Biology*, *28*(12), 1860-1871.e4. https://doi.org/10.1016/J.CUB.2018.04.033

Hamilton, L. S., & Huth, A. G. (2018). The revolution will not be controlled: natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience*, *35*(5), 573–582. https://doi.org/10.1080/23273798.2018.1499946

Heafield, K. (2011). KenLM: Faster and Smaller Language Model Queries. *Proceedings of the 6th Workshop on Statistical Machine Translation*, 187–197.

Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature*, *431*, 859–862. https://doi.org/10.1038/nature02966

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, *9*(6), 467–479. https://doi.org/10.1038/nrn2374

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

Hickok, G., & Poeppel, D. (2007). The cortical organizing of speech processing. *Nature Reviews Neuroscience*, *8*(May), 393–403. https://doi.org/10.1038/nrn2113

Hill, K. T., & Miller, L. M. (2010). Auditory Attentional Control and Selection during Cocktail Party Listening. *Cerebral Cortex*, *20*(3), 583–590. https://doi.org/10.1093/CERCOR/BHP124

Hind, S. E., Haines-Bazrafshan, R., Benton, C. L., Brassington, W., Towle, B., & Moore, D. R. (2011). Prevalence of clinical referrals having hearing thresholds within normal limits. *International Journal of Audiology*, *50*(10), 708–716. https://doi.org/10.3109/14992027.2011.582049

Holt, L. L., Tierney, A. T., Guerra, G., Laffere, A., & Dick, F. (2018). Dimension-selective attention as a possible driver of dynamic, context-dependent re-weighting in speech processing. *Hearing Research*, *366*, 50–64. https://doi.org/10.1016/J.HEARES.2018.06.014

Hope, A. J., Luxon, L. M., & Bamiou, D.-E. (2013). Effects of chronic noise exposure on speech-in-noise perception in the presence of normal audiometry. *The Journal of Laryngology and Otology*, *127*, 233–238. https://doi.org/10.1017/S002221511200299X

Horton, C., D'Zmura, M., & Srinivasan, R. (2013). Suppression of competing speech through entrainment of cortical oscillations. *Journal of Neurophysiology*, *109*(12), 3082–3093. https://doi.org/10.1152/jn.01026.2012

Huddle, M. G., Goman, A. M., Kernizan, F. C., Foley, D. M., Price, C., Frick, K. D., & Lin, F. R. (2017). The Economic Impact of Adult Hearing Loss. *JAMA Otolaryngology–Head & Neck Surgery*, *143*(10), 1040. https://doi.org/10.1001/jamaoto.2017.1243

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

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453. https://doi.org/10.1016/j.physbeh.2017.03.040

Jerger, J., & Musiek, F. (2000). Report of the Consensus Conference on the Diagnosis of Auditory Processing Disorders in School-Aged Children. *Journal of the American Academy of Audiology*, *11*, 467–474.

Johnson, K. L., Nicol, T., Zecker, S. G., Bradlow, A. R., Skoe, E., & Kraus, N. (2008). Brainstem encoding of voiced consonant-vowel stop syllables. *Clinical Neurophysiology*, *119*(11), 2623–2635. https://doi.org/10.1016/j.clinph.2008.07.277

Kassambara, A. (2021). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*.

Katz, J. (2015). *Handbook of Clinical Audiology* (M. Chasin, K. M. English, L. J. Hood, & K. L. Tillery, Eds.; 7th ed.). Wolters Kluer.

Kayser, A. S., Erickson, D. T., Buchsbaum, B. R., & D'Esposito, M. (2010). Neural Representations of Relevant and Irrelevant Features in Perceptual Decision Making. *The Journal of Neuroscience*, *30*(47), 15778–15789. https://doi.org/10.1523/JNEUROSCI.3163-10.2010

Keuleers, E., & Brysbaert, M. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. https://doi.org/10.3758/BRM.42.3.643

Kidd, G., Mason, C. R., & Rohtla, T. L. (1998). Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *The Journal of the Acoustical Society of America*, *104*, 1101. https://doi.org/10.1121/1.423246

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *116*, 2395. https://doi.org/10.1121/1.1784440

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *The Journal of the Acoustical Society of America*, *67*(3), 971–995. https://doi.org/10.1121/1.383940

Klem, G., Luders, H., Jasper, H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *52*(3), 3–6. https://doi.org/10.1016/0013-4694(58)90053-1

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26.

Lalor, E. C., & Foxe, J. J. (2010). Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *European Journal of Neuroscience*, *31*(1), 189–193. https://doi.org/10.1111/j.1460-9568.2009.07055.x

Lam, B. P. W., Xie, Z., Tessmer, R., & Chandrasekaran, B. (2017). The Downside of Greater Lexical Influences: Selectively Poorer Speech Perception in Noise. *Journal of Speech, Language, and Hearing Research*, *600*, 1662–1673. https://doi.org/10.1044/2017_JSLHR-H-16-0133

Le Prell, C. G. (2019). Effects of noise exposure on auditory brainstem response and speech-in-noise tasks: a review of the literature. *International Journal of Audiology*, *58*(S1), S3–S32. https://doi.org/10.1080/14992027.2018.1534010

Lee, A. K. C., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M. S., & Shinn-Cunningham, B. G. (2012). Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. *Frontiers in Neuroscience*, *6*, 190. https://doi.org/10.3389/FNINS.2012.00190/BIBTEX

Lenth, R. V. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means*.

Lesenfants, D., Vanthornhout, J., Verschueren, E., Decruy, L., & Francart, T. (2019). Predicting individual speech intelligibility from the cortical tracking of acoustic- and phonetic-level speech representations. *Hearing Research*, *380*, 1–9. https://doi.org/10.1016/j.heares.2019.05.006

Lesenfants, D., Vanthornhout, J., Verschueren, E., & Francart, T. (2019). Data-driven spatial filtering for improved measurement of cortical tracking of multiple representations of speech. *Journal of Neural Engineering*, *16*(6), 066017. https://doi.org/10.1088/1741-2552/AB3C92

Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016a). Toward a Differential Diagnosis of Hidden Hearing Loss in Humans. *PLoS ONE*, *11*(9), e0162726. https://doi.org/10.1371/journal.pone.0162726

Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016b). Toward a Differential Diagnosis of Hidden Hearing Loss in Humans. *PLOS ONE*, *11*(9), e0162726. https://doi.org/10.1371/JOURNAL.PONE.0162726

Lin, F. R., & Albert, M. (2014). Hearing loss and dementia - who is listening? *Aging and Mental Health*, *18*(6), 671–673. https://doi.org/10.1080/13607863.2014.915924

Madsen, S. M. K., Marschall, M., Dau, T., & Oxenham, A. J. (2019). Speech perception is similar for musicians and non-musicians across a wide range of conditions. *Scientific Reports*, *9*, 10404. https://doi.org/10.1038/s41598-019-46728-1

Madsen, S. M. K., Whiteford, K. L., & Oxenham, A. J. (2017). Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. *Scientific Reports* , *7*, 12624. https://doi.org/10.1038/s41598-017-12937-9

Mai, G., Tuomainen, J., & Howell, P. (2018). Relationship between speech-evoked neural responses and perception of speech in noise in older adults. *The Journal of the Acoustical Society of America*, *143*(3), 1333–1345. https://doi.org/10.1121/1.5024340

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/FIGURES/6

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/J.JNEUMETH.2007.03.024

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. *Interspeech*, 498–502. https://doi.org/10.21437/Interspeech.2017-1386

McHaney, J. R., Gnanateja, G. N., Smayda, K. E., Zinszer, B. D., & Chandrasekaran, B. (2021). Cortical Tracking of Speech in Delta Band Relates to Individual Differences in Speech in Noise Comprehension in Older Adults. *Ear and Hearing*, *42*(2), 343–354.

Meister, H., Rählmann, S., Lemke, U., & Besser, J. (2018). Verbal Response Times as a Potential Indicator of Cognitive Load During Conventional Speech Audiometry With Matrix Sentences. *Trends in Hearing*, *22*, 1–11. https://doi.org/10.1177/2331216518793255

Mepani, A. M., Kirk, S. A., Hancock, K. E., Bennett, K., de Gruttola, V., Liberman, M. C., & Maison, S. F. (2020). Middle-ear muscle reflex and word-recognition in “normal hearing” adults: evidence for cochlear synaptopathy? *Ear and Hearing*, *41*(1), 25–38. https://doi.org/10.1097/AUD.0000000000000804

Michel, C. M., & Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *NeuroImage*, *61*(2), 371–385. https://doi.org/10.1016/J.NEUROIMAGE.2011.12.039

Milburn, E., Dickey, M. W., Warren, T., & Hayes, R. (2021). Increased reliance on world knowledge during language comprehension in healthy aging: evidence from verb-argument prediction. *Aging, Neuropsychology, and Cognition*. https://doi.org/10.1080/13825585.2021.1962791

Moerel, M., de Martino, F., & Formisano, E. (2012). Processing of natural sounds in human auditory cortex: Tonotopy, spectral tuning, and relation to voice sensitivity. *Journal of Neuroscience*, *32*(41), 14205–14216. https://doi.org/10.1523/JNEUROSCI.1388-12.2012

Mokrian, H., Shaibanizadeh, A., Farahani, S., Jalaie, S., Mahdi, P., Amali, A., & Arian nahad, H. (2014). Evaluation of Distortion and Transient Evoked Otoacoustic Emission in Tinnitus Patients with Normal Hearing. *Iranian Journal of Otorhinolaryngology*, *26*(1), 19–24. /pmc/articles/PMC3915065/

Moore, B. C. J. (2008). Basic auditory processes involved in the analysis of speech sounds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 947–963. https://doi.org/10.1098/RSTB.2007.2152

Mulder, M. J., Van Maanen, L., & Forstmann, B. U. (2014). Perceptual Decision Neurosciences - A Model Based Review. *Neuroscience*, *277*, 872–884. https://doi.org/10.1016/j.neuroscience.2014.07.031

Mullen, T. R., Kothe, C. A. E., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., Jung, T. P., & Cauwenberghs, G. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Transactions on Biomedical Engineering*, *62*(11), 2553–2567. https://doi.org/10.1109/TBME.2015.2481482

Newman, C. W., Weinstein, B. E., Jacobson, G. P., & Hug, G. (1991). Test-retest reliability of the hearing handicap inventory for adults. *Ear and Hearing*, *12*(5), 355–357.

Newman, C. W., Weinstein, B. E., Jacobson, G. P., & Hug, G. A. (1990). The Hearing Handicap Inventory for Adults: Psychometric Adequacy and Audiometric Correlates. *Ear and Hearing*, *11*(6), 430–433.

Ng, E. H. N., & Rönnberg, J. (2020). Hearing aid experience and background noise affect the robust relationship between working memory and speech recognition in noise. *International Journal of Audiology*, *59*(3), 208–218. https://doi.org/10.1080/14992027.2019.1677951

Noble, W., & Gatehouse, S. (2004). Interaural asymmetry of hearing loss, Speech, Spatial and Qualities of Hearing Scale (SSQ) disabilities, and handicap. *International Journal of Audiology*, *43*(2), 100–114. https://doi.org/10.1080/14992020400050015

Noble, W., & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial, and Qualities of Hearing scale (SSQ). *International Journal of Audiology*, *45*(3), 172–181. https://doi.org/10.1080/14992020500376933

Noble, W., Jensen, N. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. *International Journal of Audiology*, *52*(6), 409–412. https://doi.org/10.3109/14992027.2013.781278

Noble, W., Tyler, R., Dunn, C., & Bhullar, N. (2008). Unilateral and bilateral cochlear implants and the implant-plus-hearing-aid profile: Comparing self-assessed and measured abilities. *International Journal of Audiology*, *47*(8), 505–514. https://doi.org/10.1080/14992020802070770

Oberfeld, D., & Klöckner-Nowotny, F. (2016). Individual differences in selective attention predict speech identification at a cocktail party. *ELife*, *5*, e16747. https://doi.org/10.7554/ELIFE.16747

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. In *Trends in Cognitive Sciences* (Vol. 23, Issue 11). https://doi.org/10.1016/j.tics.2019.08.004

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., & Lalor, E. C. (2014). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bht355

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-In-noise. *Ear and Hearing*, *30*(6), 653–661. https://doi.org/10.1097/AUD.0B013E3181B412E9

Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., & Kraus, N. (2011). Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *PLoS ONE*, *6*(5). https://doi.org/10.1371/journal.pone.0018082

Parthasarathy, A., Hancock, K. E., Bennett, K., Degruttola, V., & Polley, D. B. (2020). Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *ELife*, *9*, e51419. https://doi.org/10.7554/eLife.51419

Parthasarathy, A., & Kujawa, S. G. (2018). Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing. *Journal of Neuroscience*, *38*(32), 7108–7119. https://doi.org/10.1523/JNEUROSCI.3240-17.2018

Paul, B. T., Bruce, I. C., & Roberts, L. E. (2017). Evidence that hidden hearing loss underlies amplitude modulation encoding deficits in individuals with and without tinnitus. *Hearing Research*, *344*, 170–182. https://doi.org/10.1016/J.HEARES.2016.11.010

Paulon, G., Llanos, F., Chandrasekaran, B., & Sarkar, A. (2020). Bayesian Semiparametric Longitudinal Drift-Diffusion Mixed Models for Tone Learning in Adults. *Journal of the American Statistical Association*, 1–14. https://doi.org/10.1080/01621459.2020.1801448

Paulon, G., & Sarkar, A. (2021). *lddmm: Longitudinal Drift-Diffusion Mixed Models (LDDMM)*.

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing*, *39*(2), 204–214. https://doi.org/10.1097/AUD.0000000000000494

Peelle, J. E., Troiani, V., Grossman, M., & Wingfield, A. (2011). Hearing Loss in Older Adults Affects Neural Systems Supporting Speech Comprehension. *Journal of Neuroscience*, *31*(35), 12638–12643. https://doi.org/10.1523/JNEUROSCI.2559-11.2011

Peelle, J. E., Troiani, V., Wingfield, A., & Grossman, M. (2010). Neural processing during older adults' comprehension of spoken sentences: Age differences in resource allocation and connectivity. *Cerebral Cortex*, *20*(4), 773–782. https://doi.org/10.1093/cercor/bhp142

Peelle, J. E., & Wingfield, A. (2016). The Neural Consequences of Age-Related Hearing Loss. *Trends in Neurosciences*, *39*(7), 486–497. https://doi.org/10.1016/j.tins.2016.05.001

Petersen, E. B., Wöstmann, M., Obleser, J., & Lunner, T. (2016). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *Journal of Neurophysiology*, *117*(1). https://doi.org/10.1152/jn.00527.2016

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*(SUPPL. 2). https://doi.org/10.1080/14992020802307404

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37*, 5S-27S. https://doi.org/10.1097/AUD.0000000000000312

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

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002–1044. https://doi.org/10.1037/BUL0000158

Poeppel, D., Idsardi, W. J., & Van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B*, *363*, 1071–1086. https://doi.org/10.1098/rstb.2007.2160

Prendergast, G., Guest, H., Munro, K. J., Kluk, K., Léger, A., Hall, D. A., Heinz, M. G., & Plack, C. J. (2017). Effects of noise exposure on young adults with normal audiograms I: Electrophysiology. *Hearing Research*, *344*, 68–81. https://doi.org/10.1016/J.HEARES.2016.10.028

Prendergast, G., Tu, W., Guest, H., Millman, R. E., Kluk, K., Couth, S., Munro, K. J., & Plack, C. J. (2018). Supra-threshold auditory brainstem response amplitudes in humans: Test-retest reliability, electrode montage and noise exposure. *Hearing Research*, *364*, 38–47. https://doi.org/10.1016/J.HEARES.2018.04.002

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

Pryce, H., & Wainwright, D. (2008). Help-seeking for medically unexplained hearing difficulties: A qualitative study. *International Journal of Therapy and Rehabilitation*, *15*(8), 343–349. https://doi.org/10.12968/ijtr.2008.15.8.30818

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rappaport, J. M., Phillips, D. P., & Gulliver, J. M. (1993). Disturbed speech intelligibility in noise despite a normal audiogram: a defect in temporal resolution? *The Journal of Otolaryngology*, *22*(6), 447–453. https://europepmc.org/article/med/8158743

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.

Ratcliff, R., Smith, P. L., Brown, S. D., & Mckoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, *16*(2), 323–341. https://doi.org/10.1037/0882-7974.16.2.323

Ratcliff, R., Thapar, A., & McKoon, G. (2006a). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin and Review*, *13*(4), 626–635. https://doi.org/10.3758/BF03193973

Ratcliff, R., Thapar, A., & McKoon, G. (2006b). Aging, practice, and perceptual tasks: A diffusion model analysis. *Psychology and Aging*, *21*(2), 353–371. https://doi.org/10.1037/0882-7974.21.2.353

Reetzke, R., Gnanateja, G. N., & Chandrasekaran, B. (2021). Neural tracking of the speech envelope is differentially modulated by attention and language experience. *Brain and Language*, *213*, 104891. https://doi.org/10.1016/J.BANDL.2020.104891

Roark, C. L., Paulon, G., Sarkar, A., & Chandrasekaran, B. (2021). Comparing perceptual category learning across modalities in the same individuals. *Psychonomic Bulletin & Review*.

Roitman, J. D., & Shadlen, M. N. (2002). Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. *The Journal of Neuroscience*, *22*(21), 9475–9489.

Rolls, E. T., Grabenhorst, F., & Deco, G. (2010). Decision-Making, Errors, and Confidence in the Brain. *Journal of Neurophysiology*, *104*(5), 2359–2374. https://doi.org/10.1152/jn.00571.2010

Rönnberg, J., Lunner, T., Zekveld, A. A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*, 1–17. https://doi.org/10.3389/fnsys.2013.00031

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.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do You See What I Am Saying? Exploring Visual Enhancement of Speech Comprehension in Noisy Environments. *Cerebral Cortex*, *17*, 1147–1153. https://doi.org/10.1093/cercor/bhl024

Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of Musical Training on Understanding Voiced and Whispered Speech in Noise. *PLOS ONE*, *9*(1), e86980. https://doi.org/10.1371/JOURNAL.PONE.0086980

Saccone, P. A., & Steiger, J. R. (2007). Hearing handicap among adult residents of an urban homeless shelter. *Journal of Health Care for the Poor and Underserved*, *18*(1), 161–712.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*, 2400–2406. http://www.fil.ion.ucl.ac.uk/spm

Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, *26*(2), 100–107. https://doi.org/10.1016/S0166-2236(02)00037-1

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-Related Cochlear Synaptopathy: An Early-Onset Contributor to Auditory Functional Decline. *The Journal of Neuroscience*, *33*(34), 13686–13694. https://doi.org/10.1523/JNEUROSCI.1783-13.2013

Shaheen, L. A., Valero, M. D., & Liberman, M. C. (2015). Towards a Diagnosis of Cochlear Neuropathy with Envelope Following Responses. *Journal of the Association for Research in Otolaryngology*, *16*(6), 727–745. https://doi.org/10.1007/S10162-015-0539-3/FIGURES/12

Shamma, S. A. (1985). Speech processing in the auditory system I: The representation of speech sounds in the responses of the auditory nerve. *The Journal of the Acoustical Society of America*, *78*(5), 1612–1621. https://doi.org/10.1121/1.392799

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/J.1538-7305.1948.TB01338.X

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182–186. https://doi.org/10.1038/jid.2014.371

Shinn-Cunningham, B. G. (2017). Cortical and Sensory Causes of Individual Differences in Selective Attention Ability Among Listeners With Normal Hearing Thresholds. *Journal of Speech, Language, and Hearing Research*, *60*, 2976–2988. https://doi.org/10.1044/2017_JSLHR-H-17-0080

Shinn-Cunningham, B. G., & Best, V. (2008). Selective Attention in Normal and Impaired Hearing. *Trends in Amplification*, *12*(4), 283–299. https://doi.org/10.1177/1084713808325306

Singh, G., & Pichora-Fuller, M. K. (2010). Older adults' performance on the speech, spatial, and qualities of hearing scale (SSQ): Test-retest reliability and a comparison of interview and self-administration methods. *International Journal of Audiology*, *49*, 733–740. https://doi.org/10.3109/14992027.2010.491097

Skoe, E., Camera, S., & Tufts, J. (2019). Noise Exposure May Diminish the Musician Advantage for Perceiving Speech in Noise. *Ear & Hearing*, *40*(4), 782–793. https://doi.org/10.1097/AUD.0000000000000665

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, *291*, 244–252. https://doi.org/10.1016/j.bbr.2015.05.026

Smayda, K. E., van Engen, K. J., Maddox, W. T., & Chandrasekaran, B. (2016). Audio-visual and meaningful semantic context enhancements in older and younger adults. *PLoS ONE*, *11*(3), 1–14. https://doi.org/10.1371/journal.pone.0152773

Souza, P., & Arehart, K. (2015). Robust relationship between reading span and speech recognition in noise. *International Journal of Audiology*, *54*(10), 705–713. https://doi.org/10.3109/14992027.2015.1043062

Spankovich, C., Gonzalez, V. B., Su, D., & Bishop, C. E. (2018). Self reported hearing difficulty, tinnitus, and normal audiometric thresholds, the National Health and Nutrition Examination

Survey 1999–2002. *Hearing Research*, *358*, 30–36. https://doi.org/10.1016/j.heares.2017.12.001

Stamper, G. C., & Johnson, T. A. (2015a). Auditory function in normal-hearing, noise-exposed human ears. *Ear and Hearing, 36*(2), 172. https://doi.org/10.1097/AUD.0000000000000107

Stamper, G. C., & Johnson, T. A. (2015b). Letter to the Editor: Examination of potential sex influences in Stamper, G.C & Johnson, T.A. (2015). Auditory function in normal-hearing, noise-exposed human ears, Ear Hear, 36, 172-184. *Ear and Hearing*, *36*(6), 738. https://doi.org/10.1097/AUD.0000000000000228

Starr, A., Picton, T. W., Sininger, Y., Hood, L. J., & Berlin, C. I. (1996). Auditory neuropathy. *Brain*, *119*, 741–753. https://academic.oup.com/brain/article/119/3/741/396236

Summerfield, A. Q., Barton, G. R., Toner, J., Mcanallen, C., Proops, D., Harries, C., Cooper, H., Court, I., Gray, R., Osborne, J., Doran, M., Ramsden, R., Mawman, D., O'driscoll, M., Graham, J., Aleksy, W., Meerton, L., Verschure, C., Ashcroft, P., & Pringle, M. (2009). Self-reported benefits from successive bilateral cochlear implantation in post-lingually deafened adults: randomised controlled trial. *International Journal of Audiology*, *45*(1), 99–107. https://doi.org/10.1080/14992020600783079

Tierney, A., Rosen, S., & Dick, F. (2019). Speech-in-Speech Perception, Nonverbal Selective Attention, and Musical Training. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(5), 968. https://doi.org/10.1037/XLM0000767

Tremblay, K. L., Pinto, A., Fischer, M. E., Klein, B. E. K., Klein, R., Levy, S., Tweed, T. S., & Cruickshanks, K. J. (2015). Self-Reported Hearing Difficulties Among Adults With Normal Audiograms. *Ear & Hearing*, *36*(6), e290–e299. https://doi.org/10.1097/AUD.0000000000000195

Vaden, K. I., Kuchinsky, S. E., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2015). Cortical Activity Predicts Which Older Adults Recognize Speech in Noise and When. *The Journal of Neuroscience*, *35*(9), 3929–3937. https://doi.org/10.1523/JNEUROSCI.2908-14.2015

Vaden, K. I., Kuchinsky, S. E., Ahlstrom, J. B., Teubner-Rhodes, S. E., Dubno, J. R., & Eckert, M. A. (2016). Cingulo-Opercular Function during Word Recognition in Noise for Older Adults with Hearing Loss. *Experimental Aging Research*, *42*(1), 67–82. https://doi.org/10.1080/0361073X.2016.1108784

Vaden, K. I., Kuchinsky, S. E., Cute, S. L., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *Journal of Neuroscience*, *33*(48), 18979–18986. https://doi.org/10.1523/JNEUROSCI.1417-13.2013

Vaden, K. I., Teubner-Rhodes, S., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2022). Evidence for cortical adjustments to perceptual decision criteria during word recognition in noise. *NeuroImage*, *253*. https://doi.org/10.1016/J.NEUROIMAGE.2022.119042

Valderrama, J. T., Beach, E. F., Yeend, I., Sharma, M., Van Dun, B., & Dillon, H. (2018). Effects of lifetime noise exposure on the middle-age human auditory brainstem response, tinnitus and speech-in-noise intelligibility. *Hearing Research*, *365*, 36–48. https://doi.org/10.1016/J.HEARES.2018.06.003

Valero, M. D., Burton, J. A., Hauser, S. N., Hackett, T. A., Ramachandran, R., & Liberman, M. C. (2017). Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta). *Hearing Research*, *353*, 213–223. https://doi.org/10.1016/j.heares.2017.07.003

van Deun, L., van Wieringen, A., Scherf, F., Deggouj, N., Desloovere, C., Offeciers, F. E., van de Heyning, P. H., Dhooge, I. J., & Wouters, J. (2010). Earlier Intervention Leads to Better Sound Localization in Children with Bilateral Cochlear Implants. *Audiology Neurotology*, *15*, 7–17. https://doi.org/10.1159/000218358

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., & Francart, T. (2018). Speech Intelligibility Predicted from Neural Entrainment of the Speech Envelope. *Journal of the Association for Research in Otolaryngology*, *19*(2), 181–191. https://doi.org/10.1007/s10162-018-0654-z

Verhulst, S., Altoè, A., & Vasilkov, V. (2018). Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss. *Hearing Research*, *360*, 55–75. https://doi.org/10.1016/J.HEARES.2017.12.018

Wang, X.-J. (2002). Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. *Neuron*, *36*(5), 955–968. https://doi.org/10.1016/S0896-6273(02)01092-9

Welsh, L. W., Welsh, J. J., & Healy, M. P. (1985). Central presbycusis. *The Laryngoscope*, *95*(2), 128–136. https://doi.org/10.1288/00005537-198502000-00002

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wilson, R. H. (2003). Development of a Speech-in-Multitalker-Babble Paradigm to Assess Word-Recognition Performance. *Journal of the American Academy of Audiology*, *14*(9), 453–470.

Wong, P. C. M., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*, *47*(3), 693–703. https://doi.org/10.1016/j.neuropsychologia.2008.11.032

Wu, P. Z., Liberman, L. D., Bennett, K., de Gruttola, V., O'Malley, J. T., & Liberman, M. C. (2019). Primary Neural Degeneration in the Human Cochlea: Evidence for Hidden Hearing Loss in the Aging Ear. *Neuroscience*, *407*, 8–20. https://doi.org/10.1016/J.NEUROSCIENCE.2018.07.053

Xie, Z., Brodbeck, C., & Chandrasekaran, B. (2023). Cortical Tracking of Continuous Speech Under Bimodal Divided Attention. *Neurobiology of Language*, 1–26. https://doi.org/10.1162/nol_a_00100

Xie, Z., Maddox, W. T., Knopik, V. S., McGeary, J. E., & Chandrasekaran, B. (2015). Dopamine receptor D4 (DRD4) gene modulates the influence of informational masking on speech recognition. *Neuropsychologia*, *67*, 121–131. https://doi.org/10.1016/j.neuropsychologia.2014.12.013

Xie, Z., Reetzke, R., & Chandrasekaran, B. (2019). Machine learning approaches to analyze speech-evoked neurophysiological responses. *Journal of Speech, Language, and Hearing Research*, *62*(3), 587–601. https://doi.org/10.1044/2018_JSLHR-S-ASTM-18-0244

Xie, Z., Yi, H.-G., & Chandrasekaran, B. (2014). Nonnative Audiovisual Speech Perception in Noise: Dissociable Effects of the Speaker and Listener. *PLoS ONE*, *9*(12), e114439. https://doi.org/10.1371/journal.pone.0114439

Xie, Z., Zinszer, B. D., Riggs, M., Beevers, C. G., & Chandrasekaran, B. (2019). Impact of depression on speech perception in noise. *PLoS ONE*, *14*(8), e0220928. https://doi.org/10.1371/journal.pone.0220928

Yeend, I., Beach, E. F., Sharma, M., & Dillon, H. (2017). The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. *Hearing Research*, *353*, 224–236. https://doi.org/10.1016/j.heares.2017.07.006

Yi, H., Smiljanic, R., & Chandrasekaran, B. (2019). The Effect of Talker and Listener Depressive Symptoms on Speech Intelligibility. *Journal of Speech, Language, and Hearing Research*, 1–13.

Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., & Schoonhoven, R. (2006). Top-down and bottom-up processes in speech comprehension. *NeuroImage*, *32*(4), 1826–1836. https://doi.org/10.1016/j.neuroimage.2006.04.199

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing*, *32*(4), 498–510. https://doi.org/10.1097/AUD.0b013e31820512bb

Zhang, J., Tyler, R., Ji, H., Dunn, C., Wang, N., Hansen, M., & Gantz, B. (2015). Speech, Spatial and Qualities of Hearing Scale (SSQ) and Spatial Hearing Questionnaire (SHQ) Changes Over Time in Adults With Simultaneous Cochlear Implants. *American Journal of Audiology*, *24*(3), 384–397. https://doi.org/10.1044/2015_AJA-14-0074