ABSTRACT

Title of Dissertation: **EFFECTS OF AGE, HEARING LOSS AND COGNITION ON DISCOURSE COMPREHENSION AND SPEECH INTELLIGIBILITY PERFORMANCE**

Jaclyn Schurman
Doctor of Philosophy
2021

Dissertation directed by: Professor Sandra Gordon-Salant
Department of Hearing and Speech Sciences

Douglas Brungart, Ph.D.
Chief Scientist: Walter Reed National Military Medical Center

Discourse comprehension requires listeners to interpret the meaning of an incoming message, integrate the message into memory and use the information to respond appropriately. Discourse comprehension is a skill required to effectively communicate with others in real time. The overall goal of this research is to determine the relative impact of multiple environmental and individual factors on discourse comprehension performance for younger and older adults with and without hearing loss using a clinically feasible testing approach.

Study 1 focused on the impact of rapid speech on discourse comprehension performance for younger and older adults with and without hearing loss. Study 2 focused

on the impact of background noise and masker type on discourse comprehension performance for younger and older adults with and without hearing loss. The influences of cognitive function and speech intelligibility were also of interest. The impact of these factors was measured using a self-selection paradigm in both studies. Listeners were required to self-select a time-compression ratio or signal-to-noise ratio (SNR) where they could understand and effectively answer questions about the discourse comprehension passages. Results showed that comprehension accuracy performance was held relatively constant across groups and conditions, but the time-compression ratios and SNRs varied significantly.

Results in both studies demonstrated significant effects of age and hearing loss on the self-selection of listening rate and SNR. This result suggests that older adults are at a disadvantage for rapid speech and in the presence of background noise during a discourse comprehension task compared to younger adults. Older adults with hearing loss showed an additional disadvantage compared to older normal-hearing listeners for both difficult discourse comprehension tasks. Cognitive function, specifically processing speed and working memory, was shown to predict self-selected time-compression ratio and SNR. Understanding the effects of age, hearing loss and cognitive decline on discourse comprehension performance may eventually help mitigate these effects in real world listening situations.

EFFECTS OF AGE, HEARING LOSS AND COGNITION ON DISCOURSE COMPREHENSION AND SPEECH INTELLIGIBILITY PERFORMANCE

By

Jaclyn Schurman

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Sandra Gordon-Salant, Chair
Dr. Douglas Brungart
Professor Samira Anderson
Professor Matthew Goupell
Professor Catherine Carr, Dean’s Representative

Dedication

For my parents, my first teachers and most steadfast supporters.

Acknowledgements

Thank you to my advisor and teacher, Dr. Sandra Gordon-Salant, whose anatomy and physiology course changed the trajectory of my career. She has been instrumental in shaping my career path and I am grateful for her support and mentorship over the past 12 years. Thank you to my co-advisor, Dr. Douglas Brungart, whose confidence in me was a driving factor in my decision to pursue this degree. His guidance has been fundamental to the success of this project. Thank you to the other members of my committee, Dr. Samira Anderson, Dr. Matthew Goupell and Dr. Catherine Carr. I would also like to acknowledge the research staff at Walter Reed National Military Medical Center. Thank you to Dr. Kenneth Jensen, Hector Galloza and Michel Jackson, who spent countless hours helping me program these experiments. I would also like to acknowledge Dr. Stefanie Kuchinsky and Dr. Ken Grant, thank you for your thoughtful and important feedback throughout all the stages of this project.

Thank you to my fellow PhD students, especially Dr. Julie Cohen and Dr. Maureen Shader, whose friendship and feedback were essential to the completion of this degree. Most importantly, I would like to acknowledge my family. Thank you to my husband, Matt, whose patience and encouragement throughout this process have been invaluable. I would not have been able to complete this degree without his support. I would also like to acknowledge my son Benjamin, who is my greatest inspiration. Ben, you can do anything you set your mind to. Thank you to my parents and brother, our family unit has provided me with an endless amount of love and encouragement throughout this process.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

1T	1-Talker masker
AP	Anomalous Probability
ASIO	Audio Stream Input/output
BCC	Brown, Carlsen, Carstens listening comprehension test
DCT	Discourse Comprehension Test
ELU	Ease of Language Understanding Model
HINT	Hearing in Noise Test
HP	High Probability
JND	Just Noticeable Difference
LISN	Lectures, Interviews and Spoken Narrative
LMER	Linear Mixed-Effects Regression
LSPAN	Listening Span
MRT	Modified Rhyme Test
MoCA	Montreal Cognitive Assessment
NIH	National Institutes of Health
PSOLA	Pitch-Synchronous Overlap and Add
RAU	Rationalized Arcsine Units
RMS	Root-Mean-Square
RSPAN	Reading Span
RSPIN	Revised Speech Perception in Noise Test
SAT	Scholastic Aptitude Test
SNR	Signal-to-Noise Ratio
SSN	Speech-shaped noise
WIN	Words-in-Noise Test
YNH	Young Normal Hearing
ONH	Older Normal Hearing
OHI	Older Hearing Impaired
USB	Universal Serial Bus

Introduction

Discourse Comprehension

Listening comprehension for extended passages of speech, such as narratives, is referred to as discourse comprehension. Discourse comprehension is a complex and multifaceted process. This multifaceted process requires listeners to interpret the meaning of an incoming message, integrate the message into memory and use the information to respond appropriately (Schneider et al., 2010). Discourse comprehension is a skill required for verbal communication and is vital to everyday social interactions. The inability to comprehend a spoken message and communicate with others can lead to social isolation, loneliness and cognitive decline, particularly in older adults with hearing loss (Lin et al., 2011). Multiple individual and environmental factors can impact discourse comprehension performance. Environmental and speaker factors, including distracting background noise or rapid speech, can influence a listener's ability to comprehend a spoken message (Wingfield & Tun, 2007; Wingfield et al., 1999). Individual listener factors, such as hearing acuity, age, ability to benefit from contextual cues and cognitive function can also impact discourse comprehension performance (Schneider et al., 2016; Sommers et al., 2011). Further research is necessary to investigate the environmental and individual factors that modulate discourse comprehension performance in order to better understand and eventually reduce the impact of potentially adverse factors.

Discourse comprehension is essential to daily verbal communication and requires listeners to coordinate the use of bottom-up and top-down processing. Top-down processing involves the use of cognitive abilities, prior knowledge, semantics, and contextual information (Kintsch, 1988). Bottom-up processing is the analysis of acoustic properties, such as the phonetic and lexical aspects of the input signal (Schneider et al., 2016). Schneider et al. (2016) describe top-down processing as a cognitive and knowledge-driven process, and bottom-up processing as a stimulus-driven process. Discourse comprehension involves processing at all levels, from acoustic and phonetic information to pragmatics, semantics, contextual cues and the integration of prior knowledge.

Wingfield and Tun (2007) demonstrated the complexities of discourse comprehension by modeling the auditory processing levels necessary to comprehend a spoken message. Wingfield and Tun’s model outlines three main stages or systems that must work in unison to achieve discourse comprehension (Figure 1.1). The authors suggest that sensory, perceptual and cognitive systems are all required to successfully comprehend a spoken message. The sensory stage involves bottom-up processing of acoustic information and is directly impacted by the quality of the input signal. For example, the presence of background noise or rapid speech can cause distortions to the acoustic representations of the input signal, which would impact bottom-up processing at the sensory stage. Accurate processing of acoustic information during the sensory level is beneficial for the processing that occurs in the perceptual and cognitive stages. During the perceptual stage, listeners must recognize the speaker and then perform lexical identification of the spoken message. Lexical identification involves classifying the input received during the sensory stage into meaningful words (Wingfield & Tun, 2007). Next, the cognitive stage involves top-down processes such as semantic determinations, use of contextual information, previous knowledge and multiple cognitive domains. Successful discourse comprehension requires incorporation of all stages of auditory processing. Thus, discourse comprehension may be viewed within a multifaceted framework that requires the efficient processing of multiple levels of the auditory system, knowledge of the language and cognitive ability.

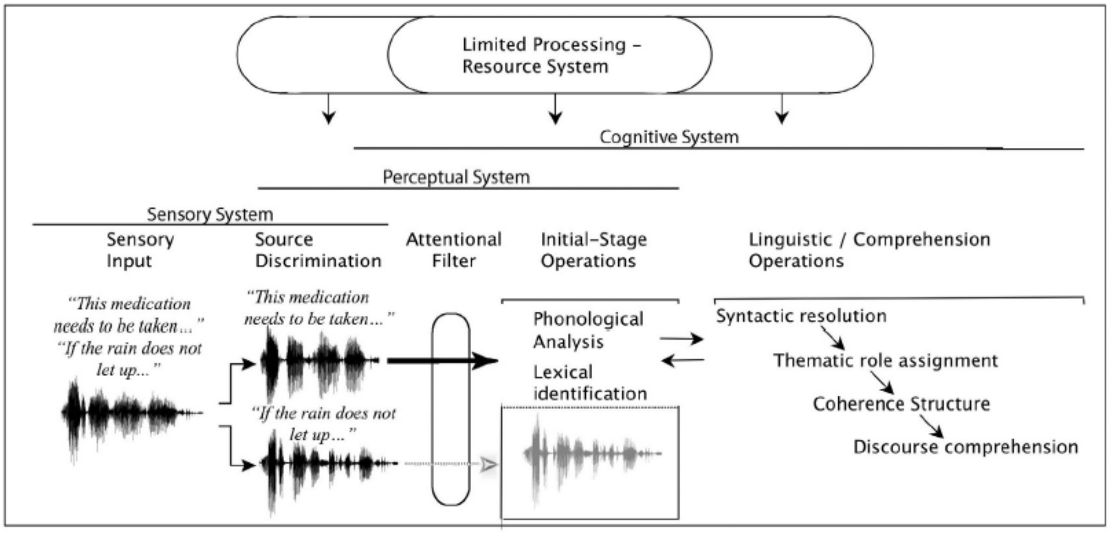

Figure 1.1. Figure from Wingfield and Tun (2007) illustrating the operations necessary for listening comprehension.

Discourse comprehension of a spoken message can be a difficult task, even under ideal listening conditions. Real-world listening environments are often far from ideal, and discourse comprehension becomes increasingly difficult in challenging listening environments that include background noise and rapid speech (Schneider et al., 2010). During these difficult listening situations, hearing loss and cognitive declines experienced by older adults can exacerbate communication difficulties (CHABA, 1988; Nagaraj, 2017; Wingfield et al., 2015). One goal of the current work is to examine how age, hearing loss and specific cognitive domains modulate discourse comprehension performance in difficult listening situations.

Speech Intelligibility

Although discourse comprehension is the skill required for daily communication, there is considerably more research focused on the factors that modulate speech intelligibility. Speech intelligibility is defined as a listener's ability to recognize a sentence or word. Speech intelligibility is typically measured during relatively simple tasks that require listeners to immediately repeat back the spoken message presented. The terms intelligibility and comprehension are often used interchangeably throughout the literature, which can cause misinterpretation of results. Speech intelligibility tasks should not be assumed to directly translate to discourse comprehension performance, which requires listeners to monitor and remember the essence of the spoken message over time (Fontan et al., 2015; Hustad, 2008; Nagaraj, 2017). According to the model outlined by Wingfield and Tun (2007), speech intelligibility and discourse comprehension tasks require different levels of processing. Speech intelligibility tasks typically require listeners to process speech up until the perceptual stage where lexical identification occurs. Therefore, the results from traditional speech intelligibility tasks may not generalize to difficulties experienced during discourse comprehension. Undoubtedly, speech intelligibility and discourse comprehension tasks require different levels of auditory and cognitive processing, and there is a lack of research directed toward understanding how these multiple processes relate. The extent to which multiple factors affect these two tasks (intelligibility and comprehension) under challenging listening situations is largely unknown. Specifically, little is known about the relationship between discourse comprehension, speech intelligibility, background noise, rapid speech, age, hearing loss and cognition. According to Wingfield and Tun (2007), asking listeners to simply repeat back a spoken message does not imply discourse comprehension. Therefore, traditional clinical assessments, which rely primarily on speech intelligibility measures, may not provide a complete representation of real-world communication performance. Real-world communication deficits may be better assessed using a comprehension task because discourse comprehension is a skill required for verbal communication.

Measuring Discourse Comprehension

Discourse comprehension is difficult to measure. In order to measure discourse comprehension performance, participants must listen to a passage and then answer questions or summarize the passage, which leads to long trial times. Long trial times lead to the inability to measure listening comprehension performance in a clinical audiology setting where test time is limited. Given that discourse comprehension is challenging and time-consuming to measure, the majority of studies evaluating aging and speech processing are concerned with the impact of age on speech intelligibility performance. However, several studies have examined the effect of aging on the comprehension of spoken passages. Sommers et al. (2011) studied discourse comprehension performance across the lifespan. The authors examined age-related declines in discourse comprehension and whether or not changes in auditory sensitivity accounted for comprehension difficulties experienced by older adults. Participants ranged in age from 20 to 89 years. Each participant was given three measures of discourse comprehension: (1) the Scholastic Aptitude Test (SAT); (2) the Brown, Carlsen, Carstens listening comprehension test (BCC) (Brown, Carlsen & Carstens, 1995); and (3) the Lectures, Interviews and Spoken Narrative (LISN) test (Tye-Murray et al., 2008). Passages were presented in quiet and ranged from two to ten minutes in length. All passages were followed by visually presented multiple-choice questions or yes/no questions.

Results from the Sommers et al. study indicated that discourse comprehension performance in quiet remained relatively stable through age 65 years and then began to decline significantly for all comprehension measures (Figure 1.2). This result suggests that with a clear non-degraded signal, older adults do not have discourse comprehension declines until age 65 or older. However, a major limitation of this study is that discourse comprehension performance was assessed in quiet. Many older people do not experience difficulty with discourse comprehension or speech intelligibility in quiet (Dubno et al., 1984; Schneider et al., 2000). However, older adults do experience speech-understanding declines in noise, and it is possible that age effects may have been observed in discourse comprehension at a younger age if more difficult listening conditions were employed (Tye-Murray et al. 2008). Thus, a challenging listening situation may be necessary in order to reveal true performance differences between younger and older listeners on discourse comprehension tasks (Rönnberg et al., 2013). In the Sommers et al. (2011) study, a multiple regression analysis was performed to identify the variance accounted for by multiple factors that may predict discourse comprehension performance. Pure tone averages for hearing thresholds accounted for 16.7% of the variance in the LISN scores, 23.6% of the variance for the SAT scores, and 14.1% of the variance for BCC scores. Age accounted for an additional 6 to 7% of variance for all measures. Overall, age and hearing loss were found to account for 23 to 30% of the variance in discourse comprehension performance. There is considerable variance left unexplained by the results of this study.

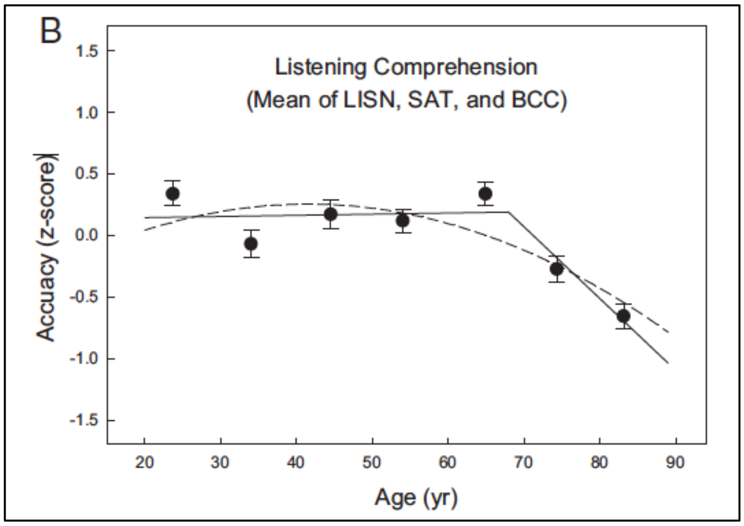

Figure 1.2. Data from Sommers et al. (2011) (Figure 2B) illustrating the decline in discourse comprehension as a function of age.

One goal of the current experiments is to determine the significance of additional factors that may contribute to a decline in discourse comprehension in older adults. Identifying additional factors that may account for the variance in discourse comprehension performance could lead to specific treatment or training paradigms aimed at maintaining discourse comprehension skills in older adults. The degree to which environmental and individual factors, such as signal-to-noise ratio (SNR), masker type, rapid speech and cognitive function, impact discourse comprehension will be studied in the current experiments.

The Role of Background Noise

Speech Intelligibility and Background Noise

The modern world is a noisy place and talkers and listeners are often required to communicate in the presence of background noise (Smeds et al., 2015). It has been well established that older adults, even those with normal hearing, require more favorable SNRs than younger adults with comparable hearing thresholds in order to achieve similar levels of intelligibility performance (Dubno et al., 1984). In a seminal paper on this topic, Dubno et al. (1984) reported that older individuals had greater difficulty understanding speech in noise compared to younger individuals, even when both groups had hearing thresholds within normal limits. Notably, speech intelligibility performance differences between younger and older groups were not found in the quiet condition. Dubno et al. (1984) emphasized the importance of testing speech intelligibility performance in the presence of noise and taking age into consideration when interpreting speech intelligibility results.

Different types of background noise produce differential effects on speech intelligibility performance (Festen & Plomp, 1990). Listeners are frequently exposed to different types of noise in real world environments and each noise type can impact speech perception differently. Specifically, energetic and informational maskers can reveal differences in speech intelligibility performance (Brungart et al., 2001). Energetic masking is typically produced by a steady-state or modulated noise and causes declines in speech intelligibility performance because of an overlap with the target signal in time and frequency. The presence of a competing speech signal can introduce additional informational masking. Informational masking occurs when the acoustic, phonetic, linguistic and semantic characteristics of the target speech and the masker are similar, which causes increased difficulty when separating the target from the masker. For example, a small number of competing talkers would be expected to produce both energetic masking and informational masking, but relatively more informational masking compared to a steady-state noise. In addition, energetic and informational maskers can affect older and younger adults differently. Younger adults typically perform better in the presence of fluctuating noise maskers than older adults, because of their ability to take advantage of a release from masking during the modulations in the noise (Festen & Plomp, 1990). Furthermore, speech-on-speech masking where the masker is intelligible, such as a 1- or 2-talker masker, has been shown to be more detrimental to speech intelligibility performance for older adults compared to younger adults (Bell et al., 2008; Koelewijn et al., 2012; Schurman et al., 2014; Tun et al., 2002; Wiley et al. 1998).

Discourse Comprehension and Background Noise

Older adults do not often complain of difficulty repeating back individual spoken words or sentences, but rather report trouble participating in conversation or comprehending a spoken message. Several studies have aimed to examine the relationship between background noise and discourse comprehension performance and the results are often conflicting. Nagaraj (2017) assessed older listeners with hearing loss on the LISN (Tye-Murray et al., 2008) in quiet and in the presence of a 12-talker babble. Passages from the LISN were 3 to 5 minutes in length and consisted of lectures, interviews and spoken narratives. Overall, listeners performed the same in quiet and in the presence of a 12-talker babble at +5 dB SNR, suggesting that the 12-talker babble may not significantly impact discourse comprehension performance for older adults with hearing loss at a positive SNR. This result is contradictory to seminal studies assessing speech intelligibility performance, which report increased difficulty in the presence of background noise compared to quiet (e.g., Festen & Plomp, 1990). The Nagaraj (2017) study also showed that degree of hearing loss did not predict discourse comprehension performance in quiet or in noise. This result indicates that factors other than hearing loss influence discourse comprehension performance. At the end of each task, participants were asked to rate their listening effort from 0 (no effort) to 10 (maximum effort) using the Borg Scale of Perceived Exertion (Borg, 1982). Participants rated listening effort higher for discourse comprehension in the presence of 12-talker babble compared to that in quiet. This result suggests that although discourse comprehension performance was the same in quiet and in noise, listeners were expending more effort in the presence of 12-talker babble.

In contrast, Tye-Murray et al. (2008) reported that younger and older normal hearing listeners performed worse on the LISN task in the presence of a 6-talker babble in an unfavorable SNR (-5 dB) compared to performance in a favorable SNR (+5 dB). It is possible that the difference in results between the Tye-Murray et al. (2008) and the Nagaraj (2017) study is the type of masker and the corresponding SNR. Results from the Tye-Murray et al. (2008) study suggest that unfavorable conditions, such as background noise with a negative SNR (-5 dB), have a detrimental impact on discourse comprehension performance for younger and older normal hearing adults. In addition, younger adults performed better than older adults in both conditions, possibly implying that discourse comprehension declines with increasing age. However, Tye-Murray et al. (2008) did not find an interaction and reported that younger and older adults were impacted to the same extent by the unfavorable 6-talker babble condition relative to their favorable SNR performance. This could mean that discourse comprehension performance for both groups was impacted to the same extent by the presence of increased background noise compared to younger adults. This result is contradictory to literature examining speech intelligiblity performance, indicating that older adults are more affected by the presence of increased masking compared to younger adults (Dubno et al., 1984; Pichora-

Fuller et al., 1995). Tye-Murray et al. (2008) also presented younger and older adults with a sentence speech intelligibility task in the same conditions as the LISN. Results showed an interaction between group and condition. Older adults were impacted to a greater extent by the unfavorable condition compared to younger adults. Taken together, these findings indicate the possibility that interactions between condition and group for speech intelligibility tasks may not generalize to discousre comprehension performance.

Schneider et al. (2000) also found declines in discourse comprehension performance in the presence of a 12-talker babble compared to a quiet condition for both younger and older adults with normal hearing, but the older adults were more affected by the 12-talker babble than the younger adults. However, when the stories used for discourse comprehension were presented at individually adjusted SNRs equated for 50% correct word intelligibility, the age effect was eliminated. Schneider et al. (2000) suggest that when the target story is audible and 50% intelligible, the presence of a 12-talker babble masker does not impact discourse comprehension performance for older adults. It is possible that older listeners may show declines in discourse comprehension performance with the use of an intelligible 1 or 2-talker masker, even after SNR adjustments are made, due to the increased processing demands of these more challenging maskers (Murphy et al., 2006; Schurman et al. 2014). A target talker presented in the presence of a 1 or 2-talker masker may be perceived as separate speech streams, which could lead to difficulty selectively attending to and focusing on the target signal for a discourse comprehension task (Humes et al., 2006). As the number of interfering talkers increases there will be less speech perceived from the masker, which will enable listeners to focus their attention on the target speech. Figure 1.3 from

Freyman et al. (2004) demonstrates the waveforms of speech maskers as the number of interfering talkers increased from a single talker to 10 talkers. Freyman et al. (2004) emphasizes the increased number of peaks and valleys present in the 1-talker waveform compared to the 10-talker waveform. The 10-talker masker waveform no longer has the speech characteristics (i.e., peaks and valleys) of a 1 or 2-talker masker (Figure 1.3).

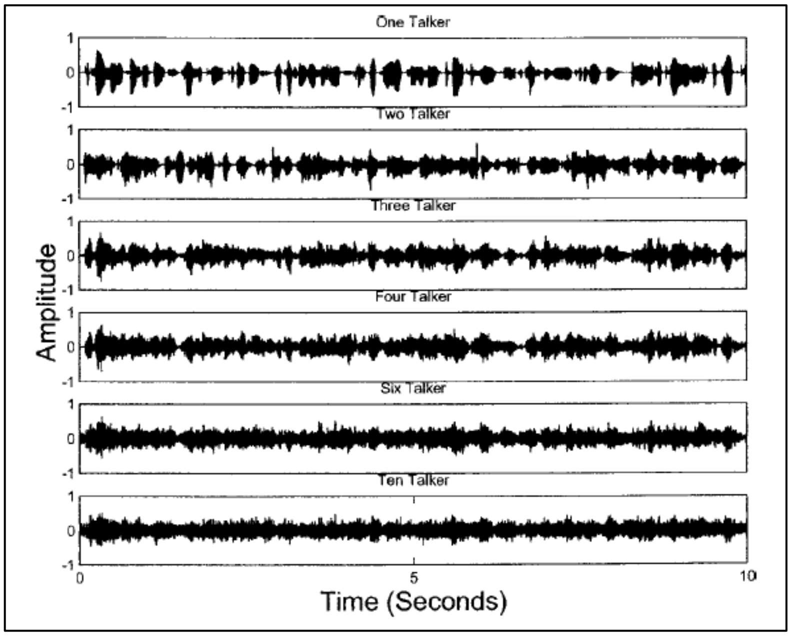

Figure 1.3. Waveforms from Freyman et al. (2004) demonstrate how the characteristics of a masker change from a single-talker masker to ten-talker babble.

In a follow up study to Schneider et al. (2000), Murphy et al. (2006) presented younger and older normal-hearing adults with a discourse comprehension task in the presence of a 2-talker spatially separated masker. Results showed that even when stories were presented at individually adjusted SNRs, discourse comprehension performance was more negatively impacted for older adults compared to younger adults in the presence of a 2-talker masker. This result suggests that older adults may be at a greater disadvantage for discourse comprehension performance compared to younger adults in the presence of this type of speech masker, even when the target signal is individually adjusted for intelligibility.

It is difficult to quantify the impact of different types of maskers on discourse comprehension performance. One reason for this difficulty is that there is often a limited number of accuracy judgments or data points that can be collected following each comprehension passage presented in the masker. Typically, there are only a few comprehension questions at the end of each passage, which results in limited possible performance outcomes across groups and individuals. It may be necessary to assess comprehension using non-traditional methods (i.e., that do not assess accuracy) in order to discover how different conditions impact discourse comprehension. One goal of the current work is to use a method of self-adjustment to discover the impact of different masker types on discourse comprehension performance.

The Role of Time-Compressed Speech

Speech Intelligibility and Rapid Speech

Real world communication unfolds quickly and the ability to process rapid speech is essential to successful participation in meaningful communication. Older adults with and without hearing loss often report difficulty understanding rapid speech. Previous investigations examining the impact of rapid speech on intelligibility performance for older adults have supported these anecdotal reports (Janse, 2009; Wingfield et al., 1985). Speech intelligibility performance for fast speech has been shown to decline more for older adults compared to younger adults (Gordon-Salant & Fitzgibbons, 1993; 2001). Intelligibility performance for rapid speech is often assessed using time-compressed speech, which simulates fast speech production without spectral distortion and is often quantified as a time-compression ratio in relation to the original stimulus length. For example, a speech signal that is time compressed by a ratio of 30% has a duration that is 30% less than the original signal.

In a seminal paper on this topic, Gordon-Salant and Fitzgibbons (1993) presented low-probability R-SPIN (Revised Speech Perception in Noise Test) (Bilger et al., 1984) sentences at multiple time-compression ratios in quiet to younger and older adults with normal hearing and hearing loss. They found that older adults performed significantly worse than younger adults with similar hearing thresholds on final word intelligibility scores (Figure 1.4). Significant differences between older and younger adults for time-compressed speech intelligibility performance have also been reported in the presence of noise (Tun, 1998). Results from both studies indicate that older adults with and without hearing loss are at a greater disadvantage compared to younger adults when processing time-compressed speech. In addition, Gordon-Salant and Fitzgibbons (1993) found that age effects were independent of hearing loss effects on intelligibility for time-compressed speech. This finding suggests that older adults with hearing loss may be at a disadvantage compared to older adults with normal hearing.

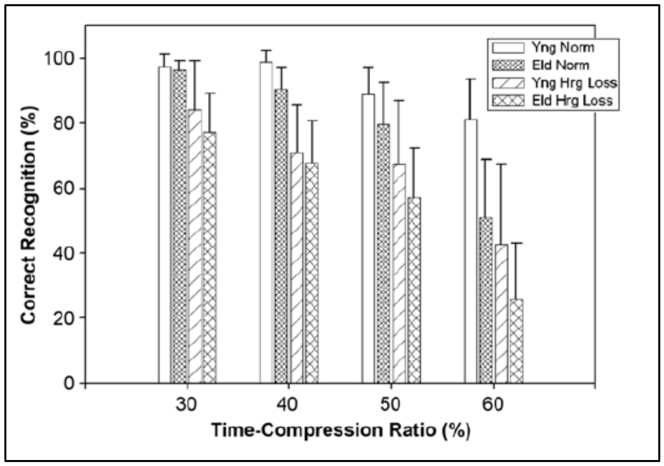

Figure 1.4. Data from Gordon-Salant and Fitzgibbons (1993). Figure redrawn for Gordon-Salant (2005) illustrates the decline in word final intelligibility at four time-compression ratios for younger and older adults with and without hearing loss.

Follow-up studies by the same group further investigated the primary source of the difficulty experienced by older adults when processing time-compressed speech. In the Gordon-Salant and Fitzgibbons (2001) study, younger and older listeners with and without hearing loss were presented with low-probability R-SPIN sentences that were processed using four different methods of time compression: (1) Uniform time compression; (2) time compression of consonant phonemes; (3) time compression of vowel phonemes; and (4) time compression of pauses. The purpose of using multiple time compression methods was to determine the source of difficulty for older adults when processing rapid speech. Results showed that the uniform time compression and the time compression of consonant phoneme conditions were particularly challenging for older adults compared to younger adults. The time compression of the consonants accounted for the most variance (53.3%) for performance in the uniform condition. This result suggests that the difficulty experienced by older adults may be caused by declines in the ability to process degraded time-compressed consonants. However, the remaining variance in performance during the uniform time-compression condition indicates that there were other factors contributing to declines in performance for older adults.

Gordon-Salant and Fitzgibbons (2004) also assessed speech intelligibility for multiple methods of time compression. Younger and older listeners with and without hearing loss were presented with low-probability R-SPIN sentences that were processed using five different methods of time compression: (1) Uniform time compression of the entire signal at a ratio of 50%; (2) 50% time compression of the first segment of the signal; (3) 50% time compression of the second segment of the signal; (4) 50% time compression of the third segment of the signal; and (5) 50% time compression of a randomly selected segment of the signal. In the 2004 study, older adults performed more poorly than younger adults on all forms of time compression. This result suggests that regardless of the time compression method, older adults with and without hearing loss are at a disadvantage compared to younger adults when processing time-compressed speech. Thus, it is possible that an age-related slowing of cognitive processing speed and/or age-related slowed neural processing of speech in the auditory pathway, and not the processing of degraded stimuli, is the dominating factor leading to declines in intelligibility of time-compressed speech for older adults. Taken together, results from the above studies suggest that declines in the processing of time-compressed speech experienced by older adults can be attributed to both the inability to process degraded speech and an overall slowing of processing speed. The above studies have provided important insights into how and why older adults experience declines in intelligibility performance for time-compressed speech. However, it remains unclear whether the declines experienced by older adults for intelligibility of time-compressed speech generalize to declines in rapid discourse comprehension performance.

Discourse Comprehension and Rapid Speech

In order to achieve adequate discourse comprehension, listeners must recognize speech, assign meaning to the speech, and encode the spoken message into memory while new information is being presented. This process must occur rapidly in order for discourse comprehension to be successful. The ability to perform these tasks efficiently declines when the speech is rapid (Wingfield et al. 1999; Wingfield et al., 2003).

Several studies have aimed to quantify the impact of time-compressed speech on discourse comprehension performance. Gordon et al. (2009) studied the impact of time compression on discourse comprehension using spoken lectures that were 10 to 15 minutes in length. They presented time-compressed lectures in quiet and in 12-talker babble to younger and older adults with normal hearing. Time-compressed lectures were presented at a normal rate (120 wpm) and twice the normal rate (240 wpm, or 50% time compression). All lectures were followed by 10 yes or no questions, which were either detail oriented or integrative. Detail questions focused on stated facts and integrative questions required listeners to infer and combine multiple parts of the story. Results showed that both groups answered fewer questions correctly when stories were time compressed compared to the normal rate condition. Gordon et al. (2009) also reported that younger and older adults were equally affected by time compression (Figure 1.5). The authors suggested that when speech is speeded uniformly and within the range of naturally occurring speech rates (i.e., 240 wpm), older and younger adults are impacted equally by time compression. These findings suggest that discourse comprehension performance declines in the presence of rapid time-compressed speech for both younger and older adults with normal hearing. This result could have real-world listening implications because rapid discourse comprehension is often required during daily communication.

Based on previous investigations that assessed the effect of time compression on speech intelligibility performance, it is hypothesized that older adults will show a greater decline in discourse comprehension performance compared to younger adults at faster rates. In contrast to the results reported by Gordon and colleagues, earlier work from Wingfield et al. (1999) supports this hypothesis. Wingfield et al. (1999) presented younger and older normal-hearing listeners with passages at a normal rate and two time-compression ratios (68% and 50%). The time-compressed passages were also restored back to their original presentation length by introducing silences at clause and sentence boundaries. Wingfield and colleagues found an interaction between age and time-compression ratio for discourse comprehension performance, in which older adults with normal hearing were at a greater disadvantage in the presence of rapid speech compared to younger adults with normal hearing. The extent to which time compression impacts discourse comprehension performance for younger and older adults will be examined in Experiment 1.

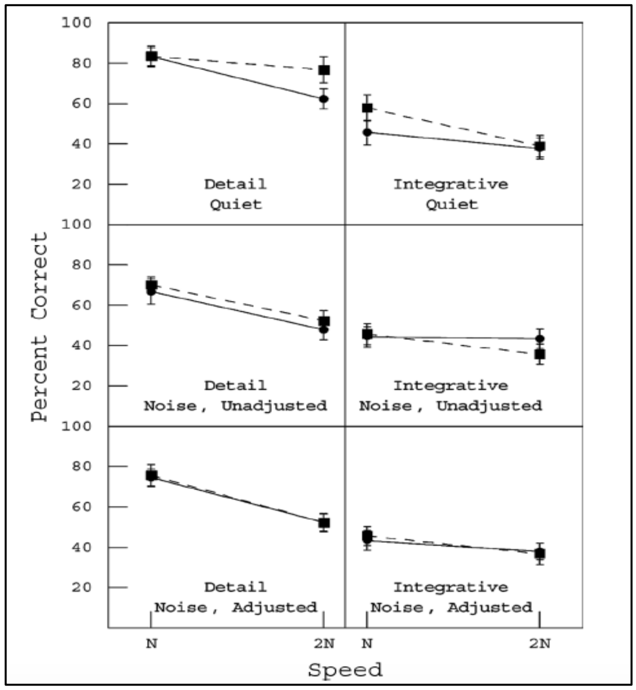

Figure 1.5. Data from Gordon et al. (2009) illustrating the decline in percent correct as a function of speed, question type and age. Circles represent older adults and squares represent younger adults.

The Role of Context in Processing Degraded Speech

Contextual cues and syntactic complexity have been shown to impact the processing and memory of degraded speech signals (Dubno et al., 2000; Pichora-Fuller et al., 1995; Wingfield & Ducharme, 1999; Wingfield et al., 2003). It has been well documented that both younger and older adults benefit from contextual cues during speech intelligibility performance (Dubno et al. 2000; Pichora-Fuller et al. 1995) and similar results have been reported for discourse comprehension performance (Wingfield & Ducharme, 1999). However, there has been some debate regarding the extent to which younger and older adults benefit differently from contextual cues.

Typically, contextual benefit for speech intelligibility performance is assessed using two types of sentence material: high-context and low or anomalous-context sentences. High-context sentences are predictable, and typically have two or more semantic cues to aid in word identification (e.g., His plans meant taking a big risk). Low-context or anomalous-context sentences are difficult to predict and do not have semantic cues to aid in word identification (e.g., Low Context: They discussed the risk; Anomalous Context: His doctor drank a lost risk). The percent correct performance for high-probability sentences is compared to percent correct performance for low-context sentences to determine the benefit or improvement in performance when contextual cues are available (Bilger et al., 1984; Kalikow et al., 1977). Pichora-Fuller et al. (1995) reported that larger set sizes of high context material were remembered compared to low-context material, and that older adults with normal hearing benefit more from contextual cues compared to younger adults with normal hearing for speech intelligibility performance in noise. It is important to note that the older adults performed more poorly for low-context sentences compared to younger adults. Therefore, older adults had more room for improved performance in the high-context condition. In contrast, Schurman et al. (2014) found that younger listeners benefited more from contextual cues compared to older adults during a 1-back sentence memory task in noise. Schurman et al. (2014) utilized high-context sentences and anomalous-context sentences. Results from Schurman et al. (2014) imply that the ability to take advantage of contextual cues may decline for older adults during a task that requires rehearsal and memory, such as a 1-back or discourse comprehension task. Contradictory to both findings, Dubno et al. (2000) argued that the two groups benefit equally from contextual cues during an immediate recall speech intelligibility task in noise when benefit is assessed at the same baseline performance level.

Discourse comprehension requires listeners to build conceptual connections and utilize contextual cues across sentences to make judgments about the spoken message. The ability to take advantage of contextual cues is thought to be critical for discourse comprehension (Wingfield & Tun, 2007). Thus, it is hypothesized that the extent to which a listener is able to benefit from contextual cues could be an influencing factor for predicting discourse comprehension performance.

One study investigated the extent to which the benefit of contextual cues on a speech intelligibility task predicts discourse comprehension performance. Nagaraj (2017) found that the difference in scores between high-probability and low-probability SPIN sentences explained variance in comprehension performance only in the quiet condition, but not in the multi-talker babble condition. This result is contradictory to the theory that suggests the ability to benefit from contextual information is essential to discourse comprehension in challenging listening environments (Wingfield & Tun, 2007). Nagaraj (2017) interpreted these results cautiously and concluded that a more systematic study that relates contextual advantage for speech intelligibility to discourse comprehension is required.

The Relationship Between Comprehension and Intelligibility

The "comprehension theory" outlined by Wingfield and Tun suggests that the sensory stage, which involves the processing of bottom-up acoustic information, must occur with high accuracy in order for discourse comprehension to be successful. Speech intelligibility, or a listener's ability to recognize a sentence or word, is considered part of the sensory stage. Thus, it may be expected that speech intelligibility influences discourse comprehension performance and there is a strong relationship between these two processes. However, the direct relationship between discourse comprehension and speech intelligibility performance is difficult to quantify.

Several studies examined the relationship between discourse comprehension and speech intelligibility (Fontan et al., 2015; Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015) and found that speech intelligibility performance did not predict discourse comprehension performance. Hustad (2008) examined the relationship between discourse comprehension and speech intelligibility for normal hearing adults when listening to passages and sentences produced by multiple individuals with dysarthria. Dysarthria is a motor speech disorder that often causes speech to be unintelligible with varying degrees of severity. In the comprehension task, participants were presented with spoken passages and asked to answer questions after each passage. During the intelligibility task, listeners were presented with individual sentences and asked to immediately recall each sentence. Twelve different dysarthric speakers produced speech stimuli. The speakers were separated into four groups of dysarthria severity: mild, moderate, severe and profound dysarthria. Results showed that there was no relationship between discourse comprehension and sentence intelligibility performance, except in the mild dysarthria condition. Intelligibility performance accounted for 12% of the variance in discourse comprehension performance in the condition with the mildly dysarthric talker and did not account for the variance in the more severe dysarthria conditions. This result suggests that at least one form of degraded speech, dysarthric speech, differentially modulates discourse comprehension and speech intelligibility. The authors noted that this finding is problematic for clinicians who may predict discourse comprehension from speech intelligibility. Another factor that was not assessed in this study, cognitive function, may explain additional variance in discourse comprehension performance for degraded speech.

More recent studies were also unsuccessful in finding a strong relationship between discourse comprehension and speech intelligibility. For example, Smith and Pichora-Fuller (2015) measured discourse comprehension in quiet and sentence intelligibility in noise for younger adults with normal hearing and older adults with and without hearing loss. Discourse comprehension was measured using the LISN, and speech intelligibility was measured using the R-SPIN and the QuickSIN (Speech-In-Noise-Test) (Killion et al., 2004). The results showed no relationship between discourse comprehension and speech intelligibility performance. However, there were limitations of the Smith and Pichora-Fuller study. Discourse comprehension was measured in quiet, speech intelligibility was measured in noise, and different talkers recorded stimuli for each type of assessment. Thus, the lack of a correlation between sentence intelligibility and discourse comprehension could have been affected as much by the differences in talkers and quiet/noise conditions as by the different nature of intelligibility versus comprehension tasks.

As stated previously, Nagaraj (2017) also examined the relationship between discourse comprehension and speech intelligibility performance for older adults with hearing loss and found no relationship between the two measures. Nagaraj (2017) aimed to determine the variance in discourse comprehension performance that was explained by hearing loss and speech intelligibility. Discourse comprehension was measured with the LISN, and speech intelligibility was measured with the R-SPIN (Bilger et al., 1984) and QuickSIN (Killion et al., 2004). All three tests were presented in quiet and at a fixed +5 dB SNR for 12-talker babble. Results revealed that neither speech intelligibility performance nor degree of hearing loss accounted for a significant proportion of variance in discourse comprehension performance either in quiet or in noise. This result indicates that speech intelligibility performance and hearing thresholds were not significant variables contributing to discourse comprehension performance for older listeners with hearing loss.

The three prior studies comparing speech intelligibility to discourse comprehension all reported findings contradicting a relationship between intelligibility and comprehension (Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015). However, the theoretical framework proposed by Wingfield and Tun (2007) suggests that high speech intelligibility performance is a necessary pre-requisite for adequate discourse comprehension performance. Results from Schneider et al. (2000) would support this notion. Schneider and colleagues suggest that if speech is intelligible, discourse comprehension is preserved. Schneider et al. (2000) presented younger and older normal hearing listeners with spoken passages in the presence of a 12-talker babble and in quiet. Results showed that discourse comprehension performance declined more in older adults compared to younger adults when passages were presented in the 12-talker babble. Age effects were eliminated when younger and older adults were presented with passages at individually adjusted SNRs for equivalent (50% correct) word intelligibility using the R-SPIN. In other words, when comprehension passages were presented at individual SNRs that yielded equivalent word intelligibility, older adults performed similarly to younger adults. The authors therefore attributed the age-related decreases in discourse comprehension performance in noise to associated declines in speech intelligibility in the presence of the masker. This outcome implies that the main source of difficulty for older adults during discourse comprehension in noise is a decline in the bottom-up sensory stage of processing, not top-down cognitive processing. However, it is conceivable that individually adjusting SNRs alters the acoustic difficulty of the stimuli and decreases the cognitive processing load. Therefore, an alternative possibility is that the cognitive demands, such as attention and working memory, of the discourse comprehension task vary with the SNR. The next section discusses the influence of cognitive function on discourse comprehension performance.

The Role of Cognitive Function

The interaction between discourse comprehension and cognition is multifaceted. Furthermore, the relationship between cognition and discourse comprehension may vary based on environmental and talker factors, such as background noise and time compression, or individual factors, such as age and hearing loss. Several studies have attempted to examine the relationship between discourse comprehension and cognition, and multiple theories have been generated to explain this relationship. Generally, there is strong evidence that supports the influence of working memory capacity, attention, and speed of processing on speech intelligibility and discourse comprehension performance (Pichora-Fuller et al., 2016; Rönnberg et al., 2013; Salthouse, 1996). There is a debate regarding which cognitive domain is most influential for discourse comprehension performance in younger adults and older adults with and without hearing loss (Schneider et al., 2010). It is also possible that there are differences in top-down processing strategies between younger and older adults for discourse comprehension (Schneider et al., 2016). However, the nature of the differences in top-down processing between younger and older adults is largely unknown. The domains of working memory, processing speed and attention are discussed below.

Working Memory

Working memory is often defined as the ability to manipulate and temporarily store incoming information (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980). Discourse comprehension also requires listeners to manipulate and briefly store incoming information. Given that discourse comprehension and working memory both require similar manipulations and storage of information, it may be anticipated that working memory capacity accounts for a significant amount of variance in discourse comprehension performance, especially in noisy listening environments. Noisy listening situations are thought to require more top-down resources in order to successfully process the target signal compared to a quiet situation. The Ease of Language Understanding (ELU) Model suggests that increased demands on bottom-up processing require listeners to rely more on top-down working memory resources (Rönnberg et al., 2013). Rönnberg and his colleagues report a strong relationship between working memory capacity and speech intelligibility performance in noisy situations (Rudner et al., 2012). This suggests that listeners with greater working memory resources are able to compensate using top-down processing when listening to a speech signal in noise. The ELU model focuses on the relationship between working memory and speech intelligibility performance. However, the authors highlight the need to evaluate the relationship between working memory capacity and more functional outcome measures, such as discourse comprehension. Several studies have attempted to quantify this relationship, but results have been inconsistent.

Gordon et al. (2009), discussed previously, also examined the relationship between age, working memory and time-compressed discourse comprehension performance. Younger and older normal hearing adults were assessed on the Reading Span (RSPAN) test and 10- to 15-minute lectures in quiet and in the presence of a 12-talker babble. Lectures were presented at a normal rate and a speed twice the normal rate. Importantly, the lectures were presented at individually adjusted SNRs for word intelligibility, similar to the method used by Schneider et al. (2000). According to the ELU model, presenting listeners with discourse comprehension material at individually adjusted SNRs ensures audibility, but it also alters the top-down cognitive demands required for the discourse comprehension task. In addition, it is not entirely known how the SNR required for intelligibility interacts with discourse comprehension performance. Gordon et al. (2009) found a correlation between working memory and discourse comprehension performance in younger listeners but did not find a correlation for older listeners. This result suggests that younger adults were more reliant on top-down cognitive processing compared to older normal hearing adults, which is inconsistent with previous literature. Previous literature examining the relationship between working memory and speech intelligibility in noise suggests that older listeners need to recruit more top-down resources in order to compensate for age-related declines in bottom-up processing (Rönnberg et al., 2013), which was not observed in this study.

One reason for the contradictory result reported by Gordon et al. (2009) is that assessing discourse comprehension at individually adjusted SNRs for equivalent intelligibility increased the amount of cognitive recourses available for passage recall, which helped to improve performance for the older listeners. Additionally, lectures were 10 to 15 minutes in length, which is normally considered a long-term memory interval. Working memory capacity is typically defined in terms of shorter processing intervals (Baddeley & Hitch, 1974). Therefore, working memory capacity may not be directly related to processing 10 to 15-minute lectures. Importantly, Gordon et al. (2009) did not measure processing speed, which may have been involved in the processing of time compressed speech.

A more recent study by Smith and Pichora-Fuller (2015) was also unsuccessful in finding a strong relationship between working memory (as measured by the RSPAN) and discourse comprehension performance in quiet for older adults with hearing loss. This result may be explained by the assessment of comprehension passages in quiet. The use of more challenging or cognitive demanding background noise, such as a 1-talker (1T) background, during the discourse comprehension task may have revealed a different result. Previous literature would suggest that a more challenging background noise, which engages top-down recourses, may show a stronger relationship with working memory capacity (Rönnberg et al., 2013). For example, Koelewjin et al. (2012) found a larger pupil size, which is thought to indicate increased listening effort, during a speech intelligibility task in the presence of a 1T masker compared to a speech-shaped noise masker. A study by Ward et al. (2016) found that working memory as measured with the RSPAN was correlated with discourse comprehension performance of vocoded one-minute passages for both younger and older adults with normal hearing. Vocoding is a signal processing technique that preserves temporal envelope information and alters the spectral detail. This result indicates that working memory capacity is related to discourse comprehension performance for degraded vocoded speech. It is clear that more work is necessary to examine the relationship between discourse comprehension performance and working memory for younger and older adults with and without hearing loss. It is also important to assess alternative measures of cognition, such as attention and speed of processing, to determine if these abilities can explain additional variance in discourse comprehension performance.

Speed of Processing

Multiple processes must occur in unison in order for discourse comprehension to be successful. A listener must be able to rapidly receive auditory information, create a mental representation, integrate the representation into memory while new information is being presented, and later retrieve the information to formulate an appropriate response. Processing speed, or how quickly an individual can perform this set of tasks, would be expected to influence discourse comprehension performance (Wingfield et al., 1999). Speed of processing is often defined as the rate at which an operation is performed in order to accomplish a given task (Salthouse, 1996). The Processing-Speed Theory suggests that speed of processing plays a primary role in language processing (Salthouse, 1996). Salthouse suggests that two mechanisms control processing speed as it relates to language processing: the limited time mechanism and the simultaneity mechanism (Salthouse, 1996). The limited time mechanism suggests that the time spent processing early operations will restrict the amount of time available for later operations, therefore all relevant operations may not be completed. Thus, accurate recall may not be feasible due to the slow processing of early operations, such as phonological processing. The simultaneity mechanism proposes that information processed early may no longer be available when later processing is complete. This mechanism would suggest that declines in processing speed could lead to the loss of relevant early processed information by the time it is needed for recall. The Processing-Speed Theory suggests that speed of processing has some influence on discourse comprehension performance.

Time-compressed speech stimuli are typically used to study the impact of processing speed on discourse comprehension and speech intelligibility performance for younger and older adults. Although age-related declines in the intelligibility of time-compressed speech have been well documented (Gordon-Salant & Fitzgibbons, 1993, 2001; Wingfield et al., 1985), the main source of difficulty for older adults has been debated. There are two central hypotheses: (1) The cognitive hypothesis, which suggests a cognitive decline in processing speed is the primary reason the older adults experience difficulty processing time-compressed speech (Salthouse, 1996) and (2) The perceptual hypothesis, which suggests that declines in the peripheral auditory system and the ability to process degraded phonetic and acoustic information introduced by time compression is the primary source of difficulty (Schneider et al. 2005). A third hypothesis, which will not be examined in the current experiments, is that a decline in central auditory processing or neural synchrony is the source of difficulty for older adults when processing time-compressed speech.

Schneider et al. (2005) argue that older adults' difficulty processing time-compressed speech is attributed to age-related auditory perceptual declines or difficulty processing distortions in the speech signal. Schneider et al. (2005) presented younger and older normal hearing listeners with sentences processed using three different methods of time compression. Two of the time-compression methods introduced acoustic and phonemic distortions (e.g., translating frequencies upward and removing phonemes) and one produced minimal acoustic distortion by time compressing pauses and steady-state portions of the speech. Listeners were asked to repeat back the final word in each sentence. Results showed that older adults performed worse than younger adults in the time-compression conditions that produced acoustic distortions, but both groups preformed similarly in the condition with limited acoustic distortions. Given a time-compressed speech sample with limited distortions, older adults and younger adults perform the same. This result suggests that the difficulty older adults experience when processing rapid speech is attributed to the ability to process acoustic distortions introduced by time compression. Thus, Schneider and colleagues would support the perceptual hypothesis and conclude that age-related auditory declines, not slowed processing speed, modulate intelligibility performance for time-compressed speech in older normal hearing adults.

The Schneider et al. (2005) study examined the impact of time compression on sentence-level material and found that a word-final intelligibility task was less difficult than a discourse comprehension task. Wingfield et al. (1999) studied how time compression affects discourse comprehension. They presented younger and older normal-hearing listeners with paragraph-length passages (M= 229 words) at a normal rate and two time-compression ratios (68% and 50%). The time-compression technique produced minimal acoustic distortions. Results showed that increasing the time compression of the passage decreased discourse comprehension performance compared to the normal rate condition for both younger and older adults. Wingfield et al. (1999) also presented passages that were time compressed and then restored to their original length by inserting silences at syntactic and clause boundaries. Therefore, the passages were still time compressed, but listeners now had more time to process the incoming information during the silences. Results indicated that the insertion of silences improved recall for both younger and older adults. The addition of processing time eliminated all effects of time compression for younger listeners, and their performance returned to the baseline-uncompressed condition. Although older adults' discourse comprehension performance improved with insertion of silences, older listeners were never able to completely recover to their baseline comprehension performance in the normal-rate condition. This result supports the influence of both hypotheses, suggesting that difficulty processing acoustic and phonetic distortions and slowed processing speed contribute to older listeners' comprehension decline with rapid speech.

Piquado et al. (2012) also examined the impact of auditory processing and cognitive function on discourse comprehension performance. Young listeners with and without hearing loss were presented with narrative passages. During the baseline condition, passages were presented at a continuous normal rate and listeners were asked to recall the narrative details. Results showed that listeners with hearing loss recalled passages more poorly than normal-hearing listeners. This result suggests that hearing loss is detrimental to discourse comprehension performance. In the second condition, silences were placed at syntactic boundaries and participants were told to self-pace the passage by pressing the keyboard when they were ready for the next segment of the narrative. In this condition, the discourse comprehension performance of participants with hearing loss improved to equal the discourse comprehension performance of normal hearing listeners. Thus, the introduction of increased processing time allowed individuals with hearing loss to perform similarly to their normal hearing counterparts. The authors propose that hearing loss may slow the processing of incoming bottom-up information, leading to declines in the top-down resources required for successful discourse comprehension. Older adults with and without hearing loss were not assessed in this study. The relationship between auditory distortions and processing speed during discourse comprehension for older adults with and without hearing loss remains unclear. One goal of Experiment 1 is to determine the relative impact of processing speed and acoustic degradation on discourse comprehension performance by younger normal hearing adults and older adults with and without hearing loss.

Attention

Attention is an important cognitive domain that may support discourse comprehension performance. Attention is a complex concept that includes orienting, selecting and focusing on environmental stimuli over time. Attention is comprised of multiple categories, including divided attention, sustained attention and selective attention (Kahneman, 1973). The following experiments will focus on selective attention, which is the focusing of attention on some aspect of the stimulus and the inhibition of other aspects (Craik & Lockhart, 1972). Selective attention is specifically of interest because all stages of speech processing in noise, from phoneme recognition to discourse comprehension, require the listener to selectively attend to a target speaker and inhibit distractions in the background (Wingfield & Tun, 2007). In addition, selective attention has been shown to decline with age (Hasher & Zacks, 1988) and may influence discourse comprehension and speech intelligibility performance in noise for older adults (Guerreiro et al., 2013; Helfer & Freyman, 2008).

A seminal theory on selective attention described by Hasher and Zacks (1988) postulates that inhibitory control, or the ability to suppress irrelevant distractions, underlies the decline in older adults' selective attention ability. Hasher and Zacks (1988) suggest that declines in selective attention are particularly damaging to reading comprehension performance. Although the Hasher and Zacks theory on selective attention is focused on reading comprehension, the theory can also be applied to spoken discourse comprehension performance. Discourse comprehension demands a high memory load, and tasks with increased memory loads are thought to be more susceptible to distraction (Hasher & Zacks, 1998). Hasher and Zacks propose that the inability to selectively attend and inhibit irrelevant information will place higher demands on memory storage capacity and therefore limit the recall accuracy of important information during discourse comprehension. Although there are strong theories that demonstrate the probable importance of selective attention on discourse comprehension performance, the majority of auditory research has focused on the relationship between selective attention and speech intelligibility performance. One goal of the present study is to determine the relationship between selective attention and discourse comprehension performance.

Summary and Overall Goal

The overall goal of this research is to determine which factors contribute most to discourse comprehension performance in difficult listening situations. Discourse comprehension is a skill required to communicate verbally with others, and the inability to comprehend a spoken message can lead to social isolation, loneliness and cognitive declines for older adults (Lin et al., 2011). In order to mitigate the negative effects of possible age-related decline in discourse comprehension, effort must be expended to examine the factors that influence discourse comprehension during real world listening situations. Real-world environments produce challenging listening conditions, including competing background noise and rapid speech. In addition to demanding listening scenarios, individual factors including age, hearing loss, cognitive function, and the ability to benefit from context can influence discourse comprehension performance. There has been much debate regarding the relative contributions of each factor in predicting discourse comprehension performance. The following experiments attempt to address this issue by examining discourse comprehension as it relates to multiple environmental and individual factors.

Experiment 1

Introduction

Real world conversations unfold quickly, and the ability to process rapid speech is essential to successful participation in meaningful communication. Older adults with and without hearing loss often report difficulty comprehending rapid speech (Gordon et al. 2009). Discourse comprehension requires listeners to interpret the meaning of an incoming message, integrate the message into memory and use the information to respond appropriately (Schneider et al., 2010). The ability to perform this series of events declines when speech is presented rapidly, especially for older adults (Wingfield et al. 1999; Wingfield et al., 2003). However, there are conflicting hypotheses regarding the reason for speed-induced comprehension difficulties.

One theory, known as the perceptual hypothesis, is that difficulty processing rapid speech arises from the degraded phonetic and acoustic information introduced by the time compression method (Schneider et al. 2005). This theory focuses on the time-compression method and rate at which the speech segments are produced. The rate at which speech segments are spoken (excluding silences) is known as the articulation rate (Jacewicz, et al., 2009). Articulation rate includes the possible distortion present in the speech signal introduced by the time compression method. In contrast, information rate or speaking rate refers to the presentation rate of the entire speech signal (i.e., speech segments and present silences) (Jacewicz, et al., 2009). In this paper, the term information rate is used to refer to the rate of the overall speech signal. The second hypothesis is that difficulty processing rapid speech is due to cognitive speed of processing constraints and the ability to process the overall information rate of the signal (Piquado et al., 2012; Salthouse, 1996).

Schneider et al. (2005) contend that older adults' difficulty processing time-compressed speech is attributed to difficulty processing distortions in the speech signal. To examine this theory, Schneider et al. (2005) presented younger and older normal hearing listeners with sentences processed using three different methods of time compression, which varied in degree of acoustic distortion. Two of the time-compression methods introduced acoustic and phonemic distortions (e.g., translating frequencies upward and removing phonemes) and one produced minimal acoustic distortion by only time compressing pauses and steady-state portions of the speech. Listeners were asked to repeat back the final word in each sentence. Schneider and colleagues showed that older adults had more difficulty than younger listeners with a word identification task when the time compression method produced phonemic and acoustic degradation (e.g., translating frequencies upward and removing phonemes). However, both groups performed similarly when the time compression method produced minimal distortions. These results suggest that if a method of time compression is utilized that limits distortion [e.g., Pitch-Synchronous Overlap and Add (PSOLA)], then older and younger adults should perform similarly on a speech perception sentence task. Taken together, results from Schneider et al. (2005) indicate that the ability to process the speech segments or the articulation rate of the signal, not the increased demands on processing speed or the overall information rate, is the source of difficulty.

Wingfield et al. (1999) examined the influence of both time-compression hypotheses on discourse comprehension performance. Younger and older normal-hearing listeners were presented with passages at a normal rate and two time-compression ratios (68% and 50%). After the completion of each passage, participants were instructed to recall as much of the story as possible. Results showed that story recall declined for both groups during the time-compression conditions compared to recall during the normal-rate condition. Then, Wingfield et al. (1999) restored the time-compressed passages back to their original presentation length by introducing silences at clause and sentence boundaries. Therefore, the spoken segments were still time-compressed, but listeners now had more time to process the incoming information during the silences (i.e., they restored the information rate). Results for younger adults showed that the introduction of silences restored discourse comprehension performance to the performance level seen during the normal-rate condition. Therefore, given additional time to process the time-compressed signal, younger listeners performed similarly in this time-restoration condition compared to the normal-rate condition. This finding would support the hypothesis that constraints on speed of processing or the ability to process the overall information rate of the signal were causing the decline in discourse comprehension performance for time-compressed speech, for younger listeners with normal hearing.

Results for the older listeners from the Wingfield et al. (1999) study were less straightforward. The insertion of pauses within the time-compressed passages improved comprehension performance for older adults, however, their performance did not return to the same level of performance observed in the normal-rate condition. Even with additional time to process the time-compressed signal, older adults nevertheless experienced difficulty. This result supports the influence of both hypotheses: (1) difficulty processing rapid speech arises from the reduced acoustic information introduced by time-compression (i.e., articulation rate) and (2) difficulty processing rapid speech is due to the cognitive speed of processing constraints introduced by time compression (i.e., information rate). Wingfield et al. (1999) suggested that both difficulty processing degraded acoustic and phonetic information and demands on processing speed may contribute to older listeners' comprehension declines with rapid speech.

Wingfield et al. (1999) noted that the assessment of cognitive function and speech intelligibility performance may lead to a better understanding of the impact of rapid speech on discourse comprehension performance. The assessment of cognitive function and speech intelligibility performance could clarify which one of the two proposed hypotheses better explains why older adults have difficulty with rapid discourse comprehension. Measuring cognitive function could indicate the extent to which a cognitive decline is related to difficulty processing rapid, time-compressed speech. Results from Vaughan et al. (2006) also support the importance of assessing cognitive function as a predictor for time-compressed speech intelligibility performance. Results indicated that working memory was strongly associated with the processing of time-compressed sentences for older adults with and without hearing loss.

Wingfield and Ducharme (1999) also examined the impact of time-compressed speech on discourse comprehension performance. They implemented an approach that required listeners to self-select their preferred time-compressed speech rate for a spoken passage. Participants were asked to turn a knob to select a time-compression ratio where they felt that the passage could be understood and accurately recalled. However, they were not asked to recall details about the stories during the self-selection task. Results indicated that older adults selected a slower rate compared to younger adults during the self-selection speech rate task. Time-compression ratio selection was reliable for both groups, leading the authors to suggest that younger and older adults are able to appropriately moderate their preferred listening rates. Next, comprehension recall for passages was assessed for three time-compression conditions: (1) 67% of the original passage; (2) 53% of the original passage; and (3) at each individual's self-selected rate. Results showed that younger adults were able to recall more information about the passages compared to older adults in all time-compression conditions. Notably, older adults showed poorer recall compared to younger adults even in the condition where passages were presented at individually selected rates.

Older adults often report difficulty understanding rapid speech. One way to quantitatively capture this complaint would be to compare the self-selected listening rates for younger and older adults. A method of self-selection is clinically feasible and could provide valuable information regarding the impact of time compression on discourse comprehension performance in younger normal-hearing adults and older adults with and without hearing loss in a relatively brief period of time. One goal of the present study is to determine the relative impact of processing speed (information rate) and reduced acoustic information (articulation rate) on discourse comprehension performance by younger normal hearing adults and older adults with and without hearing loss using a self-selection method. Comparing younger and older normal-hearing adults will address questions about processing speed changes with aging, while comparing older normal-hearing and older hearing-impaired listeners will address questions about the effect of reduced audibility and possible acoustic degradation of the speech signal associated with hearing loss.

The current study implements a self-selection method similar to that described by Wingfield and Ducharme (1999) and incorporates conditions with and without silences, in order to determine the impact of information and articulation rates. The self-selected time-compression ratio will serve as a proxy for discourse comprehension accuracy performance. Based on pilot testing, the current self-selection method is expected to produce relatively fixed levels of comprehension accuracy performance (i.e., near 80% correct) across a wide range of individual time-compression selections. In other words, there should be substantial differences in self-selected time-compression ratios required to achieve 80% comprehension accuracy performance. Therefore, the variable of interest is the time-compression ratio that yields the equivalent percent correct performance. This method of measuring discourse comprehension performance may lead to a better understanding of difficulty processing time-compressed speech, especially for older adults with and without hearing loss.

Research Questions and Hypotheses

The goals of the present experiment are to: (1) determine the effect of speech presentation rate condition (silences intact vs. silences removed), age, and hearing loss on self-selected time-compression ratio during a discourse comprehension task for young normal hearing adults and older adults with and without hearing loss; (2) determine the relationship between speech intelligibility performance for words and sentences and self-selected time-compression ratio for younger adults with normal hearing and older adults with and without hearing loss; and (3) determine which cognitive domain (i.e., working memory, selective attention or speed of processing) is the most important predictor of self-selected time-compression ratio during discourse comprehension by younger normal-hearing adults and older adults with and without hearing loss. Lastly, characterizing the unique self-selection method is also of interest in the current study.

(1) Do speech presentation rate condition (i.e., silences intact or silences removed), age and hearing loss affect the self-selected time-compression ratio during a discourse comprehension task?

This question is motivated by two theories about processing rapid speech. The two theories postulate that declines in the processing of time-compressed speech are attributed either to (1) the listener's inability to process the degraded phonetic and acoustic signal that occurs as a result of rapid speech, or (2) the listener's limitations in cognitive speed of processing. Based on results from Wingfield et al. (1999), it is hypothesized that self-selected time-compression ratio will be faster in the silences intact condition compared to the silences removed condition for all groups. This result would suggest that the presence of silences allows listeners to select a faster time-compression ratio, because listening rate is modulated by available processing time or the information rate of the entire signal.

It is hypothesized that both groups of older adults will self-select a slower time-compression ratio compared to younger listeners for both speech presentation rate conditions. This result would suggest that older adults prefer a slower listening rate during a discourse comprehension task compared to younger listeners. Based on results from Wingfield et al. (1999), an interaction between age and self-selected speed is expected. Wingfield and colleagues found that the addition of extra processing speed improved young normal hearing (YNH) listeners' performance to that observed at the baseline normal speech rate. Older normal hearing (ONH) listeners' performance improved also, but not to the baseline level of performance, suggesting that ONH listeners were not able to take full advantage of the increased processing time. Finally, it is hypothesized that older adults with hearing loss will select a slower self-adjusted listening rate in both speech presentation rate conditions compared to older normal hearing adults. Although this has not been directly measured in previous discourse comprehension studies, results from previous speech intelligibility studies with time-compressed speech showed that older adults with hearing loss performed better at slower rates compared to faster rates, suggesting that they are likely to prefer a slower speed (Gordon-Salant & Fitzgibbons, 1993; 2001).

(2) Is there a relationship between speech intelligibility performance and self-selected time-compression ratio for discourse comprehension for younger adults with normal hearing and for older adults with and without hearing loss?

The relationship between discourse comprehension and speech intelligibility, specifically for time-compressed speech, is largely unknown. No previous study has directly studied the relationship between speech intelligibility and comprehension for rapid speech. Previous studies have attempted to quantify the relationship between percent correct performance on a comprehension task and percent correct performance on an intelligibility task in background noise (Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015). These previous studies found no relationship between percent correct performance on a speech intelligibility task and a comprehension task. Based on the “comprehension theory” defined by Wingfield and Tun (2007) it is clear that measures of speech intelligibility performance and measures of discourse comprehension performance are evaluating different processes. Speech intelligibility accuracy performance and comprehension accuracy performance may not have a linear relationship and an absolute comparison of percent correct performance would not be an appropriate technique to evaluate the relationship between these two processes.

The current study will take a unique approach in order to better determine how speech intelligibility impacts discourse comprehension performance. The self-selected listening rate will be used as a proxy for a fixed level of discourse comprehension performance. It is hypothesized that speech intelligibility performance will not account for significant variance in self-selected listening rates for either younger adults or older adults during discourse comprehension. This result would suggest that additional factors, such as age and cognitive function modulate the preferred listening rate during a discourse comprehension task.

(3) Across speech presentation rate conditions, which cognitive domain contributes most to variance in self-selected time-compression ratio by listener groups that vary in age and hearing loss?

It is hypothesized that cognitive function will predict self-selected time-compression ratio, suggesting that discourse comprehension performance is influenced by working memory, attention and speed of processing. Specifically, speed of processing will be the most predictive of self-selected time-compression ratios during the discourse comprehension task (Salthouse, 1996). It is also predicted that cognitive function will be more predictive of older listeners' performance compared to younger listeners' performance, because older adults require the use of more top-down processing to compensate for a degraded auditory signal (Rönnberg et al., 2013).

Methods

Participants

Three groups of 18 individuals participated in this study: younger adults aged 18-30 years with normal hearing (YNH), older adults 60 years and older with normal hearing (ONH) and older adults (60 years and older) with hearing loss (OHI). Normal hearing was defined as thresholds ≤ 25 dB HL between 250 and 4000 Hz (re: ANSI, 2018). Participants with hearing loss had thresholds ≥ 30 dB HL or at least a mild hearing loss between 2000 Hz and 8000 Hz. Figure 2.1 illustrates the average auditory thresholds for each group. Participants had at least a high school education and were native speakers of English. The Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) was used as an assessment of mild cognitive impairment. Participants who passed the test with a score of ≥ 26 (i.e., did not have a mild cognitive impairment) participated in the study.

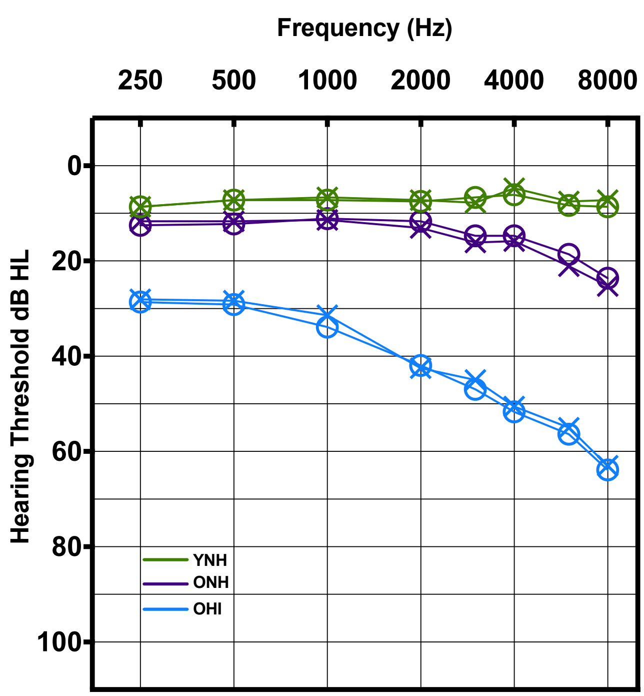

Figure 2.1. Mean audiograms for YNH, ONH and OHI groups [Right ear = O; Left ear = X).

Stimuli

All speech materials were recorded using a DPA4017B shotgun microphone with a 48kHz-sampling rate. Each story, sentence and word was stored as a separate waveform file using Adobe Audition, and the root-mean-square (RMS) levels were held constant for all stimuli.

Speech stimuli were comprised of comprehension passages, sentences and words. In order to remove talker variability as a confounding variable, one native English female talker recorded all speech materials. Comprehension stimuli were stories from the Discourse Comprehension Test (DCT) (Brookshire & Nicholas, 1993). The DCT is comprised of twelve 1-minute passages. Each narrative is followed by eight yes or no questions, which assess either the main idea of the passage (stated and implied) or story details (stated and implied). DCT questions were presented visually via a computer monitor. Sentence materials were 200 high probability (HP) R-SPIN sentences and 200 anomalous probability (AP) sentences, which were derived from the HP sentences (Bilger et al., 1984). All nouns, verbs and adjectives were considered keywords, and there were three to seven keywords in each sentence. (HP example: His PLANS MEANT TAKING a BIG RISK. AP example: His DOCTOR DRANK a LOST RISK.). Individual word intelligibility was measured using a revised version of the Modified Rhyme Test (MRT) (Bell et al., 1972).

A combination of Matlab (Matlab 2017a) and Praat (Boersma & Weenink, 2017) was utilized to generate the time compressed DCT stories, SPIN sentences and MRT words. The PSOLA method (Moulines & Charpentier, 1990) was used to create all time-compressed stimuli. PSOLA, a uniform method of time compression, was selected in order to minimize acoustic distortions and pitch changes in the resulting time-compressed signal (Moulines & Charpentier, 1990; Janse, 2009; Schlueter et al. 2014). Time compression is accomplished by deleting alternating pitch periods, concatenating the remaining speech segments, and applying a smoothing function at the signal onsets and offsets to remove extraneous clicks. In order to further decrease the possibility of distortion, the original DCT waveform files were used to generate preprocessed time-compressed or time-expanded versions of each story. Each DCT story had its own folder, which included 35 subfolders, one folder for each possible time-compression ratio. The time-compression ratios ranged from 40% of the original speed (fastest) to 200% increase from the original speed (slowest). As the participant sped up or slowed down the signal, Matlab would select the corresponding folder, which would switch the audio signal to the same relative position in a file with a slower or faster rate. A fading technique was used to ensure that no audible distortions occurred at the transition point between selections. Due to the sizable waveform files of each story, this technique was preferred, instead of time compressing or expanding the files in real-time during the adjustments. In a previous pilot study, DCT stories were time compressed in real-time, which caused noticeable clicks within the speech stimuli. In the present study, the R-SPIN sentences and the MRT words were processed immediately prior to presentation because of the small size of each waveform file. The same technique (PSOLA) used for the DCT stories was used to time compress the intelligibility stimuli. Intelligibility sentences and words were presented at the self-selected speeds chosen during the DCT stories within a given condition. Therefore, it was essential to time compress the sentences and words in real-time in order to match the time-compression ratios selected during the stories.

To further examine the impact of available processing time, the speech presentation rate of the DCT stories was manipulated using two different conditions: (1) passages with natural silences intact and (2) passages with silences removed. In the silences removed condition, any stimuli section longer than 200 ms and greater than 40 dB below the average level of the signal was considered a silent interval and was removed from the story. Stories were an average of 61 sec with silences and an average of 44.3 sec without silences. The difference between the self-selected time-compression ratio for stories with and without silences is expected to provide insight on whether the overall information rate of the signal or the articulation rate of the speech segments are the principal limiting factor in the self-selected preferred listening rate for comprehension.

Procedures

Experimental Procedures.

During the experimental conditions, participants were seated in a quiet room and heard speech material presented from an Audio Stream Input/Output (ASIO) player on a laptop computer. The speech stimuli were routed from the laptop to an RME Digiface Universal Serial Bus (USB) portable digital audio interface and delivered binaurally to the listener through Sennheiser HDA 200 headphones. Sennheiser HDA 200 headphones were chosen due to the excellent passive attenuation capabilities. All speech stimuli were presented at 80 dB SPL in the presence of continuous speech-shaped noise (SSN) fixed at +5 dB SNR. Presentation level was verified and calibrated using a Larson Davis Model 824 sound level meter connected to a flat-plate coupler. A positive SNR of 5 dB was selected following results reported by Smeds et al. (2015), which indicated real world speech-in-noise environments had SNRs near +5 dB. Therefore, a SNR of +5 dB was utilized in order to create a more realistic environment compared to presentation in quiet.

Two conditions were tested in the discourse comprehension portion of the experiment: (1) DCT stories with silences intact and (2) DCT stories with silences removed (see Table 2.1). Within each condition, participants were presented with two stories or trials. The order of conditions was randomized across participants using a block design. For example, participants would first be presented with two trials of the silences intact condition and then two trials of the silences removed conditions or vice versa. Participants were instructed to adjust the rate of the DCT passage to the fastest rate at which they felt that they could still understand and answer questions about the story. Each story began at the same original rate (indicated by a speed of 1). The scroll wheel on the computer mouse controlled the change in speed, moving the wheel towards the desk made the stimuli faster and turning the wheel in the opposite direction made the stimuli slower. The final self-selected time-compression ratio was calculated by taking an average of the presentation rate during the last 20% of the story in order to account for accidental or last-minute adjustments to the time-compression ratio. A practice story was provided to ensure that participants understood the task. Before the start of each story there was a set of spoken instructions that read, "Please start to adjust the speed now. Your goal is to play the stories as fast as you can while still being able to understand and answer questions about the story. You can also adjust the speed during the story. The story will begin after the beep." These instructions were provided before the start of the story to allow for speed adjustments to start before the story began. After the completion of each story, the participants were presented with eight comprehension questions, which yielded a percent correct performance for each story. There were four types of questions: main idea stated, main idea implied, detail stated and detail implied. Taken together, two outcome measures were collected from each story by assessing discourse comprehension performance in this manner: (1) the self-selected time-compression ratio and (2) the overall percent correct performance for the comprehension questions. The average self-selected speed for the stories in the silences intact condition and average self-selected speed for the stories in the silences removed condition were utilized in the intelligibility portion of the study.

After completing the discourse comprehension task in each speech presentation rate condition, participants completed the intelligibility measures. For example, a participant who heard two stories in the silences intact condition would then be presented with the intelligibility measures in the silences intact condition. Intelligibility sentences and words were presented at the self-selected speeds chosen during the DCT stories within a given condition. There were a total of six intelligibility conditions tested in this portion of the experiment: (1) 20 HP sentences; (2) 20 AP sentences; (3) 30 MRT words presented at the average self-selected speed from the DCT silences intact condition; and (4) 20 HP sentences; (5) 20 AP sentences; (6) 30 MRT words presented at the average self-selected speed from the DCT silences removed condition (see Table 2.1). Keywords were scored to achieve an overall percent correct score for HP and AP sentences. MRT words were presented in sets of three, and participants were asked to select each word they heard from a closed set of 6 choices for each target word. HP sentences, AP sentences and MRT words were selected in order to assess a wide range of intelligibility material from high context sentences to individual words.

Testing occurred during one session lasting approximately 2 hours at the University of Maryland, College Park. All participants were paid for their participation. The experiment was conducted with approval from the Institutional Review Board (IRB) at the University of Maryland, College Park.

Table 2.1.

Design of Experiment 1

Task	Speech Presentation Condition	Stimuli	Time Compression (TC) Ratio	Outcome Variable
Comprehension	Silences Intact	2 DCT Stories	Self-Selected	(1) Self-selected TC (2) Comprehension questions percent correct
	Silences Removed			
Intelligibility	Silences Intact	20 HP Sentences 20 AP Sentences 30 MRT Words	Average self-selected TC from the same speech presentation condition in the comprehension task	(1) Percent Correct
	Silences Removed			

Cognitive Measures.

Previous literature has emphasized working memory, speed of processing and attention as cognitive domains that decline as a result of aging and should influence discourse comprehension and speech intelligibility performance (Schneider et al., 2010). In addition, vocabulary knowledge has been shown to impact performance on difficult listening tasks for older adults (Schneider et al., 2016). Therefore, assessments were selected that would evaluate individual cognitive function and vocabulary knowledge. All cognitive and vocabulary measures were administered prior to experimental testing. Cognitive measures were presented in the same order for each subject: (1) Trail-Making Test; (2) Flanker; (3) Picture Vocabulary Test; and (4) RSPAN.

Working memory is the ability to manipulate and temporarily store incoming information (Baddeley & Hitch, 1974). Working memory capacity was assessed using the RSPAN (Danemann & Carpenter, 1980; Baddeley et al. 1985, Rönnberg, 1998). The RSPAN is thought to assess working memory because it requires individuals to initially process and manipulate the incoming information and store that information for later recall. The RSPAN was selected to measure verbal working memory capacity in order to control for hearing loss as a confounding variable. In addition, studies have demonstrated the relationship between the RSPAN and speech perception in noise (Zekveld et al., 2014). During the RSPAN, participants were asked to read sentences that were presented word-by-word and determine if the sentence was true or false. Participants responded "yes" if the sentence was true and "no" if the sentence was false. Participants were then instructed to remember the last word in each sentence to be recalled at the end of a set of sentences. As the test progressed, the number of sentences in each set increased from three to six. Typically, participants are given a SPAN score based on how many final words were recalled in the correct serial order. However, according to Freidman and Miyake (2005), a more accurate representation of an individual's working memory capacity is found when the total number of final words recalled correctly is calculated, regardless of the order. For example, if there were a total of 50 final words and 30 words were recalled correctly, regardless of the order, this would yield a score of 60% correct. This method of scoring was found to be more correlated with reading comprehension performance compared to a SPAN score (Freidman & Miyake, 2005). Scoring the total number of final words recalled, regardless of the order, leads to greater variability across participants and groups compared to a SPAN score which only yields a limited number of scoring options (e.g., 3, 3.5, 4, 4.5, etc) Therefore, the total words correct method of scoring was utilized in the current study in order to capture more variability in working memory capacity across individuals.

The Trail-Making test was utilized to measure speed of processing (Salthouse et al., 2000). Speed of processing is the rate at which an operation is performed in order to accomplish a given task. Salthouse (2011) reported that the Trail-Making test uniquely measures speed of processing because the task is correlated with other measures of processing speed, such as a Pattern Comparison task. The Trail-Making test requires participants to draw lines to connect circled numbers or letters as quickly as possible in ascending order. There are two forms: Form A contains only a sequence of numbers (1-2-3, etc) and Form B is alternating letters and numbers (1-A-2-B-3-C, etc). Salthouse (2011) reported that both forms showed strong relationships to other speed of processing assessments, suggesting that either form would be an appropriate outcome measure of speed of processing ability. In the present study, Form B was used in all analyses.

Selective attention and inhibition were assessed with the Flanker Task (Erikson & Erikson, 1974), which is administered via the National Institutes of Health (NIH) Cognitive ToolBox (Gershon et al., 2013). Attention is a multifaceted concept that includes orienting, selecting and focusing on presented stimuli. Flanker task is thought to capture some aspects of attention and inhibition abilities. The Flanker task requires individuals to focus on a set of arrows and indicate the direction of the middle arrow as quickly as possible. In some conditions the arrows are all pointing in the same direction (congruent) and sometimes the arrows are in the opposite direction (incongruent). Scoring is based on a combination of accuracy and reaction time. This combination score is then converted to a scale score with a mean of 100 and a SD of 15. Salthouse (2010) reported that it is appropriate to analyze the combination score of accuracy and reaction time for the Flanker task. Previous studies interested in the relationship between attention and the mechanisms responsible for understanding speech in noise have also used the Flanker combination score from the NIH Toolbox (Presacco et al., 2016). Lastly, linguistic competency was measured by assessing vocabulary knowledge using the NIH Toolbox Picture Vocabulary Test. Vocabulary knowledge is known to be preserved in older adults and may provide a benefit to older adults when listening tasks become difficult (Schneider et al., 2016). Participants were asked to select the picture that best represents the meaning of the given word.

Statistical Analysis

Data analysis incorporated correlations, multivariate repeated-measures analyses of variance (ANOVAs) and a linear mixed-effects regression (LMER) model. The first portion of the results section describes the unique self-selection paradigm. In order to determine if listeners were reliable when selecting the time-compression ratio, a bivariate correlation between the speed selected in trial 1 and trial 2 for each individual was performed for both speech presentation rate conditions. Next, the self-selected time-compression ratios were characterized in two ways: (1) raw time-compression ratio selections over the duration of the story and (2) the time-compression ratio difference from the ending selection over the duration of the story. Comprehension question accuracy scores were analyzed with an ANOVA with one within-subjects variable [speech presentation rate condition (two levels: silences intact and silences removed)] and one between-subjects variable [Group (three levels: YNH, ONH, and OHI)]. Based on pilot data, comprehension accuracy scores were expected to be similar (on average near 80% correct) for all groups and both speech presentation rate conditions. Lastly, a bivariate correlation between the average self-selected speed and the average comprehension question accuracy score for each individual was performed for both speech presentation rate conditions.

The LMER was performed to determine the effects of speech presentation rate condition (i.e., information rate and articulation rate), age, hearing loss, speech intelligibility and cognitive function on self-selected time-compression ratio. The LMER was performed using a maximal model structure with a backward selection method following Barr et al. (2013). The maximal model included all fixed effects, interactions and random effects of interest. The random effects structure included a random intercept of participant and a random slope of condition. The maximal model contained fixed main effects of speech presentation rate condition [two levels = silences intact (reference) and silences removed], group [three levels: YNH (reference), ONH, and OHI], speech intelligibility scores [two levels: HP (reference) and AP sentences], RSPAN score, Trail Making speed (Version B), Flanker Score and Vocabulary score. The assessments from the NIH Cognitive Toolbox (Flanker and Vocabulary Score) were input into the model as raw scores instead of age-corrected scores, because age was taken into account in the model. All cognitive tests (RSPAN, Trail Making, Flanker and Vocabulary score) were assessed for collinearity ($r \geq .7$) and standardized to z-scores before being entered into the model. No collinearity was present between cognitive measures. Speech intelligibility measures were also assessed for collinearity and converted to Rationalized Arcsine Units (RAU) prior to being added to the model. MRT performance was found to be highly correlated with HP and AP performance ($r \geq .7$) and was removed from the model in order to reduce collinearly in that predictor variable. Speech intelligibility contextual benefit (i.e., difference in percent correct performance for HP and AP sentences) was also strongly correlated with AP intelligibility performance ($r \geq .7$) in both speech presentation rate conditions and was not entered into the model.

The buildmer package (version 1.6) in R (Voeten, 2019) was utilized to implement the backwards elimination approach outlined by Barr et al. (2013). Buildmer reduces fixed and random effects terms in order to determine the final best fitting model for the data. Buildmer determines the order of the effects in the model and then a backward stepwise elimination is performed. Terms that were not significant ($p > .05$) were not included in the final reported model. In addition, significant interactions were further analyzed by re-leveling the reference variables involved in the interactions.

Results

Self-Selection Method

Two outcome measures were collected: (1) the self-selected time-compression ratio and (2) the overall percent correct performance for the comprehension questions. The use of a self-selected time-compression paradigm is relatively unique. Therefore, it is important to characterize the method of self-selection used in this study. First, a bivariate correlation was conducted between time compression ratios selected on trial 1 and on trial 2 for all groups, and revealed a significant positive correlation between these self-selected ratios for the two trials for both speech presentation conditions [Silences Intact ($r=.91$; $p<0.001$), Silences Removed ($r =.94$; $p<0.001$)] (See Figure 2.2). Participants selected consistent time-compression ratios over two repetitions for both speech presentation rate conditions. This result suggests that the self-selection method produced reliable results across conditions and participants.

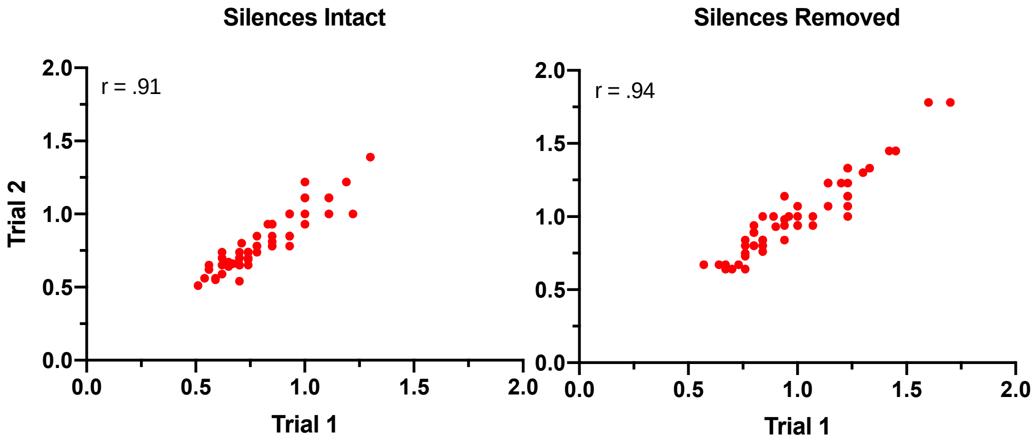

Figure 2.2. Correlation data between the self-selected time-compression ratio in trial 1 and trial 2 for both speech presentation conditions (Left panel: Silences Intact; Right panel: Silences Removed).

Next, the self-selected time-compression ratios over the duration of the story were quantified. The scroll wheel on the computer mouse controlled the adjustments, and the time-compression ratio at each time point throughout the stimulus file was recorded (i.e from 0 to 100 percent completion). The time-compression ratios over the duration of the story were characterized in two ways: (1) raw time-compression ratio selections for the duration of the story (Figure 2.3) and (2) the time-compression ratio difference from the ending selection over the duration of the story (Figure 2.4). The dashed lines in Figures 2.3 and 2.4 represent the end of the spoken instructions and the beginning of the discourse comprehension story. The averaged self-selected time-compression ratios for each group from 0 to 100 percent completion are shown in Figure 2.3. This figure shows that the majority of large adjustments to the time-compression ratio occurred during the instructions (i.e., before the dashed line). However, it appears that OHI listeners made more small adjustments throughout the story in the silences removed condition compared to the YNH and ONH groups.

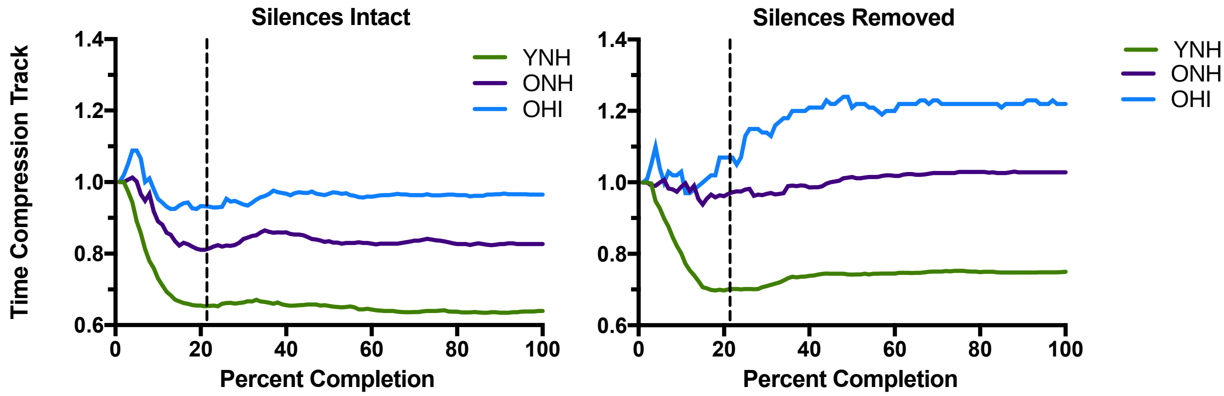

Figure 2.3. Average time-compression ratio selection (Y-axis) as a function of percent completion of the listening interval (instructions and story) for all groups (YNH, ONH and OHI) and both speech presentation rate conditions (silences intact and silences removed). The dashed line represents the end of the spoken instructions and the start of the discourse comprehension story. Note, 1.0 is original speech rate, .6 means that the speech is 60% of the original duration (shorter), and 1.2 is 120% of the original duration (longer).

Figure 2.4 shows the average difference in time-compression ratio (i.e., percent difference) over the course of the instructions and the story compared to the final time-compression ratio. Viewing the data in terms of the time-compression ratio difference from the ending selection over the duration of the story provides insight into how near listeners were to their ending speed at each time point throughout the instructions and story. On average, at the start of the story, YNH listeners were 7% (silences intact condition) and 8% (silences removed condition) away from the final self-selected time-compression ratio. At the start of the story, ONH listeners were 7% (silences intact condition) and 10% (silences removed condition) away from the final self-selected time-compression ratio. OHI listeners were 8% (silences intact condition) and 16% (silences removed condition) away from the final self-selected time-compression ratio at the start of the story. This suggests that at the start of the discourse comprehension story (signified by the dashed line in the figure), listeners were relatively close to their final time-compression ratio. However, OHI listeners were 8 to 12% further from their final self-selected speed at the start of the story compared to YNH and ONH groups.

An ANOVA was performed to determine if there were significant differences in the change in time-compression ratio at the start of the story from the final time-compression ratio across groups and speech presentation rate conditions. The analysis had one within-subjects variable [speech presentation rate condition (two levels: silences intact and silences removed)] and one between-subjects variable [Group (three levels: YNH, ONH, and OHI)]. The results revealed a significant main effect of speech presentation rate condition [$F(1,50)=14,62$, $p<.0001$, $\eta p^2 = .226$], a significant main effect of group [$F(2,50)=4.59$, $p<.05$, $\eta p^2 = .155$], and a significant interaction between group and presentation rate condition [$F(2,50)=5.77$, $p<.05$, $\eta p^2 = .188$]). Post hoc paired-comparison t-tests with Bonferroni corrections revealed no significant differences in performance between speech rate conditions for the YNH and ONH groups ($p>.05$) and no significant differences between the YNH and ONH groups when compared within the same speech presentation rate condition ($p>.05$). However, the OHI group showed significant differences across conditions. Specifically, this group was significantly closer to their final self-selected speed in the silences intact condition compared to the silences removed ($p<.05$). In addition, the OHI group was only significantly different from the YNH and ONH groups in the silences removed condition ($p<.05$). Taken together, it appears that the OHI listeners had more difficulty settling on an ideal time-compression ratio in the silences removed condition. Although OHI listeners were further from their final time-compression ratio at the start of the story compared to YNH and ONH listeners in the silences removed condition, it is important to consider that listeners were consistent in their speed selections across trials and their comprehension accuracy performance was not negatively impacted in the silences removed condition.

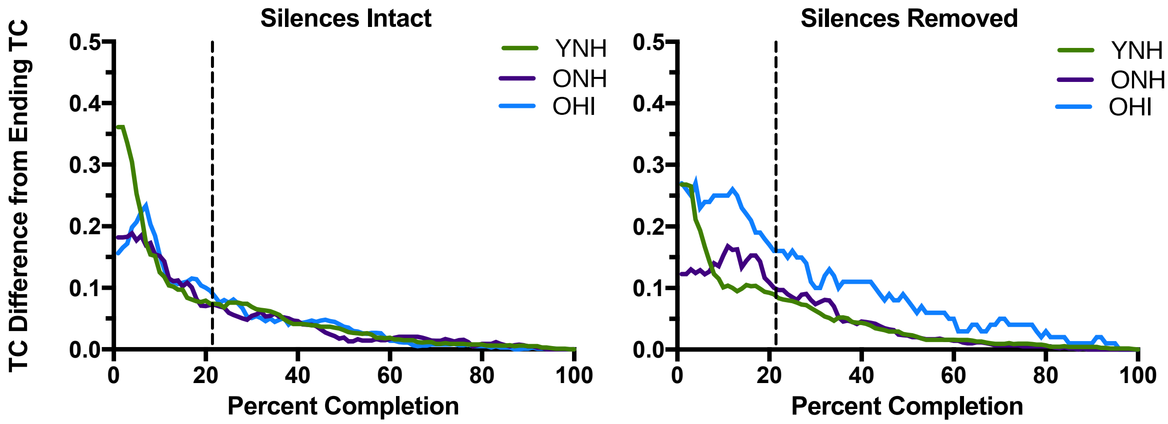

Figure 2.4. Average time-compression difference from the ending time-compression ratio (Y-axis) as a function of percent completion of the listening interval (instructions and story) for all groups (YNH, ONH and OHI) and both speech presentation rate conditions (silences intact and silences removed). The dashed line represents the end of the spoken instructions and the start of the discourse comprehension story. (TC = time compression)

Comprehension Accuracy Performance

Figure 2.5 shows percent correct comprehension performance for the questions that followed each story for young normal hearing listeners and older listeners with and without hearing loss. An ANOVA was conducted to determine the impact of condition and group on comprehension accuracy. The dependent variable was the comprehension accuracy score. The ANOVA had one within-subjects variable [speech presentation rate condition (two levels: silences intact and silences removed)] and one between-subjects variable [Group (three levels: YNH, ONH, and OHI)]. The ANOVA revealed a significant main effect of speech presentation rate condition [$F(1,51)=5.327$, $p<.05$, $\eta_p^2 = .095$]. However, the ANOVA revealed that neither the main effect of group [$F(2,51)=2.28$, $p>.05$, $\eta_p^2 = .082$] nor the interaction were significant [$F(2,51)=.075$,

p>.05, $\eta p^2 = .003$] .The main effect of speech presentation rate condition indicates that comprehension accuracy performance was better in the silences intact speech presentation rate condition compared to the silence removed condition. Overall, participants in all groups were able to achieve an average comprehension accuracy score of 80% in the silences removed condition and an average of 84% in the silences intact condition. It is conceivable that the general lack of variance in comprehension accuracy scores made it possible for an average difference of 4% between conditions to reach significance; nonetheless, this difference is not clinically relevant (Table 2.2).

Table 2.2.

Means and standard deviations for comprehension accuracy performance at individually selected time-compression ratios.

	Speech Presentation Rate Condition			
	Silences Intact		Silences Removed	
Group	*M*	*SD*	*M*	*SD*
YNH	0.82	0.1	0.78	0.11
ONH	0.87	0.07	0.83	0.08
OHI	0.84	0.08	0.79	0.1
Average	**0.84**		**0.80**	

Note. M and *SD* represent mean and standard deviation, respectively.

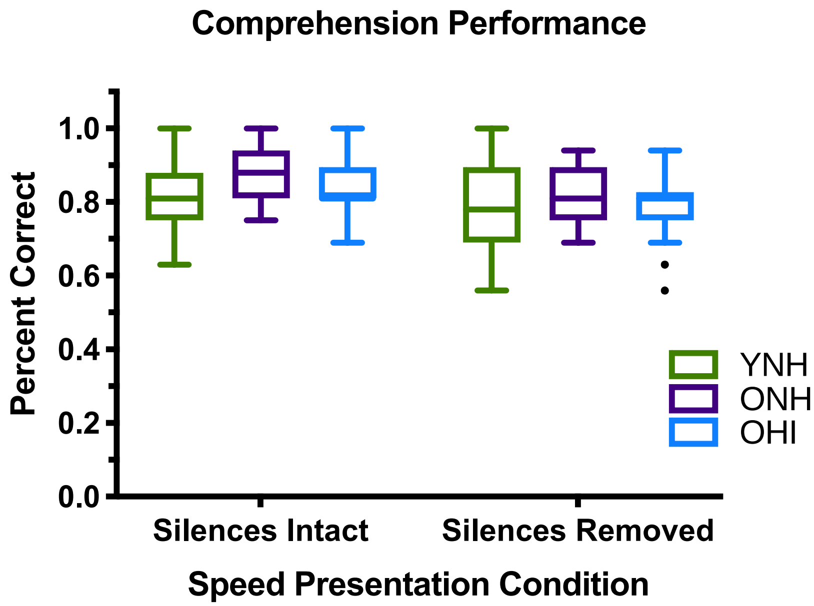

Figure 2.5. Comprehension accuracy performance for each group and both speech presentation rate conditions for eight yes and no questions presented directly following the completion of each story. The box and whisker plots were created using the Tukey method, which calculates an interquartile range (i.e., the difference between the 25th and 75th percentiles). Any value greater or less than 1.5 times the interquartile range is plotted as individual outlier points.

Another approach to analyzing the comprehension accuracy performance is to correlate it with the self-selected time-compression ratios that yielded a given comprehension score. It is clear from Figure 2.5 that comprehension accuracy performance was relatively constant for all groups and both speech presentation rate conditions. However, Figure 2.6 indicates that the self-selected time-compression ratios to achieve these scores varied widely between the groups. A bivariate correlation, including all groups, revealed no correlation between the self-selected time-compression ratio and comprehension accuracy performance [Silences Intact (r=.12); Silences Removed (r =-.05)]. These results suggest that individual self-selected time-compression ratios produced similar comprehension accuracy performance regardless of age and hearing acuity (Figure 2.5). Additionally, these findings verify that the self-selected time-compression ratio enabled the majority of listeners to achieve a consistent comprehension accuracy score. Therefore, the variable of interest is the self-selected time-compression ratio which produced the equivalent percent correct performance.

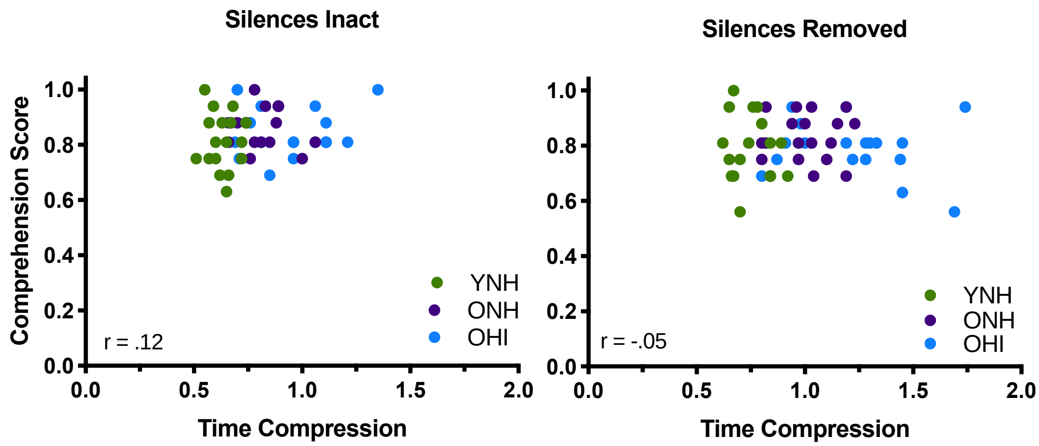

Figure 2.6. Bivariate correlation between self-selected time-compression ratio and comprehension accuracy performance for both speech presentation rate conditions.

Effects of Age and Hearing Loss

The results of the LMER are shown in Table 2.3. The dependent variable was the self-selected time-compression ratio. The maximal model contained fixed main effects of speech presentation rate condition [two levels = silences intact (reference) and silences removed], group [three levels: YNH (reference), ONH, and OHI], speech intelligibility scores [two levels: HP (reference) and AP sentences], RSPAN score, Trail Making speed (Version B), Flanker Score and Vocabulary score. Significant fixed effects were group (p<0.001), condition (p<0.001) RSPAN score (p<0.05) and Trail Making speed (p<0.005). Group and condition were also involved in a significant interaction [OHI x condition (p<0.001) and ONH x condition (p<0.005)]. None of the higher-level interactions of fixed-effects (i.e., three-way, four-way) were significant. The reference variables were re-leveled in order to compare performance between the two older groups in the silences intact condition, examine the relationship between groups in the silences removed condition, and further analyze the significant interactions. In this section, the results regarding the impact of condition and group will be discussed. LMER results involving cognitive function and speech intelligibility performance will be discussed in detail in the forthcoming results section.

Table 2.3.

Final LMER model for self-selected time-compression ratio. Bolded rows indicate significant terms ($p \leq 0.05$).

Fixed Effects	Coefficient	SE	t	p
Intercept	0.70	0.03	23.03	**<0.001**
Group [OHI]	0.21	0.05	4.72	**<0.001**
Group [ONH]	0.12	0.04	2.82	**0.005**
Condition [SR]	0.12	0.02	5.66	**<0.001**
RSPAN	-0.04	0.02	-2.01	**0.045**
Trail Making (Version B)	0.05	0.02	2.81	**0.005**
Interactions				
Group[OHI] *Condition [SR]	0.14	0.03	4.64	**<0.001**
Group[ONH] *Condition[SR]	0.07	0.03	2.24	**0.025**
Random Effects	**Variance**			**SD**
Subject (intercept)	0.012			0.11
Subject/Mode	0.007			0.083
Residual	0.002			0.048

Figure 2.7 shows the self-selected time-compression ratios for all groups and both speech presentation rate conditions. Overall, given additional time to process the time-compressed speech, listeners in all groups were able to select a faster listening rate in the silences intact condition compared to the silences removed condition. This suggests that

the constraint on overall information rate is the principal source of difficulty when listening to rapid speed during a discourse comprehension task. The main effect of group was evaluated for each speech presentation rate condition. The LMER revealed that YNH listeners were able to select faster rates in both speech presentation rate conditions compared to ONH and OHI groups, indicating an overall advantage for YNH listeners. Significant differences were also seen between ONH and OHI groups for both speech presentation rate conditions, suggesting OHI listeners preferred slower listening rates in order to reach equivalent comprehension accuracy performance compared to ONH listeners.

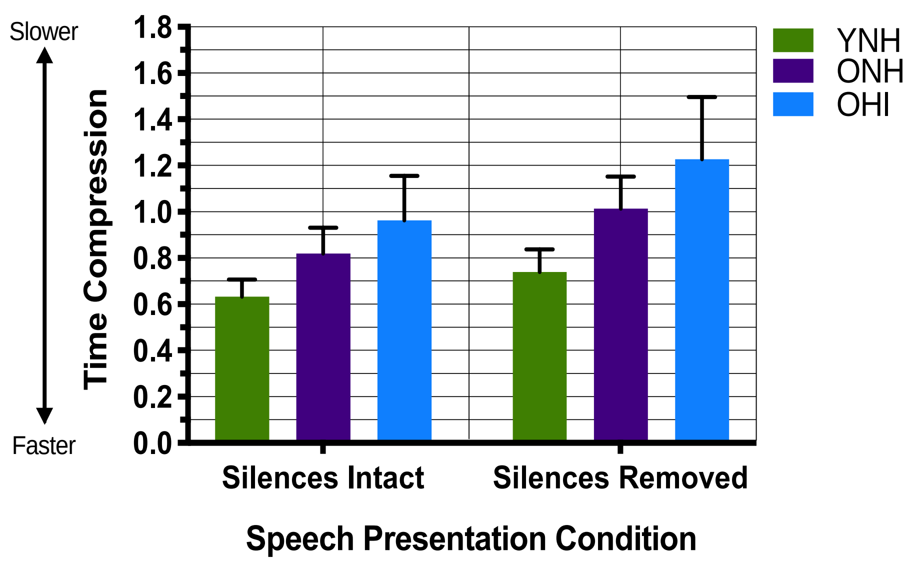

Figure 2.7. Self-selected time-compression ratio for all groups (YNH, ONH, OHI) for both speech presentation rate conditions (silences intact and silences removed). Note, 1.0 is original speech rate, .6 means that the speech is 60% of the original duration (shorter), and 1.2 is 120% of the original duration (longer). Error bars represent standard deviation.

The significant interaction between group and speech presentation rate condition was evaluated further during the re-leveling process. The source of the interaction appears to be a greater difference in the condition effect for the two older groups compared to the younger group.   Greater differences between the self-selected time-compression ratios in the silences intact condition and the silences removed condition were found for ONH [β = .18, SE = .02, t = 8.81, p < .0001] and OHI [β = 0.26, SE = .02, t = 12.22, , p <.0001] groups compared to YNH listeners [β = 0.12, SE = 0.02, t = 5.65, p <.0001]. This result suggests that both older groups were more impacted by the presence of silences in the silences intact condition compared to younger listeners. However, PSOLA, a uniform method of time compression or time expansion, was used. The PSOLA algorithm equally time compresses and/or time expands both the speech segments and periods of silence. Older adults selected a slower time-compression ratio and subsequently had longer periods of silences. Therefore, it is possible that older adults only showed a larger difference between the two conditions compared to the younger listeners because they had longer periods of silence.

In order to account for the length of silences in the silences intact condition, self-selected time-compression ratios were converted to syllables per second. Figure 2.8 shows self-selected time-compression ratios converted into syllables per second for all groups for both speech presentation rate conditions. Syllables per second were calculated by dividing the total number of syllables per story by the duration of the story. The average duration of the stories in the silences intact condition was 61 seconds and the average duration of the stories in the silences removed conditions was 44.3 seconds. Each participant's individual time-compression ratio was taken into account by multiplying the original story duration by the self-selected time-compression ratio to get the final individual story durations. Then, the average number of syllables (260.7) was divided by the final story duration (in sec) for each participant to get the syllable/sec rate. For example, a time-compression ratio of 70% in the silences intact condition yielded a final story duration of 42.7 seconds and a syllable rate of 6.1 syllables/sec. A time-compression ratio of 70% in the silences removed condition yielded a final story duration of 31 seconds and a syllable rate of 8.4 syllables/sec, a faster syllable rate than in the silences intact condition. However, this method of calculating syllable rate does not account for the length of silences in the silences intact condition.

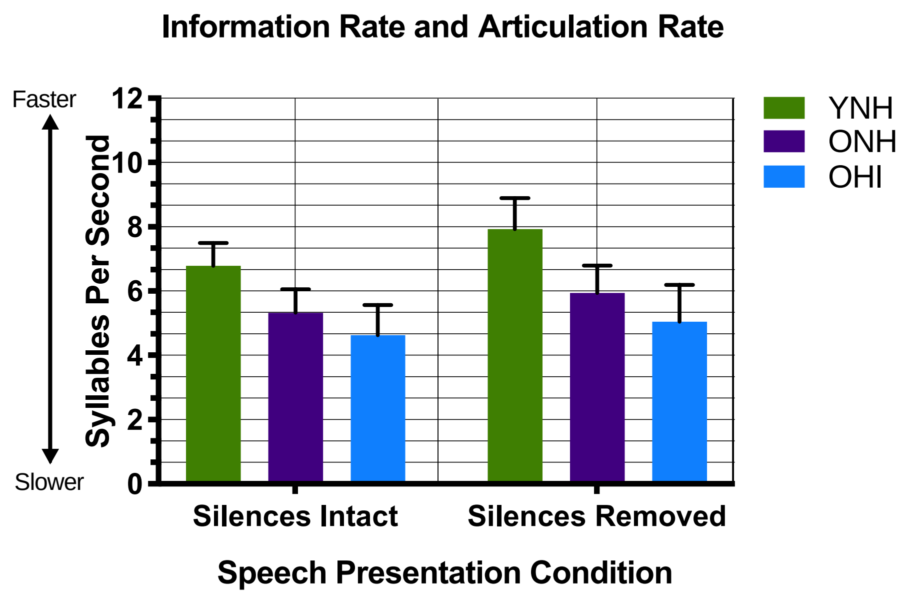

Figure 2.8. Self-selected time-compression ratio shown in syllables per second for all groups (YNH, ONH, OHI) for both speech presentation rate conditions (silences intact and silences removed). A higher number of syllables per second indicates a faster self-selected rate. Note this figure does not take into account the length of silences in the silences intact condition. It is simply a representation of the first step in accounting for the length of silences in the silences intact condition. The silences intact condition still represents the information rate (syllables per sec including silences) in this figure. Error bars represent one standard deviation.

In Figure 2.8, the calculation of the total story duration in the silences intact condition included both the duration of the spoken syllables and the duration of silent periods. Therefore, further analysis was undertaken to remove the silent periods in the silences intact condition and quantify only the duration of the spoken syllables that comprised the speech passage. Figure 2.9 shows the syllable rate after the silences were removed from the silences intact condition. To calculate the final duration and syllables/sec in the silences intact condition, the time-compression ratio was multiplied by 44.3 sec (duration of the story with silences removed) instead of the previous duration of 61 sec. Then, the average number of syllables (260.7) was divided by the re-calculated story duration for each participant to derive a syllable rate corresponding to the speech "on time." Note the syllables per second in the silences removed condition remain the same in both Figure 2.8 and 2.9. Now, the speech presentation rate is quantified only by the speech "on time" for both conditions and conveys the articulation rate of the signal (Jacewicz, et al., 2009). Figure 2.9 is a direct comparison between the articulation rates in the silences intact condition and the silences removed condition.

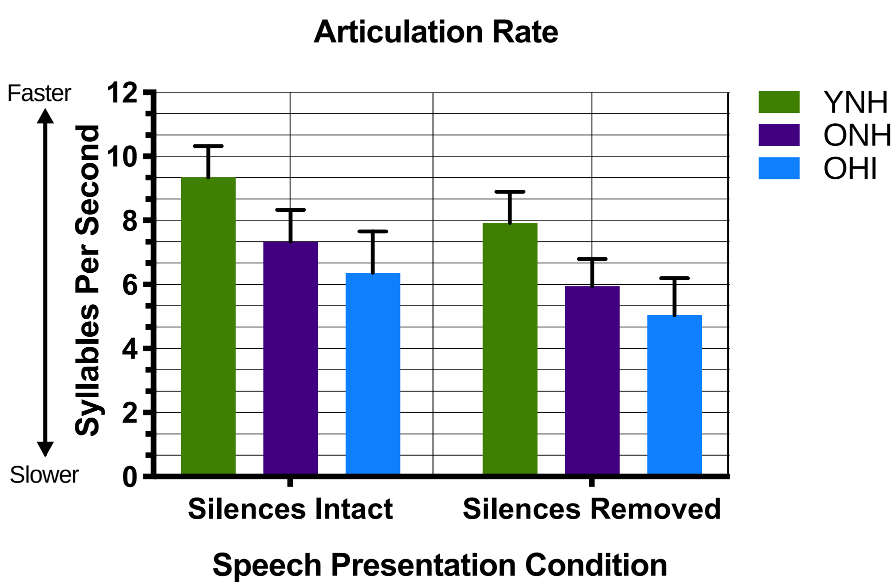

Figure 2.9. Self-selected listening rate shown in syllables per second for only the speech on time (articulation rate) for all groups and both speech presentation rate conditions. A higher number of syllables per second indicates a faster self-selected rate. Error bars represent standard deviation.

One additional caveat remained before the syllables per second data could be analyzed. The relationship between syllables per second for time-compressed speech and syllables per second for time-expanded speech is not a linear scale (Quene, 2007). Quene (2007) aimed to determine the just noticeable difference (JND) for speaking rate for time-compressed and time-expanded speech samples without silences (Figure 2.10). During a two-alternative forced choice task and a same-different task, Quene found smaller d' values or a larger JND for time-expanded speech compared to time-compressed speech. Thus, listeners needed a larger change from the original speech signal when speech was time expanded compared to time compressed in order to notice a difference. The difference in slope for the JND responses between time-compressed and time-expanded speech suggests that syllables per second cannot be analyzed on a linear scale.

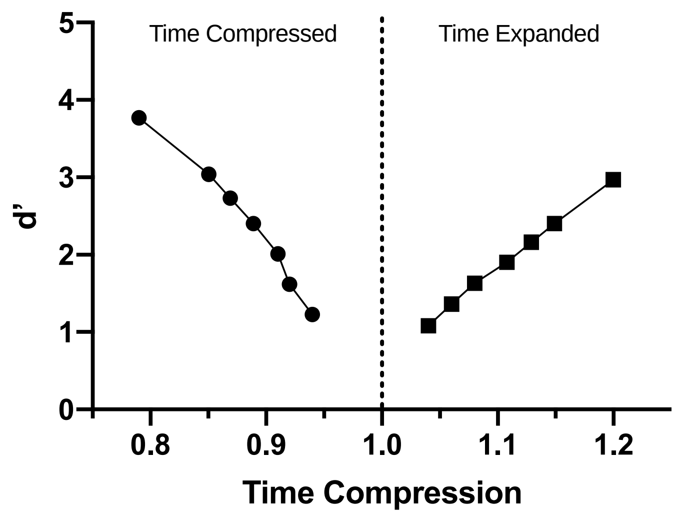

Figure 2.10. Adapted from Figure 1 in Quene (2007). D' (y-axis) is shown as a function of time-compression ratio (x-axis) during a same-different task. Similar to the current study, a lower number on the x-axis signifies a faster rate, a value of 1 represents the original speech rate and a number above 1 signifies a time-expanded signal. Larger d' values indicate smaller JNDs for stimuli in the time-compressed range.

In the current study, many older listeners, particularly those in the OHI group, selected a time-expanded signal in the silences removed condition. Results from Quene (2007) suggest that the perception of syllables per second for a time-compressed signal and the perception of syllables per second for a time-expanded signal are not on a linear scale. In other words, a single unit change in time-compression ratio has a larger impact on performance for a time-expanded signal than for a fast time-compressed signal. Therefore, results from the current study in syllables per second (Figure 2.9) were converted to a logarithmic scale before analysis. The same LMER maximal model described above was performed with the new dependent variable (i.e., syllable per second rate on a logarithmic scale). Results are shown in Table 2.4 and Figure 2.11. The purpose of this LMER was to analyze the interaction between group and condition after accounting for the length of silences in the silences intact condition and the non-linear scale of syllables per second.

Results show significant main effects of group (p<0.001) and condition (p<0.001). Group and condition were involved in a significant interaction [OHI x condition (p<0.05)]. The reference variables were re-leveled in order to further analyze the significant interaction. Significantly greater differences between the self-selected syllables per second in the silences intact condition and the silences removed condition were found for OHI listeners only ([β = -.102 SE = .008, t = -12.53, p < .0001]; ONH [β = -1.3, SE = .13, t = -10.37, p = >.05]; YNH [β = -.07, SE = 0.008, t = -9.04, p >.05]). This result suggests that, after accounting for the periods of silence, OHI listeners were still more impacted by the presence of silences compared to younger listeners. However, these results show that once the periods of silence were accounted for, ONH listeners were impacted to the same extent by the presence of silences compared to YNH listeners. Taken together, these results imply that self-selected listening rate is primarily modulated by available processing time or the information rate of the entire signal for all listeners, but to a greater extent for the OHI listeners.

Table 2.4.

Final LMER model for syllables per second on a logarithmic scale. Bolded rows indicate significant terms (p<0.05). Reference = YNH and Silences Intact.

Fixed Effects	Coefficient	SE	t	p
Intercept	0.93	0.01	59.74	**<0.001**
Group [OHI]	-.11	0.02	-4.86	**<0.001**
Group [ONH]	-.06	0.02	-3.13	**0.002**
Condition [SR]	-.07	0.00	-9.04	**<0.001**
RSPAN	.02	0.01	2.53	**0.01**
Trail Making (Version B)	-.02	0.01	-2.1	**0.04**
Interactions				
Group [OHI]* Condition[SR]	-.02	0.01	-2.470	**.016**
Group [ONH]*Condition[SR]	-.01	0.01	-1.42	.15
Random Effects	**Variance**			**SD**
Subject (intercept)	0.002			0.05
Subject/Mode	0.0004			0.02
Residual	0.0007			0.02

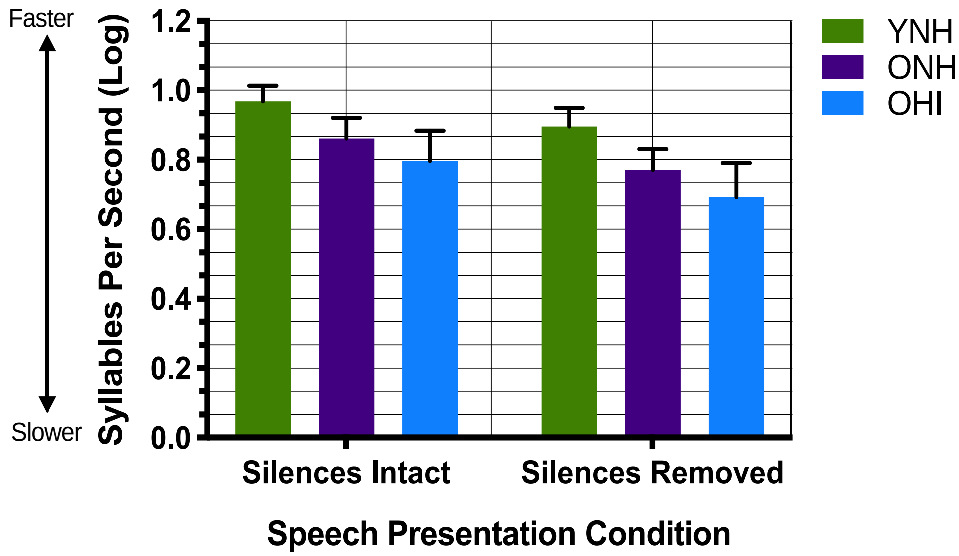

Figure 2.11. Self-selected listening rate shown in syllables per second for only the speech on time (articulation rate) plotted on a logarithmic scale for all groups for both speech presentation rate conditions (silences intact and silences removed). A higher number on logarithmic scale indicates a faster rate. Error bars represent one standard deviation.

Impact of Speech Intelligibility and Cognitive Function

The LMER model reported in Table 2.3 above also examined the effects of speech intelligibility and performance on several cognitive measures on self-selected time compression ratio. Speech intelligibility performance for the HP and AP sentences were tested at the average self-selected time-compression ratio for each condition. Measuring speech intelligibility at the self-selected time-compression ratios chosen during the discourse comprehension task was intended to provide insight into whether or not listeners were using the intelligibility of the story as a criterion for their time-compression ratio selection. Results of the LMER showed that HP and AP speech intelligibility scores were not significant predictors of self-selected time-compression ratios during a discourse comprehension task (Table 2.3). Scatter plots of the self-selected time-compression ratios and speech intelligibility scores (Figures 2.12 and 2.13) are shown to assist in the visualization of the relationships reflected in the LMER. Figures 2.12 and 2.13 show no relationship between speech intelligibility scores and self-selected time-compression ratios.

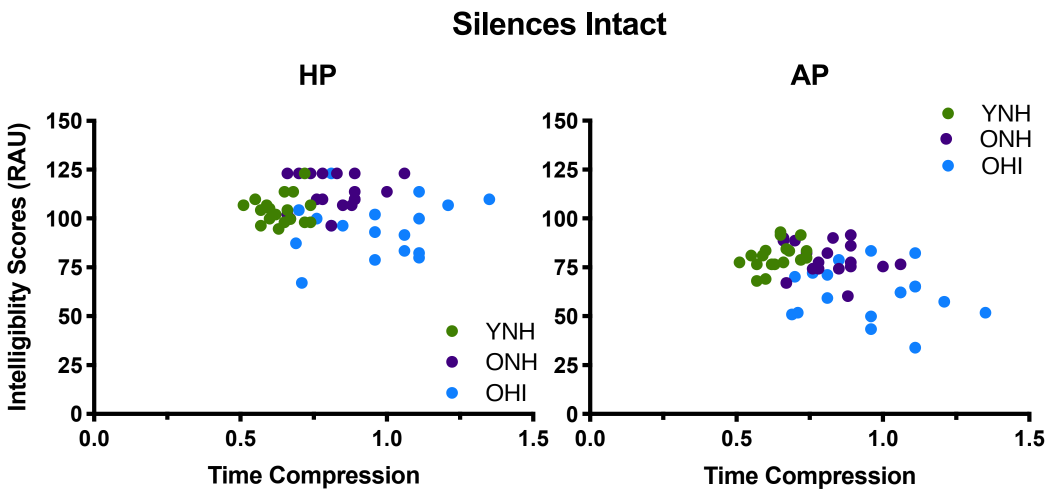

Figure 2.12. Scatterplots showing the spread of speech intelligibility scores (y-axis) and self-selected time-compression ratios (x-axis) for the three listener groups in the silences intact condition for HP and AP sentences. (HP = High Probability sentences; AP = Anomalous Probability sentences).

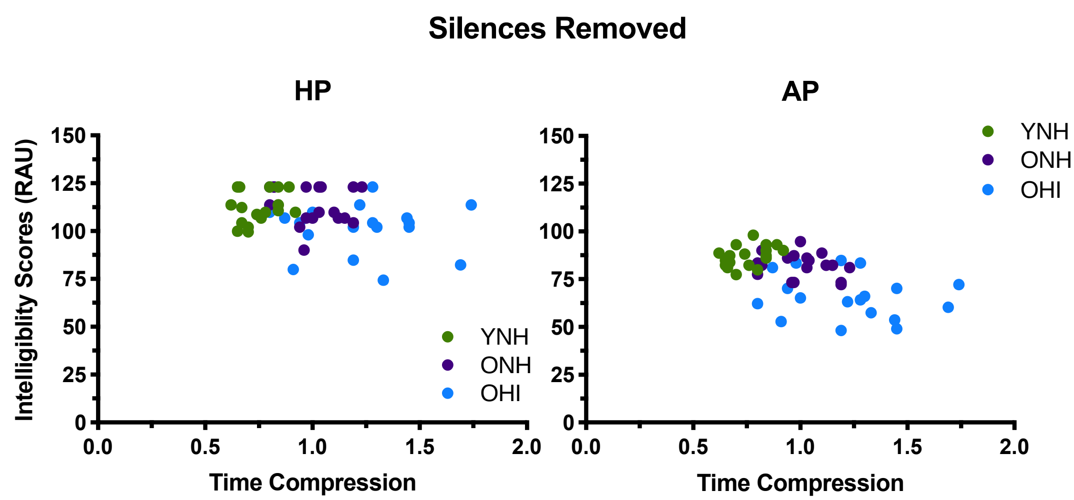

Figure 2.13. Scatterplots showing the spread of speech intelligibility scores (y-axis) and self-selected time-compression ratios (x-axis) for the three listener groups in the silences removed condition for HP and AP sentences. (HP = High Probability sentences; AP = Anomalous Probability sentences).

Cognitive function in the domains of processing speed (Trail Making task), working memory (RSPAN), and attention (Flanker Task) were evaluated to determine the cognitive domains that may predict self-selected time-compression ratios during a discourse comprehension task. Vocabulary knowledge (Picture Vocabulary Test) was also evaluated as a predictor of self-selected time-compression ratio. Note, all measures were assessed for collinearity ($r \geq .7$) and standardized to z-scores before being entered into the model. The results of the LMER reported above in Table 2.3 showed that the RSPAN and Trail Making scores (Version B) were significant predictors of self-selected time-compression ratio during the discourse comprehension task. No significant interactions were found in the LMER model between listener group and either RSPAN or Trail Making. This finding suggests that working memory capacity and speed of processing influenced the ability to process rapid speech during a discourse comprehension task similarly for all three groups and both speech presentation rate conditions. Scatter plots of the self-selected time-compression ratios and RSPAN scores (Figure 2.14) and self-selected time-compression ratios and Trail Making speeds (Figure 2.15) are shown to assist in the visualization of the relationships reflected in the LMER. Figure 2.14 shows a significant relationship between working memory capacity (RSPAN) and self-selected time-compression ratio. Figure 2.15 shows a significant relationship between processing speed (Trail Masking task) and self-selected time-compression ratio.

Lastly, ANOVAs and appropriate post hoc comparisons were performed in order to determine the differences in performance for all groups on each cognitive measure. For each cognitive measure (RSPAN, Trail Making, and Flanker), a separate one-way

ANOVA and post hoc multiple comparisons with Bonferroni corrections were conducted to evaluate group differences in each cognitive domain. The ANOVAs revealed a main effect of listener group for each cognitive measure. Post hoc multiple comparisons revealed significant differences between the YNH group and both older groups for all cognitive assessments ($p<.001$). YNH listeners showed higher RSPAN scores, faster Trail making speeds and higher Flanker scores compared to the ONH and OHI groups. Post hoc multiple comparisons revealed no significant differences in RSPAN score, Trail Making speed or Flanker score ($p>.05$) between ONH and OHI groups. This result suggests that significant differences between self-selected time-compression ratios for the ONH and OHI groups cannot be specially attributed to disparities in cognitive function between these two groups. Lastly, ANOVA revealed no significant differences in Vocabulary score ($p>.05$) between all three groups.

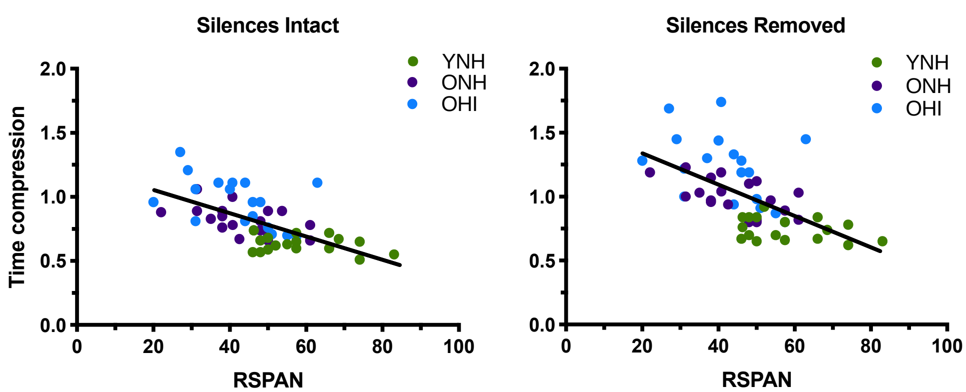

Figure 2.14. Scatterplots showing the spread of RSPAN (x-axis) and self-selected time-compression ratio (y-axis) for the three listener groups and both speech presentation ratio conditions with a regression line inserted.

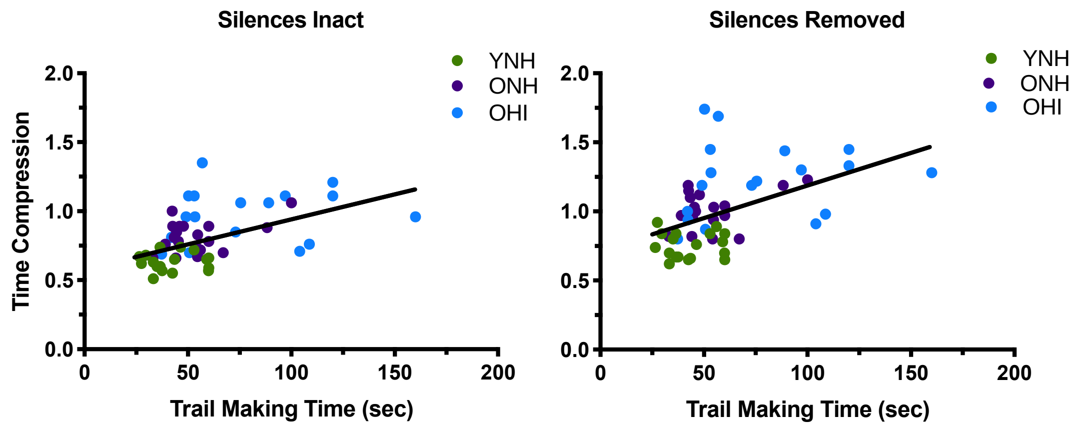

Figure 2.15. Scatterplots showing the spread of Trail Making time (x-axis) and self-selected time-compression ratio (y-axis) for the three listener groups and both speech presentation ratio conditions with a regression line inserted.

Discussion

The first goal of this study was to determine the effects of speech presentation rate condition (silences intact vs. silences removed), age, and hearing loss on self-selected time-compression ratio during a discourse comprehension task for young normal hearing adults and older adults with and without hearing loss. The results revealed an effect of age and hearing loss on self-selected time-compression ratio for both speech presentation rate conditions. The results also showed that self-selected time-compression ratio was generally modulated by information rate, especially for OHI listeners. The second goal of the study was to examine the relationship between speech intelligibility performance and self-selected time-compression ratio across the different listener groups. The results revealed that speech intelligibility performance at the self-selected time-compression ratio for adequate comprehension was not a significant predictor of the self-selected time-compression ratios of individual listeners. The final aim of this study was to determine which cognitive domain (i.e., working memory, selective attention or speed of processing) was predictive of self-selected time-compression ratio during a discourse comprehension task. Results showed that processing speed and working memory were significant predictors of self-selected time-compression ratio.

Effects of Age and Hearing Loss

The first aim of this study was to examine the effect of age and hearing loss on the self-selected time-compression ratio during a discourse comprehension task. It was hypothesized that older adults with and without hearing loss would select slower listening rates in both conditions compared to younger adults. LMER results showed that older adults with and without hearing loss required a slower listening rate in both conditions to reach 80% performance for comprehension accuracy compared to younger listeners. This finding is consistent with previous discourse comprehension research that reported older adults had more difficulty answering comprehension questions when the passage was time-compressed compared to younger adults (Wingfield et al., 1999; Wingfield & Ducharme, 1999). It was also hypothesized that older adults with hearing loss would select slower listening rates compared to older adults with normal hearing. This hypothesis was based on previous time-compression speech intelligibility and sentence comprehension studies that found age effects were independent of hearing loss effects on intelligibility for time-compressed speech (Gordon-Salant & Fitzgibbons, 1993; 2001; Wingfield et al., 2006). In the current study, older adults with hearing loss selected a slower listening rate during the discourse comprehension task compared to older normal hearing listeners. This indicates that OHI listeners are at a disadvantage compared to the ONH group when listening to rapid speech during a discourse comprehension task. This difference between older normal hearing adults and older hearing-impaired adults has not been measured in previous discourse comprehension studies. Thus, the current findings provide evidence that the negative effects of hearing loss on time-compressed speech intelligibility and sentence comprehension tasks may generalize to discourse comprehension performance. Taken together, these results may have real-world implications because they suggest that older adults, especially with hearing loss, require a slower listening rate compared to younger listeners in order to perform similarly on the comprehension accuracy task. Real-world conversation unfolds quickly, listeners must identify the spoken message, interpret the message and encode the message for later retrieval. This process must occur rapidly in order for discourse comprehension to be successful. However, listeners who prefer a slow listening rate may have difficulty participating in meaningful communication.

Effect of Speech Presentation Rate Condition

It was hypothesized that self-selected listening rate would be faster in the silences intact condition compared to the silences removed condition for all groups. LMER results showed that all groups selected a faster time-compression ratio in the silences intact condition compared to the silences removed condition. This result suggests that given additional time to process a spoken passage, listeners were able to select a faster listening rate, a finding that supports the cognitive hypothesis. This hypothesis states that listeners are limited by processing speed and/or the overall information rate of the signal and are able to benefit (i.e., select a faster listening rate) from the presence of silent periods in the silences intact condition. In contrast, the perceptual hypothesis (not supported by the current findings) states that listeners are limited by the distortion of the signal or the articulation rate. This hypothesis predicts that listeners should select similar time-compression ratios in both presentation rate conditions, regardless of the extra processing time provided by the silences and given the minimal acoustic distortions associated with PSOLA.

A significant interaction between the effects of speech condition and listener group was also observed, which is attributed to differences in the magnitude of the condition effect across the three groups. Essentially, the older groups showed a larger difference between self-selected time-compression ratio in the silences intact and silences removed conditions compared to the YNH listeners. This result was interpreted to reflect that older adults were more impacted by the presence or absence of extra processing time (i.e., silent periods) compared to the younger listeners. However, this result could simply be attributed to the fact that older listeners selected slower time-compression ratios. The PSOLA algorithm uniformly compresses and/or expands both the speech segments and periods of silence. Thus, the slower the self-selected time-compression ratio, the longer the periods of silence. Consequently, it is possible that older adults showed a larger difference between the two conditions compared to the younger listeners because they had longer periods of silence in the silences intact condition.

In order to account for the duration of silences in the silences intact condition and further understand the source of the significant interaction, the self-selected time-compression ratios were converted into syllables per second and transformed to a logarithmic scale. Data were plotted on a logarithmic scale, because the relationship between the syllables per second rate for time-compressed speech and the syllables per second for time-expanded speech is non-linear (Quene, 2007). In the current study, the majority of OHI listeners time expanded the signal in the silences removed condition in order to perform similarly on the comprehension accuracy questions. After controlling for the duration of silences and transforming the data to a logarithmic scale, the significant interaction between group and condition was attributed to the OHI group. Specifically, the OHI group showed a greater condition effect compared to the ONH and YNH listeners. These results imply that self-selected listening rate is influenced by available processing time or overall information rate for OHI listeners. It is possible that individuals with hearing loss are at a disadvantage in the sensory stage of bottom-up processing compared to normal hearing listeners. Therefore, OHI listeners likely compensate by relying on top-down cognitive resources. Previous literature suggests that when bottom-up sensory processing is compromised, additional top-down systems compensate in order to complete a given task (Piquado et al., 2012; Reuter-Lorenz & Cappell, 2008; Rönnberg et al., 2013)

The result that OHI listeners were the most impacted by the presence of silences supports results reported for younger listeners with hearing loss by Piquado et al. (2012). Their results showed that listeners with hearing loss recalled passages more poorly than normal-hearing listeners. However, when silences were placed at syntactic boundaries and participants could self-pace the passage, participants with hearing loss improved to the level of discourse comprehension performance seen in normal hearing listeners. Thus, the introduction of increased processing time allowed individuals with hearing loss to perform similarly to the normal hearing group. The authors proposed that hearing loss may slow the processing of incoming bottom-up acoustic information, leading to an increase in the top-down resources required for successful discourse comprehension by individuals with hearing loss. However, it is important to note that cognitive function was not measured directly, and older adults with and without hearing loss were not assessed, which limits comparisons between the results of the current study and those of Piquado et al. (2012).

Although results from Piquado and colleagues support the results found in the current study, it was hypothesized, based on results from Wingfield et al. (1999), that younger adults would show the greatest difference between the two conditions compared to the older groups. Results from the current study showed that younger adults and older adults with normal hearing were impacted to the same extent by the presence or absence of silences. Wingfield et al. (1999) presented YNH and ONH listeners with normal-rate passages and time-compressed passages. The time-compression conditions included: (1) 68% time compression and (2) 68% time-compressed passages that were restored to their original length with the insertion of silent periods. Wingfield et al. (1999) used an overall time-compression method that equally compressed the speech signal and the natural pauses, which is similar to the method used in the current study (i.e., PSOLA). Results for younger adults showed that the introduction of silences restored discourse comprehension performance to performance in the normal rate condition. This finding supports the notion that processing speed or the ability to process the overall information rate of the signal was causing the decline in discourse comprehension performance, for younger listeners with normal hearing. In the Wingfield et al. (1999) study the insertion of silences improved performance for the older adults, however, performance did not return to the same level observed in the normal-rate condition. Wingfield et al. (1999) suggested that this result implies that older adults with normal hearing had constraints on processing speed (i.e., information rate) and difficulty from the reduced acoustic information introduced by time-compression (i.e., articulation rate).

The results of the current study are somewhat contradictory to those of Wingfield et al. (1999), which found that YNH listeners were more impacted by the presence of silences compared to ONH listeners. Results of the current study show that YNH and ONH groups were impacted to the same extent by the presence or absence of silences, and that both normal hearing groups were less impacted compared to the OHI group. It is difficult to directly compare the results in the present study to results in the Wingfield et al. (1999) study due to the differences in testing methods. In addition, Wingfield et al. (1999) measured changes in percent correct performance, while the current study measured differences in self-selected time-compression ratios at a fixed level of comprehension performance. In addition, Wingfield and colleagues introduced artificial silent periods at clause and phrase boundaries in order to restore the length of the passage. In the current study, only natural silent periods were present in the silences intact condition. Thus, it is possible that the presence of natural silences (the current study) compared to synthetic silent periods (Wingfield et al., 1999) differentially impacted the results of both studies. Lastly, Wingfield et al. (1999) did not measure the influence of cognitive function on discourse comprehension performance. Measuring cognitive function is another way to evaluate the two hypotheses that suggest declines in the processing of time-compressed speech should be attributed primarily to either cognitive decline or to the ability to process somewhat distorted acoustic information.

Impact of Cognitive Function

It was hypothesized that cognitive function, specifically speed of processing, would predict self-selected time-compression ratio. LMER results showed that processing speed and working memory capacity were significant predictors of self-selected time-compression ratios across all groups and both speech presentation rate conditions. During a discourse comprehension task, a listener must be able to receive auditory information, integrate the spoken message into memory while new information is being presented, and later retrieve the information to respond appropriately. Processing speed, or how quickly an individual can perform this set of tasks, would be expected to influence the self-selected time-compression ratio that produced a fixed level of comprehension accuracy performance. This relationship between processing speed and discourse comprehension performance has been theorized, but no previous study has directly assessed processing speed in relation to discourse comprehension performance for rapid or time-compressed speech. Working memory, which is the ability to manipulate and temporarily store incoming information (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980), would also be expected to influence self-selected time-compression ratios during a discourse comprehension task. Given that discourse comprehension and working memory both require similar manipulations and storage of information, it is not surprising that working memory capacity was a predictor of self-select listening rate. The results of the current study are consistent with those of Ward et al. (2016), despite the fact that this earlier study degraded speech spectrally with vocoding, whereas the current study degraded speech temporally with time compression. Results from Ward et al. (2016) showed a significant relationship between RSPAN scores and discourse comprehension performance for vocoded speech for younger and older adults. Although it may not be surprising that processing speed and working memory predict self-selected time-compression ratio, many previous studies have not found a relationship between discourse comprehension performance and cognitive function (Nagaraj, 2017; Smith & Pichora-Fuller, 2015).

It was also anticipated that cognitive function would be more predictive of older listeners' performance compared to that of younger listeners. Results of the current study revealed no interaction between the cognitive measures and listener group, which suggests that cognitive function was equally predictive of self-selected time-compression ratios for all three groups. Older adults showed declines in all cognitive measures compared to younger adults; they also selected slower time-compression ratios. However, individual differences in cognitive performance influenced self-selected time-compression ratio similarly for all groups. This result is consistent with results from Ward et al. (2016), who found that working memory capacity impacted discourse comprehension equally for younger and older groups. Taken together, cognitive abilities, specifically speed of processing and working memory capacity, are involved in the processing of rapid speech during a discourse comprehension task for younger adults and older adults with and without hearing loss.

Impact of Speech Intelligibility

It was hypothesized that speech intelligibility would not be a significant predictor of self-selected time-compression ratio. This hypothesis was based on previous literature showing no relationship between discourse comprehension and speech intelligibility performance in difficult listening situations (Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015). The relationship between speech intelligibility and discourse comprehension for time-compressed speech has not been measured directly in previous literature. The results of the current study showed that speech intelligibility was not a significant predictor of self-selected time-compression ratio during the discourse comprehension task. This result is consistent with three prior studies that have measured the relationship between speech intelligibility and discourse comprehension under different challenging listening situations (Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015). These prior studies compared the percent correct performance for speech intelligibility to the percent correct performance for discourse comprehension and found no relationship between the two measures. However, traditional measures of speech intelligibility and discourse comprehension evaluate different processes. Discourse comprehension involves the interpretation of a spoken message, which is measured by answering questions or summarizing a passage. Speech intelligibility is the recognition of a spoken message, which is typically measured with an immediate recall task. According to Wingfield and Tun (2007), discourse comprehension involves more top-down processing compared to speech intelligibility. Therefore, a direct comparison between percent correct performance may not be an appropriate technique to evaluate the relationship between these two processes. The current study utilized a different methodological approach than previous studies (Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015), but nonetheless did not reveal a significant systematic relationship between speech intelligibility and discourse comprehension performance. Although the results of the current study tend to support previous literature, it is critical to consider the design of the current study. In the current study, speech intelligibility performance was not assessed at the same time-compression ratio across listeners; rather, speech intelligibility was only measured at each individual's self-selected time-compression ratio. Consequently, it is difficult to make firm conclusions about the relationship between relative speech intelligibility performance and self-selected time-compression ratio in this study. Due to the nature of the testing paradigm, there may be alternative explanations to the LMER results.

In the current study, the correlation between speech intelligibility performance and time-compression ratio was not significant, with a relatively shallow slope observed between speech intelligibility scores and time-compression ratios. This result would ordinarily mean that speech intelligibility did not predict or modulate self-selected time-compression. However, a possible explanation for this finding is that speech intelligibility scores were limited in range relative to the large range in self-selected time-compression ratios. It is conceivable that when listeners self-selected their time-compression ratios, they required a certain level of speech intelligibility performance, and thereby minimized the systematic variability in speech intelligibility performance that would be required of a significant predictor variable. In this case, it is possible that listeners were using speech intelligibility to aid in their time-compression selection because the systematic variation in intelligibility was minimized. This notion would suggest that speech intelligibility performance was somewhat modulating the self-selected time-compression ratio. For example, if speech intelligibility was completely irrelevant to self-selection, then a listener with a very fast self-selected time-compression ratio might show poor intelligibility. In other words, if speech intelligibility did not impact the self-selected time-compression ratios, then we would expect to see a steeper slope between speech intelligibility and time compression. Taken together, the fact that speech intelligibility is not a predictor should not be interpreted to mean that speech intelligibility was irrelevant in the self-selection process. Due to the nature of the testing paradigm, the null result cannot conclusively rule out speech intelligibility as a contributing factor.

Further study is required to authoritatively confirm the presence of a relationship between speech intelligibility performance and the self-selected time-compression ratio during a comprehension task. In future studies, it would be critical to measure speech intelligibility performance at a variety of time-compression ratios (i.e., faster and slower relative to the listener's selection during the story). This technique would provide more insight into how speech intelligibility at a variety of performance levels and time-compression ratios relates to preferred listening rate during a discourse comprehension task. Alternative speech intelligibility material should also be considered.

Self-Selection Method

The use of a self-selected time-compression paradigm in the current study is relatively unique. Therefore, it was important to confirm that the method of self-selection was reliable and sensitive to differences between listener groups. Listeners across the three groups and both speech presentation rate conditions selected consistent time-compression ratios over multiple trials. This result shows that the self-selection method produced reliable results, and participants were systematically selecting a preferred time-compression ratio across trials. The reliable self-selected time-compression ratios consistently yielded performance near 80% correct for all groups and both speech presentation rate conditions. Based on these results it can be concluded that listeners, even with hearing loss, self-select a time-compression ratio that leads to adequate comprehension performance. Therefore, if this self-selection method were used clinically, it is possible that comprehension accuracy questions may not need to be assessed. If the comprehension questions did not need to be assessed, then the method of self-selection could quickly provide valuable information regarding discourse comprehension performance (i.e., 1-minute per story). Long trial times in discourse comprehension are problematic in the clinical setting because of excessive time requirements. The current paradigm has the potential to measure the impact of rapid speech on discourse comprehension in a relatively brief period of time.

Conclusions

This study evaluated the impact of age, hearing loss, speech intelligibility and cognitive function on self-selected time-compression ratios during a discourse comprehension task. Results showed that younger listeners selected faster time-compression ratios in both speech presentation conditions compared to the older groups. These findings indicate that ONH and OHI listeners are more adversely affected by rapid speech than YNH listeners during a discourse comprehension task. In addition, OHI listeners selected slower time-compression ratios for both conditions compared to ONH listeners. This suggests that hearing loss, in addition to age, affects the rate at which listeners can comprehend a spoken passage. This study is the to first demonstrate that OHI listeners may have more difficulty processing rapid speech during a discourse comprehension task. In addition, OHI listeners benefited more from the presence of silences in the silences intact condition compared ONH and YNH groups. This result suggests that the OHI group was relying more on cognitive processing resources and perception of overall information rate compared to the two normal hearing groups. The results of the current study suggest that discourse comprehension cannot be predicted from speech intelligibility performance. However, due to the nature of the testing paradigm it cannot be definitively ruled out that speech intelligibility did not aid in the self-selection process. Quantifying the relative contributions of speech intelligibility during difficult and realistic discourse comprehension tasks is an important area for future research. This study also examined the relationship between cognitive function and self-selected time-compression ratio. Results showed that processing speed and working memory were significant predictors of self-selected time-compression ratio, which suggests that cognitive function impacts the processing of rapid speech during a discourse comprehension task for all groups.

Experiment 2

Introduction

A common complaint of older adults with and without hearing loss is difficulty comprehending speech in noisy environments. However, little is known about how background noise affects listening comprehension for extended passages of speech in this population. This type of listening comprehension is referred to as discourse comprehension. Discourse comprehension is a skill required for verbal communication and is vital to everyday social interactions. Unfortunately, due to the complex nature of discourse comprehension, performance is often challenging to measure. In order to measure discourse comprehension performance, participants must listen to a passage and then answer questions or summarize the passage, which results in long trial times. Lengthy experimental trial times lead to the inability to parametrically measure a psychometric function or perform an adaptive tracking procedure for discourse comprehension performance. Therefore, the influences of SNR and masker type on discourse comprehension performance have remained difficult to quantify (Best et al., 2016).

The majority of auditory research and clinical assessments is focused on measuring speech intelligibility. Speech intelligibility is defined as a listener's ability to recognize and repeat a sentence or word. It has been well established that older adults, even with normal hearing, require more favorable SNRs compared to younger adults with similar hearing thresholds in order to achieve comparable levels of speech intelligibility performance (Dubno et al., 1984). In contrast, very little is known about the differences in SNRs required for younger and older adults with and without hearing loss to achieve similar levels of discourse comprehension performance. A goal of the present research is to determine the self-selected SNR required for younger and older adults with and without hearing loss to perform similarly on a discourse comprehension task. The ability to comprehend a message may be more relevant than intelligibility for daily communication (Hygge et al., 1992; Schneider et al., 2016; Wingfield & Tun, 2007) and more research is required in order to better understand the effects of SNR and masker type on discourse comprehension. There is also a critical need for a clinical measure of discourse comprehension performance that can be utilized to address this issue. Understanding how self-selected SNR and masker type impact discourse comprehension for older adults with and without hearing loss could provide insight into these patients' abilities to comprehend speech in the real world. A limited number of studies have aimed to examine the relationship between background noise and discourse comprehension performance using clinically feasible methods.

Hygge et al. (1992) implemented a self-selection task to assess the impact of background noise on discourse comprehension performance for normal-hearing and hearing-impaired individuals under the age of 65 years. Participants were asked to listen to a story spoken by a female speaker and then adjust the level of background noise to a level at which they could subjectively comprehend what was said. There were three types of background noise: (1) speech shaped noise (SSN); (2) a male 1-talker (1T) masker; and (3) a male reversed 1T masker. Each condition lasted for 3 minutes and the signal level was attenuated randomly between 10 and 25 dB at 30, 65, 100 and 135 seconds throughout the story. When the signal level was attenuated between 10 and 25 dB, the story became more difficult to hear. Each time the signal level was attenuated, the participants were instructed to raise the level of the signal until they could just follow the passage again.

Results from the Hygge et al. (1992) study revealed differences in self-selected SNRs between the normal-hearing and hearing-impaired listeners. The hearing-impaired listeners self-selected a higher signal level, and consequently a higher SNR, across all masker types compared to the normal hearing listeners, suggesting that hearing loss has an impact on the SNR required for discourse comprehension. Furthermore, hearing-impaired listeners selected similar SNRs for all masker types, signifying that the type of masker did not influence comprehension for the hearing-impaired group. In contrast, the normal-hearing listeners selected a lower signal level (and thus, lower SNR) in the 1T masker condition compared to the speech-shaped noise masker condition. This result indicates that during discourse comprehension, normal-hearing listeners were able to benefit from the amplitude modulations in the 1T masker in order to select a lower SNR compared to the steady-state SSN condition. The authors concluded that for normal-hearing listeners, energetic masking observed in the SSN condition was more detrimental to comprehension performance compared to interference of the 1T masker. This result also indicates that the 1T masker was not an effective speech masker because there was less masking with this speech masker than with the non-speech masker. One possible reason for this outcome was that listeners took advantage of differences in voice pitch between the target talker (female) and the 1T masker (male).

Hygge et al. (1992) implemented a clinically feasible method to measure the impact of masker type on discourse comprehension performance. However, it is important to note that the listeners were not specifically tested on comprehension accuracy performance. Therefore, it is possible that when listeners adjusted the level of the story, they were adjusting for the audibility of the speech material in real time, rather than adjusting to a level that would enable them to remember and answer comprehension questions about the 3-minute story. Thus, it is unknown whether or not listeners would have self-selected a higher signal level in the 1T condition if the added cognitive demands of memory and passage interpretation were required (Baddeley & Salame, 1983; Bell et al., 2008). Previous literature examining the impact of cognitive function on speech intelligibility performance in the presence of a 1T masker suggests that listeners require increased cognitive resources to separate the target signal from a 1T masker compared to a SSN masker (Koelewjin et al., 2012). Schurman et al. (2014) reported that younger and older normal-hearing adults had more difficulty recalling sentences during an n-back task with a 1T masker compared to SSN. Given that there is an increase in difficulty in the comprehension task when story memorization and recall are required, it is possible that the minimal demands of the listening task in the Hygge et al. study did not capture the expected effect of energetic masking compared to a speech masker on discourse comprehension. Lastly, cognitive function was not assessed in this prior study. Measuring cognitive function across multiple domains should provide insight into the cognitive processes that are involved in processing discourse comprehension in background noise.

In the current experiment, a method of self-selection was used to quantify the impact of masker type on self-preferred SNR during a discourse comprehension task. Comprehension accuracy was also measured by requiring listeners to answer questions about the story. Based on pilot testing, the self-selection method in the current study is expected to produce a fixed level of comprehension accuracy performance (i.e., near 80% correct). However, the self-selected SNRs to achieve this level of accuracy are expected to vary greatly across groups and masker types. The variable of interest in the present study is the self-selected SNR, which yields a fixed level of discourse comprehension performance. In addition, listeners were separated into groups based on age and hearing status in order to determine the impact of age and hearing loss on self-selected SNR. The influence of cognitive function on discourse comprehension was also of interest in the current study. Lastly, characterizing the unique self-selection method was also an aim of the current study.

Research Questions and Hypotheses

The goals of Experiment 2 are to: (1) determine the impact of masker type, age and hearing loss on self-selected SNR during a discourse comprehension task for younger normal hearing adults and older adults with and without hearing loss; (2) determine the relationship between self-selected SNR for discourse comprehension and speech intelligibility performance for words and sentences for young normal hearing adults and older adults with and without hearing loss; and (3) determine which cognitive domain (i.e., working memory, selective attention and/or speed of processing) is the most important predictor for self-selected SNR for younger normal hearing adults and older adults with and without hearing loss.

(1) Does masker type, age and hearing loss impact self-selected SNR during a discourse comprehension task?

It is well known that older adults, even with normal hearing, require a higher SNR to perform similarly to younger normal-hearing listeners during a speech intelligibility task (Dubno et al., 1984). Therefore, it is hypothesized that older adults in both groups will select more favorable (i.e., higher) SNRs across all masker conditions compared to younger adults during the discourse comprehension task. It is hypothesized that older adults with hearing loss will select higher SNRs compared to older normal hearing listeners across all masker types, a result consistent with that reported by Dubno and colleagues (1984). In addition, it is expected that all groups will select a higher SNR in the 1T masker condition compared to speech shaped noise (SSN) and Cafeteria noise. Schurman et al. (2014) reported that younger and older adults had more difficulty recalling sentences during an n-back task with a 1T masker compared to SSN, possibly implying that listeners may require a more favorable SNR for a 1T masker compared to non-speech maskers during a comprehension task. Self-selecting a higher SNR in the 1T condition would suggest that the presence of a 1T masker impacts discourse comprehension performance to a greater extent compared to the SSN and Cafeteria maskers. Finally, it is expected that there will be an interaction between masker type and group, in which the magnitude of the effect of a speech masker, as observed with a 1T masker, is larger for older adults than for younger adults. This result would signify that a 1T masker more adversely impacts older adults with and without hearing loss during a discourse comprehension task compared to younger normal hearing adults, a finding consistent with that observed for speech intelligibility performance in the presence of a speech masker (Tun et al., 2002).

(2) Is there a relationship between self-selected SNR and speech intelligibility performance for younger normal hearing adults and older adults with and without hearing loss?

The relationship between discourse comprehension and speech intelligibility performance is largely unknown. Based on previous literature, it is hypothesized that speech intelligibility performance will not predict the self-selected SNR in the discourse comprehension task for either younger or older listeners (Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015). This result would suggest that SNR selection was not modulated by intelligibility across masker type and groups.

(3) Which cognitive domain contributes most to variance in self-selected SNR during a discourse comprehension task across masker types and listener groups varying in age and hearing loss?

Listeners who are better able to identify the target stimuli in the presence of background noise will likely have a larger working memory capacity, more attentional resources and faster speed of processing (Salthouse, 1985; Rönnberg et al. 2013). Working memory requires individuals to manipulate and temporarily store incoming information, and discourse comprehension requires this same set of skills (Ward et al. 2016). Therefore, it is hypothesized that working memory capacity will be the most predictive of self-selected SNR during discourse comprehension compared to attention and speed of processing. Furthermore, working memory is hypothesized to have the strongest relationship to the self-selected SNR in the 1T condition, because increased cognitive resources may be required to process a signal in the presence of a speech masker (Baddeley & Salame, 1983; Bell et al., 2008). Lastly, it is hypothesized that the relationship between working memory and self-selected SNR will be stronger for older adults with and without hearing loss compared to younger adults, because older adults require the use of more top-down processing to compensate for age-related declines in bottom-up processing (Rönnberg et al., 2013).

Methods

Participants

The same participants from Experiment 1 were recruited for Experiment 2. Three groups of 18 listeners participated: younger adults aged 18-30 years with normal hearing (YNH), older adults 60 years and older with normal hearing (ONH) and older adults 60 years and older with hearing loss (OHI). Normal hearing is defined as thresholds ≤ 25 dB HL between 250 and 4000 Hz (re: ANSI, 2018). Participants with hearing loss had thresholds greater than 25 dB HL between 250 and 4000 Hz and a pure tone average (PTA) of 50 dB HL or better. Participants had at least a high school education and were required to be native speakers of English. The Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) was used as an assessment of mild cognitive impairment. Participants who passed the test with a score of ≥ 26 (i.e., did not have a mild cognitive impairment) participated in the study.

Stimuli

The comprehension and intelligibility stimuli in Experiment 2 were the same as those used in Experiment 1. All speech materials were recorded using a DPA4017B shotgun microphone with a 48kHz-sampling rate. Each story, sentence and word were stored as a separate waveform file using Adobe Audition, and the RMS levels were held constant for all stimuli.

In order to remove talker variability as a confounding variable, one native English female talker recorded all speech materials. Comprehension stimuli were stories from the Discourse Comprehension Test (DCT) (Brookshire & Nicholas, 1993). The DCT is comprised of twelve 1-minute passages. Each narrative is followed by eight yes or no questions, which assess either the main idea of the passage (stated and implied) or story details (stated and implied). DCT questions were presented visually via a computer monitor. Sentence materials were 200 high probability (HP) R-SPIN sentences and 200 anomalous probability (AP) sentences, which were derived from the HP sentences (Bilger et al., 1984). All nouns, verbs and adjectives were considered keywords, and there were three to seven keywords in each sentence. (HP example: His PLANS MEANT TAKING a BIG RISK. AP example: His DOCTOR DRANK a LOST RISK.). Individual word intelligibility was measured using a revised version of the Modified Rhyme Test (MRT) (Bell et al., 1972).

Stories and sentences were chosen randomly for each experiment. Stories and sentences presented in Experiment 1 were not presented in Experiment 2. All speech stimuli were presented in three different types of background noise: (1) Speech-shaped noise (SSN); (2) Cafeteria noise (Cafe) (Kayser et al., 2009); and (3) a 1T masker. The cafeteria noise was recorded at the University of Oldenburg during lunchtime in a crowded cafeteria. The cafeteria recordings consist of unintelligible babble in conjunction with the sounds of dishes and silverware clinking and chairs moving. The 1T masker was comprised of selected passages from Grimm's Fairy Tale Classics. A native English female talker, different from the target talker, recorded the Grimm's Fairy Tale stories. All maskers were equated in RMS level.

Procedures

Experimental Procedures.

During the experimental conditions, participants were seated in a quiet room and heard speech material presented from an ASIO player on a laptop computer. The speech stimuli were routed from the laptop to an RME Digiface USB portable digital audio interface and delivered binaurally to the listener through Sennheiser HDA 200 headphones. Sennheiser HDA 200 headphones were chosen due to the excellent passive attenuation capabilities.

All target speech stimuli were presented at 80 dB SPL. Presentation level was verified and calibrated using a sound level meter connected to a flat-plate coupler. Three conditions were tested in the discourse comprehension portion of the experiment: (1) DCT stories in the presence of an SSN masker; (2) DCT stories in the presence of a Cafe masker; and (3) DCT stories in the presence of a 1T masker. In each trial, the DCT story levels were fixed at 80 dB SPL and the SNR originated at +5 dB. Participants were presented with two stories within each masker type (Trial 1 and Trial 2). The order of conditions was randomized across participants using a block design. For example, participants were first presented with two trials in the 1T condition, then two trials in the Cafe condition and two trials in the SSN condition. Listeners were instructed to vary the level of the background noise to the loudest level at which they felt that they could still understand and answer questions about the story. A custom Matlab program controlled the SNR adjustment during each story. SNRs ranged from -15 dB to +15 dB. The scroll wheel on the computer mouse controlled the change in masker level. For example, moving the wheel down towards the desk made the SNR lower or less favorable and moving the wheel in the opposite direction made the SNR higher or more favorable. The final self-selected SNR was calculated by taking an average of the SNRs during the last 20% of the story in order to account for accidental or substantial last-minute adjustments to the SNR. A practice story was provided to ensure that participants understood the task. Before the start of each story, there was a set of spoken instructions that read, “Please start to adjust the noise level now. Your goal is to play the stories with as much noise as you can while still being able to understand and answer questions about the story. You can also adjust the noise level during the story. The story will begin after the beep.” These instructions were provided before the start of the story to allow for the adjustments to begin before the story started. After the completion of each story, the participants were presented with eight comprehension questions. There were four types of questions: main idea stated, main idea implied, detail stated and detail implied. Two outcome measures were collected from each story by assessing discourse comprehension performance in this manner: (1) the self-selected SNR and (2) the overall performance accuracy on the comprehension questions.

After the discourse comprehension task in each masker condition, participants completed the intelligibility measures. For example, a participant who heard two stories in the SSN condition was then presented with the intelligibility measures in the SSN condition. The average self-selected SNRs for the stories in each masker condition were used to assess sentence and word intelligibility performance. There were a total of nine speech intelligibility conditions tested in this portion of the experiment: (1) 20 HP sentences; (2) 20 AP sentences; and (3) 30 MRT words presented at the average self-selected SNR from the DCT SSN condition; (4) 20 HP sentences; (5) 20 AP sentences;

and (6) 30 MRT words presented at the average self-selected SNR from the DCT Cafe noise condition; (7) 20 HP sentences; (8) 20 AP sentences; and (9) 30 MRT words presented at the average self-selected SNR from the DCT 1T masker condition (see Table 3.1). Sentences presented in Experiment 1 were not presented in Experiment 2.

Table 3.1.

Design of Experiment 2

Task	Masker Conditions	Stimuli	SNR presentation level	Outcome Variable
Comprehension	Speech Shaped Noise	2 DCT stories	Self-selected	(1) Self-selected SNR (2) Comprehension questions percent correct
	Cafeteria Noise			
	1-Talker masker			
Intelligibility	Speech Shaped Noise	20 HP sentences 20 AP sentences 30 MRT words	Average self-selected SNR from the same masker in the comprehension task	(1) Percent Correct
	Cafeteria Noise			
	1-Talker masker			

Cognitive Measures.

All cognitive measures were the same as those used in Experiment 1. Cognitive function and vocabulary knowledge were assessed using a variety of test materials. Cognitive measures were presented in the same order for each subject: (1) Trail-Making Test (processing speed); (2) Flanker (attention); (3) Picture Vocabulary Test; and (4) RSPAN (working memory).

Statistical Analysis

Data analyses included correlations, multivariate repeated-measures analyses of variance (ANOVA) and a Linear Mixed-Effects Regression (LMER) model. The first portion of the results section describes the unique self-selection paradigm. First, in order to determine if listeners were reliable when selecting SNR, a bivariate correlation between the SNR selected in trial 1 and trial 2 for each individual was performed separately for all masker conditions. Next, the self-selected SNRs were characterized in two ways: (1) raw SNR selections for the duration of the story and (2) the SNR difference from the ending selection over the duration of the story. Comprehension question accuracy scores were analyzed with a multivariate repeated-measures ANOVA with one within-subjects variable [Masker type (three levels: SSN, Cafe, and 1T)] and one between-subjects variable [Group (three levels: YNH, ONH, and OHI)]. Based on pilot data, comprehension accuracy scores were expected to be similar (on average near 80% correct) for all groups and masker conditions. Lastly, a bivariate correlation between the average self-selected SNR and the average comprehension question accuracy score for each individual was performed with combined groups for each masker type.

An LMER was performed to determine the impact of masker type, age, hearing loss, speech intelligibility performance and cognitive function on self-selected SNR. The LMER was performed using a maximal model structure with a backward selection approach following Barr et al. (2013). The maximal model included all fixed effects, interactions and random effects of interest. The model included random intercept of participant and a random slope of condition. The maximal model contained fixed main effects of masker condition [three levels = SSN (reference), Cafe, and 1T], group [three levels: YNH (reference), ONH, and OHI], speech intelligibility scores [two levels: HP (reference) and AP sentences], RSPAN score, Trail Making speed (Version B), Flanker Score and Vocabulary score. The assessments from the NIH Toolbox (Flanker and Vocabulary Score) were input into the model as raw scores instead of age-corrected scores, because age was taken into account in the model. All cognitive tests (RSPAN, Trail Making, Flanker and Vocabulary score) were assessed for collinearity ($r \geq .7$) and standardized to z-scores before being entered into the model. Speech intelligibility measures were also assessed for collinearity and converted to RAUs prior to being added to the model. MRT performance was found to be correlated with HP and AP performance ($r \geq .7$) and was removed from the model in order to reduce collinearity in that predictor variable. Speech intelligibility contextual benefit (i.e., difference in percent correct performance for HP and AP sentences) was correlated with AP intelligibility performance ($r \geq .7$) and was not entered into the model. The buildmer package (version 1.6) in R (Voeten, 2019) was utilized to implement the backwards elimination approach outlined by Barr et al. (2013). Buildmer reduces fixed and random effect terms in order to determine the final best fitting model for the data. Buildmer determines the order of the effects in the model and then a backward stepwise elimination procedure is performed. Terms that were not significant ($p > .05$) were not included in the final model reported. In addition, significant interactions were analyzed further by re-leveling the reference variables involved in the interactions.

Results

Self-Selection Method

Two outcome measures were collected: (1) the self-selected SNR and (2) the overall percent correct performance for the comprehension questions. The use of a self-selection paradigm is relatively unique. Therefore, it is important to verify that the method of self-selected SNR used in this study is valid and reliable. First, a bivariate correlation, conducted with data from all groups, revealed a significant positive correlation between the self-selected SNR in trial 1 and trial 2 for all masker conditions

(Figure 3.1). All groups selected a consistent SNR over multiple repetitions for all masker types [SSN (r=.97; $p<0.001$), Cafe (r =.96; $p<0.001$), 1T (r =.93; $p<0.001$)]. This result suggests that the self-selection method produced reliable results across conditions and participants.

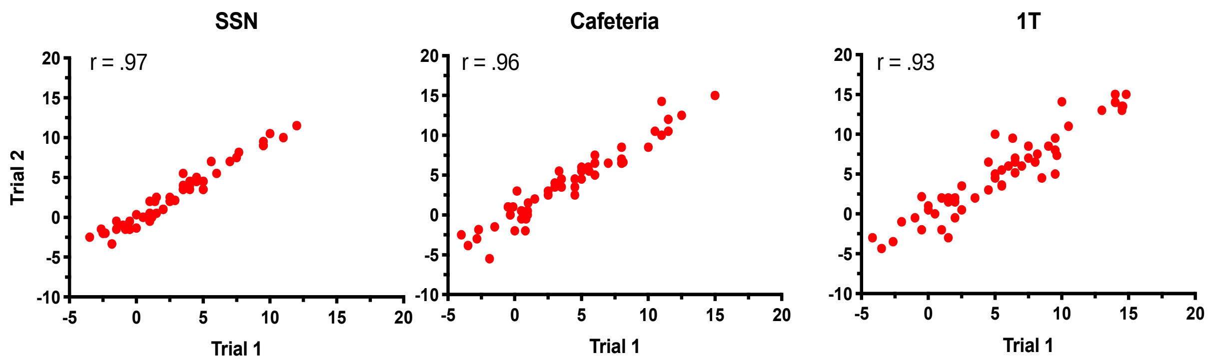

Figure 3.1. Bivariate correlation between the self-selected SNR in trial 1 and trial 2 for each masker type.

Next, the self-selected SNRs over the duration of the story were quantified. The scroll wheel on the computer mouse controlled the adjustments, and the SNR at each time point throughout the listening interval was recorded (i.e from 0 to 100 percent completion). The self-selected SNRs over the duration of the story were characterized in two ways: (1) raw SNR selections over the duration of the story (Figure 3.2) and (2) the SNR difference from the ending selection over the duration of the story (Figure 3.3). The dashed lines in Figures 3.2 and 3.3 represent the end of the spoken instructions and the start of the discourse comprehension story. The averaged self-selected SNRs for each group from 0 to 100 percent completion are shown in Figure 3.2. Figure 3.2 shows that the majority of large adjustments to SNR were made during the instruction period. This result demonstrates the importance of the presentation of spoken instructions prior to the start of the story.

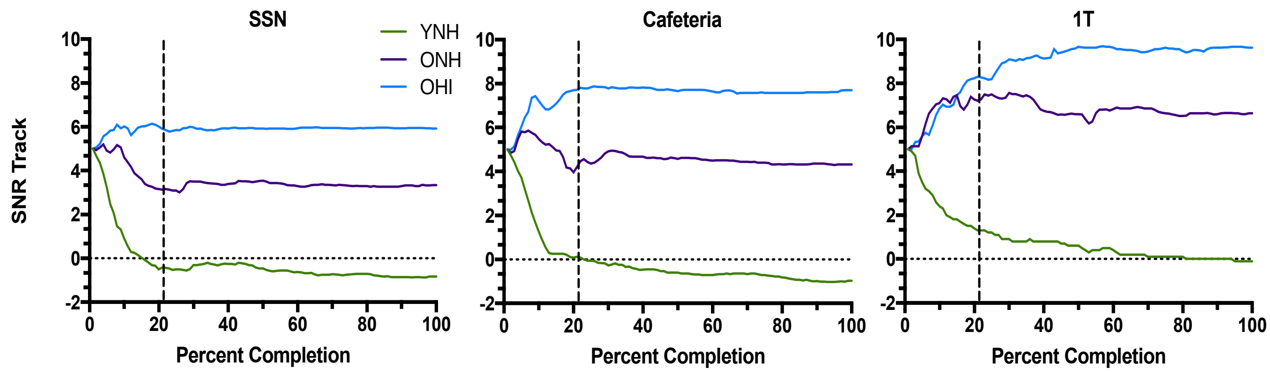

Figure 3.2. Average SNR selection (y-axis) as a function of percent completion of the listening interval (instructions and story) for all groups (YNH, ONH and OHI) and all masker types (Left Panel: SSN, Middle Panel: Cafeteria, Right Panel: 1T). The dashed line represents the end of the spoken instructions and the start of the discourse comprehension story.

Figure 3.3 demonstrates the absolute difference in self-selected SNR over the course of the stimulus compared to the ending SNR averaged for all participants in each group. Viewing the data in terms of the absolute difference in SNR from the ending SNR is expected to provide insight into how close listeners were to their final SNR over the course of the stimulus (i.e., 0 to 100 percent completion). On average, at the start of the story, YNH listeners were 1.4 dB (SSN), 2 dB (1T) and 1.5 dB (Cafe) away from the final self-selected SNRs. At the start of the story, ONH listeners were 1.2 dB (SSN), 2.4 dB (1T) and 1.1 dB (Cafe) away from the final self-selected SNRs. Finally, at the start of the story, OHI listeners were .6 dB (SSN), 1.4 (1T) an .7 dB (Cafe) away from the final self-selected SNRs. Thus, listeners in all groups appear to be relatively close to their final self-selected SNRs across masker types at the start of the discourse comprehension story.

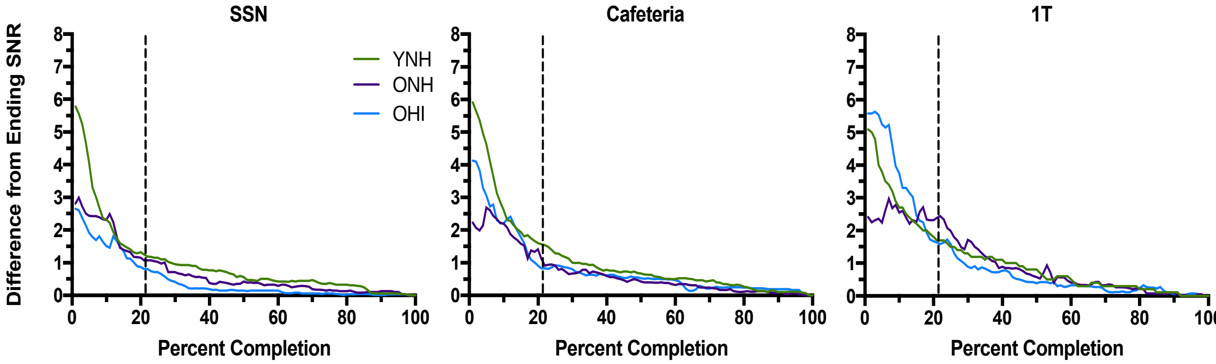

Figure 3.3. Average SNR difference from the ending SNR (Y-axis) as a function of percent completion of the listening interval (instructions and story) and the story for all groups (YNH, ONH and OHI) and all masker types (SSN, Cafe, and 1T). The dashed line represents the end of the spoken instructions and the start of the discourse comprehension story.

An ANOVA was performed to determine if there were significant differences in the absolute difference in SNR at the start of the story from the final SNR across groups and masker types. The dependent variable was the difference in self-adjusted SNR between the start of the story and the end of the story. The ANOVA had one within-subjects variable [Masker type (three levels: SSN, Cafe, 1T)] and one between-subjects variable [Group (three levels: YNH, ONH, and OHI)]. This analysis revealed a significant main effect of masker type [$F(2,51)=8.21$, $p<.005$, $\eta p^2 = .186$], no main effect of group [$F(2,50)=.68$, $p>.05$, $\eta p^2 = .096$], and no interaction [$F(2,51)=2.72$, $p>.05$, $\eta p^2 = .053$]). Paired comparison t-tests with Bonferroni corrections were conducted on data collapsed across groups to assess the main effect of masker type. The effect of masker type was due to a higher absolute difference from the final SNR at the start of the story for the 1T masker compared to that observed for both the SSN ($p<.001$) and Cafe ($p<.001$) maskers. There was no significant difference between the SSN and the Cafe maskers ($p>.05$). These results suggest that listeners in all groups took longer to reach their final SNR in the 1T masker condition compared to the SSN and Cafe maskers. Although this difference is significant, it is important to note that listeners were nevertheless reasonably close to their ending SNR at the start of the story in the 1T condition (YNH: 2 dB; ONH: 2.4 dB; OHI: 1.4 dB). The reason for this difference is likely that the final self-selected SNR in the 1T condition is the furthest from the original starting SNR, and listeners needed to make a larger SNR adjustment to reach their final preferred SNR.

Comprehension Accuracy Results

Figure 3.4 shows comprehension percent correct performance for the questions that followed each story for the three listener groups. Comprehension question accuracy scores were analyzed with an ANOVA with one within-subjects variable [masker type (three levels: SSN, Cafe, 1T)] and one between-subjects variable [Group (three levels: YNH, ONH, and OHI)]. An ANOVA revealed no main effect of group [$F(2,51)=1.97$, $p>.05$, $\eta p^2 = .072$], no main effect of masker type [$F(2,50)=.68$, $p>.05$, $\eta p^2 = .013$], and no interaction [$F(2,51)=.77$, $p>.05$, $\eta p^2 = .03$]). These results suggest that participants in all groups were able to achieve similar comprehension accuracy performance across all masker types.

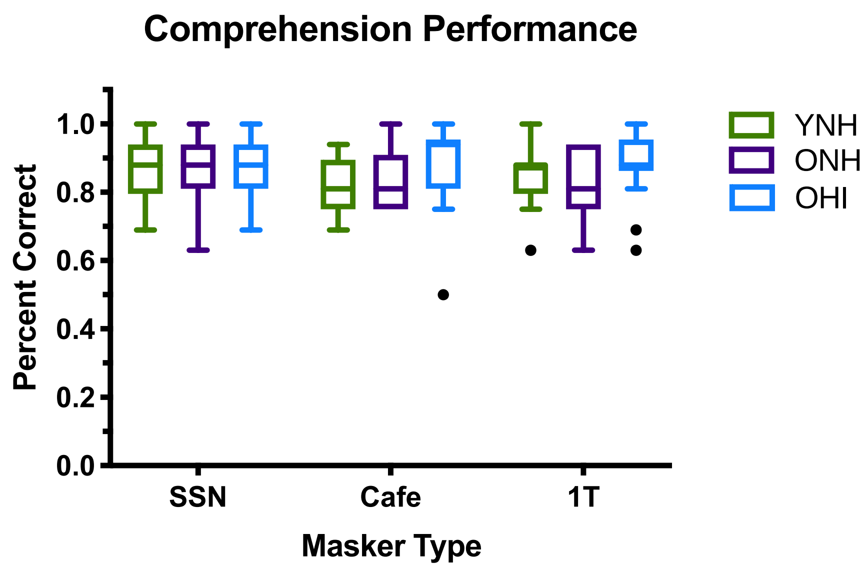

Figure 3.4. Comprehension accuracy performance for each group and all masker types for eight yes or no questions presented directly following the completion of each story. The box and whisker plots were created using the Tukey method, which calculates an interquartile range (i.e., the difference between the 25th and 75th percentiles). Any value that is greater or less than 1.5 times the interquartile range is plotted as individual outlier points.

It is clear from Figure 3.4 that comprehension accuracy performance was held relatively constant for all groups and all masker types. This result suggests that a wide range of self-selected SNRs yields a relatively fixed level of comprehension accuracy performance regardless of age and hearing acuity (see Figure 3.5). A bivariate correlation, including all groups, revealed no correlation between the self-selected SNR and comprehension accuracy performance [SSN (r=.01), Cafe (r =.1), 1T (r =-.002)] (Figure 3.5). These results verify that the variable of interest is the self-selected SNR that enabled the majority of listeners to achieve a consistent comprehension accuracy score.

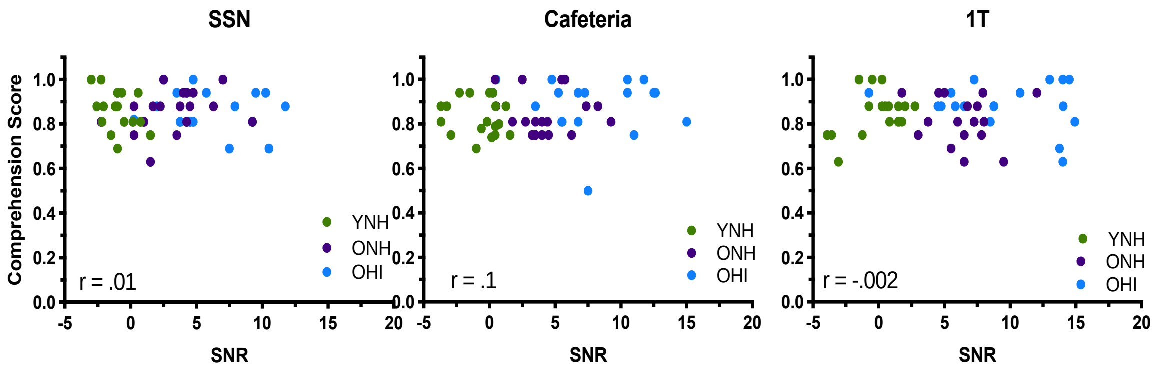

Figure 3.5. Bivariate correlations between self-selected SNR and comprehension accuracy performance for all masker conditions.

Based on these results it can be concluded that listeners in all groups are able to self-select an SNR that leads to adequate comprehension accuracy performance. Due to the overall consistency of the comprehension accuracy performance, it is possible that if this assessment were used clinically, then comprehension questions would not need to be assessed. The method of self-selection could provide valuable information regarding discourse comprehension performance relatively quickly (i.e., 1-minute per story). Long trial times in discourse comprehension are problematic in the clinical setting because of excessive time requirements. The current paradigm has potential to assess the impact of different types of background noise on discourse comprehension in a brief period of time.

Effect of Age and Hearing Loss

The results of the LMER are shown in Table 3.2. The dependent variable was self-selected SNR. The maximal model contained fixed main effects of masker condition [three levels = SSN (reference), Cafe and 1T], group [three levels: YNH (reference), ONH, and OH], speech intelligibility scores [two levels: HP (reference) and AP sentences], RSPAN score, Trail Making speed (Version B), Flanker Score and Vocabulary score. The model also included the random intercept of participant and a random slope of condition. In this section only the results regarding the impact of condition and group and will be discussed. In the final model, there were significant fixed effects of group ($p<0.001$), condition ($p<0.001$) and RSPAN score ($p<0.005$). Group and condition were also involved in a significant interaction (ONH x condition and OHI x condition). None of the higher-level interactions of fixed-effects (i.e., three-way or four-way) were significant. The reference variables were re-leveled in order to compare performance between the groups for all masker types, examine the relationships within groups for all masker types and further analyze the significant interactions. In this section, the results regarding the impact of condition and group will be discussed. LMER results involving speech intelligibility performance and cognitive function will be discussed in detail in the forthcoming results section.

Table 3.2.

Final LMER model for self-selected SNR. Bolded rows indicate significant terms (p<0.05).

Fixed Effects	Coefficient	SE	t	p
Intercept	0.25	0.60	0.42	0.676
Group [ONH]	2.51	0.85	2.95	**<0.005**
Group [OHI]	5.07	0.87	5.85	**<0.001**
RSPAN	-1.40	0.36	-3.86	**<0.005**
Condition [Cafe]	0.09	0.41	0.23	0.818
Condition [1T]	0.80	0.50	1.59	0.118
Interactions				
Group [ONH] * Condition [Cafe]	0.94	0.59	1.61	0.112
Group [OHI] * Condition[Cafe]	1.59	0.59	2.73	**<0.05**
Group [ONH] * Condition [1T]	2.41	0.71	3.38	**0.001**
Group [OHI] * Condition[1T]	2.62	0.71	3.67	**<0.001**
Random Effects	**Variance**			**SD**
Subject (intercept)	4.82			2.19
Subject/Condition [Cafe]	2.76			1.66
Subject/Condition [1T]	4.26			2.06
Residual	0.63			0.79

LMER results indicate that younger listeners were able to select lower SNRs across all masker types compared to ONH and OHI listeners during the discourse comprehension task (see Figure 3.6). In addition, there was no effect of masker type for YNH listeners (p >.05). This result indicates that masker type did not impact self-selected SNR during a discourse comprehension task for the YNH group. However, there was a significant interaction involving masker type and both older groups. Significant differences between all masker types were found for ONH and OHI groups, suggesting that masker type had an impact on self-selected SNR for older listeners with and without hearing loss. Although the results showed that ONH listeners were able to select lower

SNRs compared to OHI listeners for all masker types, the pattern of relative differences between SNRs across masker types was similar for both older groups. Both older groups selected the highest SNR in the 1T masker condition, then the Cafe condition and the lowest SNR was selected in the SSN condition [all comparisons between conditions were significant (p<.05)].

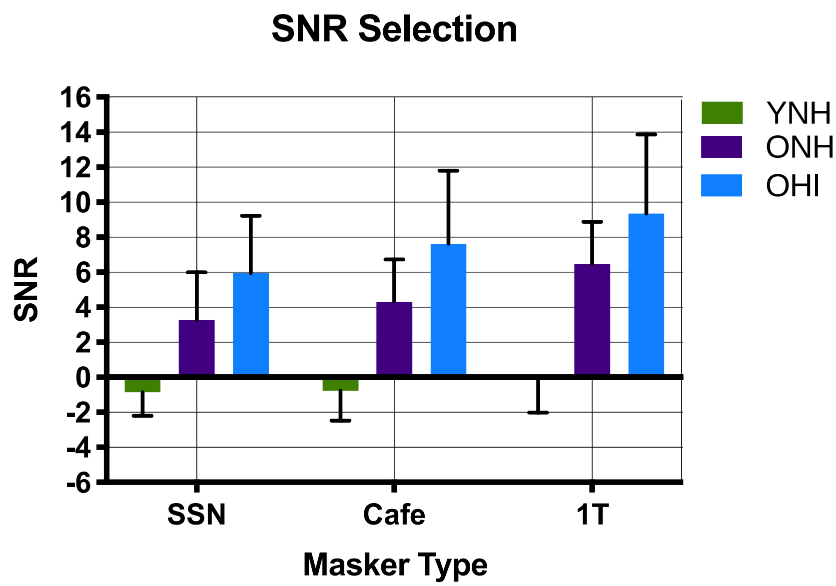

Figure 3.6. Self-selected SNR for all groups (YNH, ONH, OHI) for all masker types (SSN, Cafe, 1T). Error bars represent one standard deviation.

Impact of Speech Intelligibility and Cognitive Function

The LMER model reported in Table 3.2 above also examined the effects of speech intelligibility and performance on several cognitive measures on self-selected SNR. Speech intelligibility scores for the HP and AP sentences were tested at the self-selected SNR for each condition. Measuring speech intelligibility at the self-selected SNRs chosen during the discourse comprehension task was expected to provide insight into whether or not listeners used the intelligibility of the story as a criterion for their

SNR selection during the passage comprehension task. Results of the LMER showed that HP and AP speech intelligibility scores were not significant predictors of self-selected SNR during the discourse comprehension task (Table 3.2). Scatter plots of the self-selected SNRs and speech intelligibility scores (Figures 3.7 and 3.8) are shown to assist in the visualization of the relationships reflected in the LMER. Figures 3.7 and 3.8 show no relationship between speech intelligibility scores and self-selected SNRs.

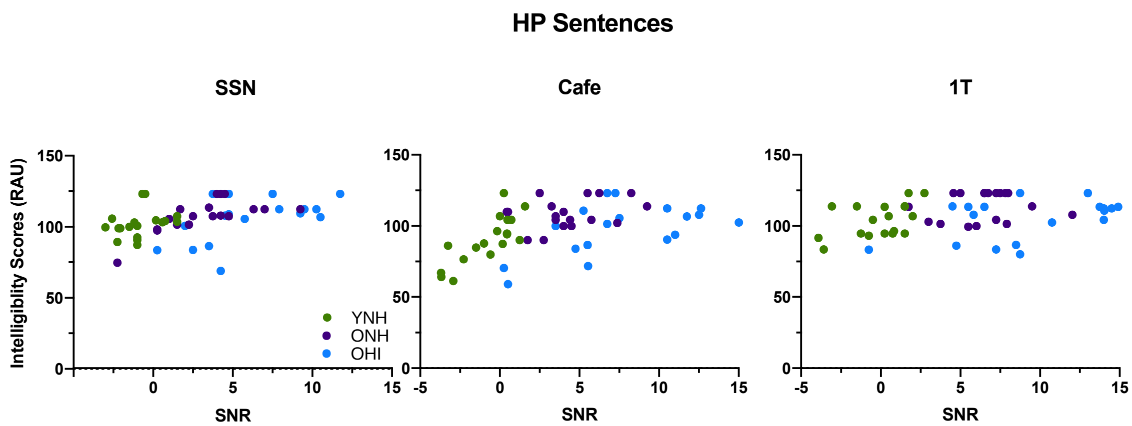

Figure 3.7. Scatterplots showing the spread of speech intelligibility scores (y-axis) and SNRs (x-axis) for all listener groups and masker types for HP sentences.

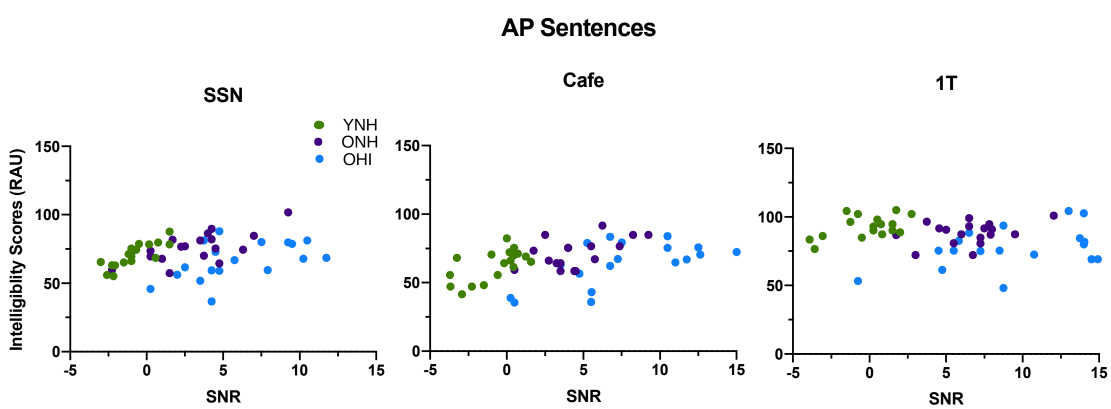

Figure 3.8. Scatterplots showing the spread of speech intelligibility scores (y-axis) and SNRs (x-axis) for all listener groups and masker types for AP sentences.

Cognitive function in the domains of speed of processing (Trail Making Task), working memory (RSPAN), and attention (Flanker Task) were evaluated to determine the cognitive variables that may predict self-selected SNR during a discourse comprehension task. Vocabulary knowledge (Picture Vocabulary Test) was also evaluated as a predictor of self-selected SNR. RSPAN was the only significant predictor of self-selected SNR during the discourse comprehension task (Table 3.2). No significant interactions were found in the LMER model between group, condition and RSPAN. This suggests that working memory capacity influences SNR preference during a discourse comprehension task equally for all three groups and masker types. Scatter plots of self-selected SNR as a function of RSPAN (Figure 3.9) allow for a visualization of the relationship reflected in the LMER results. Figure 3.9 shows a significant relationship between working memory capacity and self-selected SNR across all masker types.

Lastly, ANOVAs and appropriate post hoc comparisons were performed in order to determine the differences in performance for all groups on each cognitive measure. These results were discussed in detail in Experiment 1. YNH listeners showed higher RSPAN scores, faster Trail making speeds and higher Flanker scores compared to the ONH and OHI groups. There were no significant differences in RSPAN score, Trail Making speed or Flanker score between ONH and OHI groups.

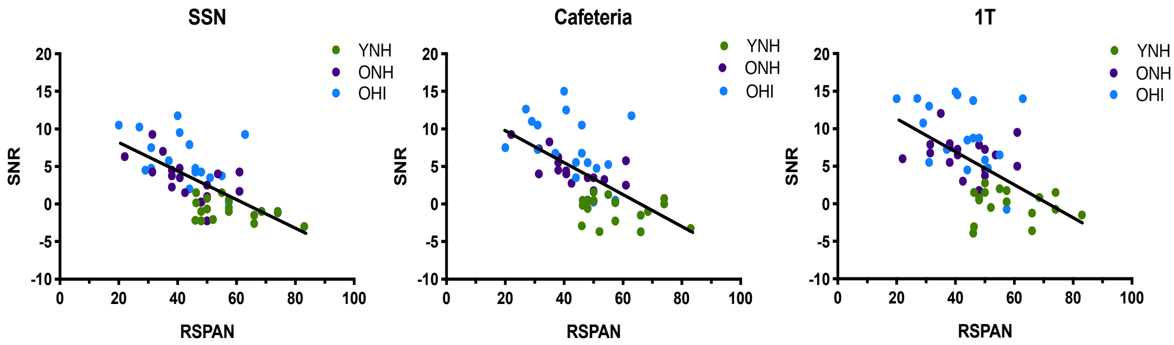

Figure 3.9. Scatterplots showing the spread of RSPAN (x-axis) and self-selected SNR (y-axis) in each masker type for the three listener groups with a regression line inserted.

Discussion

The first goal of this study was to determine the effects of masker type, age, and hearing loss on self-selected SNR during a discourse comprehension task for young normal hearing adults and older adults with and without hearing loss. The results revealed an effect of masker type and group, showing that YNH listeners selected less favorable SNRs compared to ONH and OHI groups and indicating increased difficulty in the presence of a speech masker compared to non-speech maskers for both older groups. The second goal of the study was to examine the relationship between speech intelligibility performance and self-selected SNR across the different listener groups. The results revealed that speech intelligibility performance was not a significant predictor of self-selected SNR. However, due to the nature of the task in the current study speech intelligibility performance cannot be definitively ruled out as a modulating factor of self-selected SNR. The final aim of this study was to determine which cognitive domain (i.e., working memory, selective attention or speed of processing) was predictive of self-selected SNR during a discourse comprehension task. Results showed that working memory was a significant predictor of self-selected SNR.

Effects of Age, Hearing Loss and Masker Type

The first aim of the present study was to determine if younger listeners with normal hearing and older adults with and without hearing loss self-select different SNRs in order to achieve similar levels of comprehension accuracy performance. Relatively little is known about the impact of age, hearing loss and masker type on discourse comprehension. It was hypothesized that older adults would select more favorable (i.e., higher) SNRs across all masker conditions compared to younger adults during the discourse comprehension task. It was also hypothesized that older adults with hearing loss would select more favorable SNRs compared to older normal hearing listeners across all masker types. The results revealed that both groups of older adults selected higher SNRs to reach a similar comprehension accuracy performance compared to the YNH group for all masker types. This suggests that ONH and OHI listeners may perform more poorly than YNH listeners in the presence of equivalent amounts of background noise during a discourse comprehension task. Murphy et al. (2006) and Tye-Murray et al. (2008) also found that YNH listeners performed better (i.e., answered more questions correctly) on a discourse comprehension task in the presence of background noise compared to older adults with normal hearing. These results and the results of the current study imply that discourse comprehension in a complex listening environment declines among older adults compared to younger adults. Furthermore, results of the current study also showed that OHI listeners required a higher SNR across all masker types compared to ONH listeners. This result implies that hearing loss, in addition to age, leaves OHI listeners at a disadvantage compared to ONH listeners for all masker types during a discourse comprehension task. A direct comparison of ONH and OHI listeners on a discourse comprehension task in the presence of background noise has not been studied previously. The results of the current study extend prior findings that age and hearing loss negatively and independently impact speech intelligibility performance in the presence of background noise to performance on a discourse comprehension task (Dubno et al., 1984; Goossens et al., 2017; Souza & Turner, 1994). Taken together, these results may have real-world implications because they suggest that older adults with and without hearing loss require a more favorable SNR compared to younger listeners in order to perform similarly on a comprehension accuracy task. Real-world conversation often occurs in the presence of background noise and listeners who need more favorable SNRs to comprehend a spoken message may have difficulty participating in meaningful communication.

It was hypothesized that the effect of masker type would be larger for older adults than for younger adults. LMER results revealed a significant interaction between listener group and masker type. The interaction is driven by the difference in the extent to which masker type impacted the self-selected SNRs for each group. There were no significant differences between self-selected SNRs across masker types within the YNH group. This result suggests that YNH listeners were not impacted by the presence of different types of maskers during the discourse comprehension task. In contrast, there were significant differences in self-selected SNRs across masker types for the ONH and OHI groups, suggesting that older groups were more impacted by masker type compared to YNH listeners. These group effects confirm seminal work examining the impact of background noise on speech intelligiblity performance (Dubno et al., 1984; Tun et al., 2002) and one discourse comprehension study (Murphy et al., 2006), which indicated that older adults were more affected by the presence of masking compared to younger adults. Murphy et al. (2006) found an interaction when measuring discourse comprehension performance in the presence a 2-talker masker. The presence of a speech masker may lead to increased difficulty attending to the target signal for older adults compared to an SSN masker (Humes et al., 2006; Tun et al., 2002).

It was also hypothesized that listeners would self-select the most favorable SNR in the 1T masker condition. This hypothesis was based on previous speech intelligibility literature indicating that increased cognitive recourses are required to separate the target signal from a 1T masker during speech recognition tasks with an increased memory load (i.e., a difficult n-back sentence task compared to an immediate recall task) (Koelewjin et al., 2012; Schurman et al., 2014). Given that memorization and recall were required for the discourse comprehension task, listeners were expected to require the most favorable SNR in the 1T condition in order to accurately answer comprehension questions. YNH listeners selected similar SNRs across all masker conditions, indicating that they were not impacted by the increased processing demands of the 1T masker. This result extends findings from Tun et al. (2002), who reported that younger adults were not impacted by the presence of a meaningful speech masker during a speech intelligibility task. Tun et al. (2002) suggested that younger adults are less vulnerable to a meaningful and intelligible masker compared to older adults. In the current study, both older listener groups needed to select a more favorable SNR in the 1T condition compared to the SSN and Cafe conditions in order to achieve equal comprehension accuracy performance. This result suggests that the presence of an intelligible 1T masker was more detrimental to older than younger adults, even those with normal hearing. In addition, the 1T masker and the target signal were both spoken by female talkers. The differences between the self-selected SNRs in the 1T condition compared to the SSN and Cafe conditions were likely exacerbated because the target and masker were both spoken by female talkers (Brungart, 2001). Brungart (2001) showed poorer speech intelligibility performance when the target and the masker talkers were the same sex than when they were of the opposite sex. Taken together, these results suggest that the 1T speech masker was the most detrimental to discourse comprehension performance for older adults with and without hearing loss. This result could have real-world implications because a single competing talker or other intelligible maskers are often present in real-world environments, which may cause older adults to have difficulty comprehending the target spoken message.

The results of the current study are somewhat contradictory those reported by Hygge et al. (1992). They also implemented a self-selected SNR method and also found that hearing impaired individuals selected a higher SNR across all masker types, compared to normal hearing listeners, suggesting that hearing loss influences the SNR necessary for a certain level of performance on a discourse comprehension task. The results of Hygge et al. (1992) also revealed an interaction between masker type and group. The significant interaction resulted from hearing-impaired listeners selecting the same SNRs for all masker types (SSN, male 1T and male reversed 1T), unlike the normal-hearing listeners. Thus, individuals with hearing loss were not impacted by masker type, but individuals with normal hearing were affected by masker type. This result is contrary to the results reported in the current study. In the current study, the significant interaction resulted from YNH listeners selecting the same SNRs for all masker types. It is possible that in the Hygge et al. (1992) study, listeners with normal hearing were able to take advantage of differences in voice pitch between the target talker (female) and the 1T masker (male) and hearing-impaired individuals did not receive this benefit, yielding similar SNR selections across maskers for the hearing-impaired group. In addition, the Hygge et al. study (1992) did not measure comprehension accuracy. Notably, listeners were simply required to "just follow the story." It is possible that the hearing-impaired listeners would have self-selected a more favorable SNR in the 1T condition compared to the SSN condition if more cognitive resources were required to interpret and accurately respond to the passage. Lastly, participants within each group in the Hygge et al. (1992) study ranged in age from 17 to 62 (M=43). It is possible that results from Hygge and colleagues were confounded by age effects within groups.

In the current study, the magnitude of the effect of masker type across the two older groups was also of interest. Older adults, even those with normal hearing, typically report difficulty understanding speech in realistic noisy environments. Therefore, it was important to assess performance in the presence of complex maskers (i.e., Cafe and 1T) compared to an SSN masker. A comparison between older adults with and without hearing loss across different masker types during a discourse comprehension task has not been examined previously. That OHI listeners selected higher SNRs across all masker types compared to ONH listeners in the current study suggests that hearing loss, in addition to age, impacts performance on a discourse comprehension task. However, there was no difference in the magnitude of the masker effect between the two older groups. In other words, both older groups were impacted to the same extent relative to each masker type during a discource comprehension task. Both older groups selected the most favorable SNR for the 1T masker, the next most favorable SNR for the Cafe masker, and the least favorable SNR for the SSN masker. Taken together, these results suggest that advancing age is the main contributor to the magnitude of change in self-selected SNR during a discourse comprehension task across different masker types. It is possible that the pattern of differences between masker types was similar between the two older groups because listeners in these groups were using similar compensatory strategies and top-down processing resources in order to select SNRs that led to equal and adequate comprehension accuracy performance across all masker types.

Impact of Speech Intelligibility Performance

It was hypothesized, based on previous literature, that speech intelligibility would not be a significant predictor of self-selected SNR (Hustad, 2008; Nagaraj, 2017; Smith & Pichora-Fuller, 2015). The LMER results revealed that speech intelligibility was not a significant predictor of self-selected SNR. Although the current study used a new technique, the results are consistent with those reported by Hustad, (2008), Nagaraj, (2017) and Smith and Pichora-Fuller, (2015). The results of the current study are also generally consistent with those of Murphy et al. (2006). These investigators presented younger and older normal-hearing adults with a discourse comprehension task at individually adjusted SNRs for speech intelligibility in the presence of a 2-talker spatially separated masker. Murphy and colleagues found that older adults performed more poorly than younger adults on the comprehension task, even when the stories were presented at individually adjusted SNRs for speech intelligibility. The results from Murphy et al. (2006) suggest that, in general, it is not possible to predict discourse comprehension from speech intelligibility performance alone. In addition to the Murphy et al. (2006) findings, the results from the current study suggest that factors other than the speech intelligibility of the target story modulate comprehension performance for older adults. However, due to the nature of the paradigm in the current study it is critical to consider additional explanations for the outcome. In the current study, speech intelligibility performance was not assessed at the same SNR across listeners. Speech intelligibility was only measured at each individual's self-selected SNR. Consequently, it is difficult to make firm conclusions about the relationship between relative speech intelligibility performance and self-selected SNR in this study and there may be alternative explanations to the LMER results.

Another possible explanation for the result is that speech intelligibility scores were limited in range relative to the large range in SNRs. Ordinarily, when there is a large range of SNRs, one would expect a relatively large range in speech intelligibility performance (i.e., a steep slope) (MacPherson & Akeroyd, 2014). However, in the current study, there was no relationship between speech intelligibility and self-selected SNR. It is conceivable that the systematic variation normally seen in speech intelligibility performance across a large range of SNRs was minimized as a result of the self-adjustment procedure. This possibility may suggest that listeners utilized the speech intelligibility of the signal to aid in their SNR selection. For example, if speech intelligibility was completely irrelevant to self-selection, then a listener with a very low SNR would show poor intelligibility. If speech intelligibility did not modulate the self-selected time-compression ratios, then we would expect to see a steeper slope between speech intelligibility and time-compression. If this notion is true, it would suggest that listeners used speech intelligibility to aid in their SNR selection, because the systematic variation in intelligibility performance was minimized. This notion somewhat supports findings from Schneider et al. (2000). These authors found that when stories during a discourse comprehension task were presented at individually adjusted SNRs equated for word intelligibility, discourse comprehension was preserved. This result implies that listeners' speech intelligibility modulates performance on a discourse comprehension task in the presence of 12-talker babble.

Due to the nature of the task in the current study, it is difficult to determine the true impact of speech intelligibility during the discourse comprehension task. In future studies, it would be beneficial to measure speech intelligibility performance at a variety of SNRs (i.e., higher and lower relative to the listener's selection during the story) that would yield fixed levels of speech intelligibility performance. This technique would provide insight into changes in speech intelligibility across a range of SNRs. This approach may provide insight into how different levels of speech intelligibility performance influence self-selected SNR during discourse comprehension performance. The fact that speech intelligibility was only measured at one SNR (i.e., the self-selected SNR) makes it difficult to generalize the results of the current study to a broader understanding of the relationship between speech intelligibility and discourse comprehension. Alternative speech intelligibility material should also be considered in future studies.

Impact of Cognitive Function

It was hypothesized that cognitive function, specifically working memory, would predict self-selected SNR. LMER results showed that working memory capacity was a significant predictor of self-selected SNRs across all groups and masker types. Working memory capacity proved to be the only cognitive measure that was significant and retained in the final LMER model. This result suggests that there is a relationship between self-selected SNR necessary to reach a comprehension accuracy performance near 80% correct and working memory capacity. Working memory was predicted to have a strong relationship with the self-selected SNR in a discourse comprehension task, because working memory and discourse comprehension require a similar set of temporary storage skills (Ward et al. 2016). Working memory requires the ability to manipulate and temporarily store incoming information (Baddeley & Hitch, 1974; Daneman & Carpenter 1980), and discourse comprehension requires this same set of processing skills. Although this result may not be surprising, previous literature aimed at assessing the relationship between working memory and discourse comprehension performance has yielded mixed results (Gordon et al., 2009; Nagaraj, 2017; Smith & Pichora-Fuller, 2015; Ward et al., 2016).

A study by Ward et al. (2016) found that RSPAN scores were correlated with discourse comprehension performance of vocoded one-minute passages for both younger and older adults with normal hearing. This result indicates that working memory capacity is related to discourse comprehension performance for degraded vocoded speech. The results of the current study are consistent with those of Ward et al. (2016), despite the fact that in the previous study, speech was degraded spectrally with vocoding, and in the current study, speech was degraded with masking. Results from the Ward et al. (2016) study and the current study also support predictions of the ELU Model, which suggest that speech processing that occurs under difficult listening situations interferes with bottom-up processing and requires listeners to rely more on top-down working memory resources (Rönnberg et al., 2013). The ELU model predicts that individuals with decreased working memory capacity will perform more poorly than individuals with higher working memory capacity on difficult listening tasks. Importantly, the ELU model was based on speech intelligibility performance, therefore the current study may provide evidence that the ELU model can be applied to discourse comprehension performance as well.

It was also anticipated that cognitive function would be more predictive of older listeners' performance compared to younger listeners. Results of the current study revealed no interaction between the cognitive measures and listener group, which suggests that cognitive function was equally predictive for all groups. Although older adults showed declines in all cognitive measures compared to younger adults and selected more favorable SNRs, individual differences in cognitive performance influenced self-selected SNRs similarly for all groups. This result is consistent with results from Ward et al. (2016) who found that working memory capacity impacted discourse comprehension equally for younger and older groups. Cognitive function was also hypothesized to have the strongest relationship to the self-selected SNR in the 1T condition, because increased cognitive resources are required for this condition compared to the SSN and Cafe masker conditions (Baddeley & Salame, 1983; Bell et al., 2008; Koelewjin et al., 2012; Schurman et al., 2014). However, the predictor variable of working memory was not involved in any interactions with masker type. This result suggests that during a discourse comprehension task, working memory was equally important in performing the task in the presence of all masker types tested in this study. Taken together, the impact of WM on self-adjusted SNR during a speech comprehension task was independent of listener group or masker type.

Self-selection Method

Listeners in all groups and across all masker types selected consistent SNRs over multiple trials. This result shows that the self-selection method produced reliable results and participants were systematically selecting preferred SNRs across trials and masker types. The self-selected SNRs consistently yielded performance near 80% correct for comprehension accuracy performance for all groups and masker types. This result suggests that listeners in all groups were able to select an SNR that allowed them to perform similarly on the comprehension accuracy task. In addition, the self-selection method was sensitive to differences between groups and masker types.

Conclusions

This study investigated the effects of masker type, age, hearing loss, speech intelligibility performance and cognitive function on self-selected SNRs during a discourse comprehension task. Results demonstrated that ONH and OHI listeners required a more favorable SNR across all masker conditions compared to YNH listeners. This suggested that age has an impact on the SNR required to perform a discourse comprehension task. In addition, hearing loss was shown to have an additional detrimental impact on discourse comprehension performance. The self-selection task revealed differences between all three groups, suggesting that this task may have clinical utility to identify individuals who have difficulty understanding speech in noisy environments. It was also hypothesized that the self-adjusted SNRs of older adults during discourse comprehension would be affected more by the different masker types than the SNRs selected by YNH listeners. The results showed that YNH listeners selected similar

SNRs across masker types, which suggests that during a discourse comprehension task, YNH listeners were not impacted by the presence of different maskers. Older adults selected different SNRs depending on the masker type. Older adults preferred the highest SNRs in the Cafe and 1T conditions compared to the SSN condition. This result implies that older adults with and without hearing loss are impacted to a greater extent compared to younger adults by realistic and single competing speech maskers. These findings are consistent with the speech intelligibility literature (Dubno et al., 1984; Tun et al., 2002); however, this study is the first to demonstrate that both older adults with and without hearing loss are impacted to a greater extent by masker type compared younger adults during a discourse comprehension task.

Speech intelligibility was also assessed in order to determine the relationship between intelligibility and discourse comprehension. The results showed that speech intelligibility was not a predictor of self-selected SNR. However, due to the nature of the study design it is difficult to conclude definitively that speech intelligibility was not a factor in the self-selection process. In future studies, it would be beneficial to measure speech intelligibility performance at a variety of SNRs to create a psychometric function around the listener's SNR selection during the discourse comprehension task (i.e., higher and lower selections during the story). This paradigm may provide insight into changes in speech intelligibility performance across a range of SNRs related to their preferred SNR during a discourse comprehension task. This is an important area for future research with the objective of quantifying the relative contributions of speech intelligibility during difficult and realistic discourse comprehension tasks. Lastly, it was hypothesized that cognitive function, specifically working memory, would predict self-selected SNRs. The

LMER results showed that working memory predicted performance equally for all listener groups and masker types. This result is an important finding that suggests working memory influences discourse comprehension performance in the presence of background noise for young listeners and older listeners with and without hearing loss.

General Discussion

The overall goal of this research was to identify the various individual and environmental factors that modulate discourse comprehension performance in difficult listening situations. Specifically, the goals of the current experiments were to: (1) determine the effects of age and hearing loss, and either speech presentation rate condition or masker type, on self-selected time-compression ratio or SNR, respectively, during a discourse comprehension task; (2) determine the relationship between speech intelligibility performance and either self-selected time-compression ratio or SNR; and (3) determine which cognitive domain (i.e., working memory, selective attention or speed of processing) is the most important predictor of self-selected time-compression ratio and SNR during a discourse comprehension task for young normal hearing adults and older adults with and without hearing loss.

Effects of Age and Hearing Loss on Discourse Comprehension in Degraded Conditions

Several consistent findings were observed across both experiments pertaining to the effects of age and hearing loss on self-selected time-compression ratio and SNR. First, older normal hearing listeners required slower time-compression ratios and more favorable SNRs across all conditions to reach 80% performance for comprehension accuracy compared to younger listeners. This result suggests that older adults with normal hearing are at a disadvantage compared to younger adults in the presence of rapid speech and background noise during a discourse comprehension task. These results are generally consistent with previous literature examining the impact of age on discourse comprehension and speech intelligibility under degraded conditions (Gordon-Salant & Fitzgibbons, 1993; Ward et al., 2016; Wingfield et al., 1999; Wingfield & Ducharme, 1999). However, the effect of hearing loss among older listeners has not been assessed in previous discourse comprehension studies. Results of the current study showed that older adults with hearing loss selected slower time-compression ratios and more favorable SNRs compared to older adults with normal hearing. This indicates that hearing loss, in addition to age, affects the rate and SNR at which OHI listeners can comprehend a spoken passage.

Taken together, these results may have real-world implications because they suggest that older adults, especially those with hearing loss, require a slower listening rate and a more favorable SNR compared to younger listeners in order to perform similarly on a comprehension accuracy task. Real-world conversation unfolds quickly and often in the presence of background noise. As a result, listeners must rapidly identify the spoken message, interpret the message, and encode it for later retrieval. This process must occur efficiently even in the presence of background noise in order for discourse comprehension to be successful. However, listeners who prefer slower listening rates or more favorable SNRs may have difficulty participating in meaningful communication under conditions where speech is produced rapidly and/or in the presence of noise.

Impact of Rapid Speech and Background Noise on Discourse Comprehension

Both studies aimed to identify the impact of degraded listening situations (i.e., rapid speech or background noise) on discourse comprehension performance. In the first study, comparisons between the speech presentation rate conditions (i.e., silences intact and silences removed) were used to evaluate the relative impact of information rate and articulation rate during a discourse comprehension task. It was hypothesized that given additional time to process the story during the silent periods, all listeners would be able to select a faster time-compression ratio in the silences intact condition compared to the silences removed condition. Results confirmed this hypothesis. This result suggests that to some extent, self-selected time-compression ratios were modulated by processing speed and/or the overall information rate of the signal, as well as the ability to benefit (i.e., select a faster listening rate) from the presence of silent periods in the silences intact condition for all groups. However, Study 1 also revealed a significant interaction in which OHI listeners had the largest difference in self-selected time-compression ratio between conditions compared to both normal-hearing groups. This result suggests that OHI listeners were more impacted by the available processing time or the overall information rate of the signal compared to the normal hearing groups. It is possible that individuals with hearing loss were compensating for declines in bottom-up processing by relying on top-down processing speed to modulate their time-compression ratio selections. Although the ELU model does not specifically focus on processing speed, the concept can be applied to these results. The ELU model suggests that increased demands on bottom-up processing (i.e., hearing loss in this case) require listeners to compensate by relying more on top-down cognitive resources (Rönnberg et al., 2013). It appears that the OHI listeners attempted to compensate for declines in bottom-up processing by utilizing the additional processing time available in the silent intervals. As a result, their performance showed that they could self-select a faster rate in the silences intact condition compared to the silences removed condition. If OHI listeners were unable to compensate for declines in bottom-up processing by utilizing the additional processing time when silent intervals were intact, then these listeners may have selected similar time-compression ratios across conditions. Study 2 measured the effects of masker type, age and hearing loss on self-selected SNR during a discourse comprehension task. Results showed a significant interaction between listener group and masker type. YNH listeners selected similar SNRs across all masker types, suggesting they were not differentially impacted by masker type. However, both older groups selected the highest SNR in the 1T condition compared to the other masker conditions.

Taken together, the studies suggest that different effects of aging and hearing loss emerge when using a self-selection procedure for time compression ratio versus SNR during a discourse comprehension task. Specifically, in Study 1 OHI listeners were impacted to a greater degree by rapid speech compared to both normal hearing groups. It is important to note that the time-compression task in Study 1 was presented in the presence of noise at +5 dB in order to make the task more realistic compared to quiet (Smeds et al. (2015). These results suggest that OHI listeners may be at a greater disadvantage compared to ONH listeners when speech is rapid and in the presence of noise (Study 1) versus a normal rate signal in the presence of noise (Study 2).

Impact of Speech Intelligibility on Discourse Comprehension

The effect of speech intelligibility performance on self-selected stimulus variables (time-compression ratio and SNR) during a discourse time compression task was measured across both studies. The goal was to determine if the speech intelligibility of the signal was a major factor contributing to self-selected time-compression ratios and SNRs during a discourse comprehension task. Therefore, speech intelligibility materials were presented at the self-selected time-compression ratios and SNRs chosen during the discourse comprehension task. Results from both studies revealed that intelligibility was not a significant predictor of self-selected time-compression ratio nor SNR during the discourse comprehension task. Previous work in this area, utilizing different techniques, has not shown a systematic relationship between discourse comprehension and speech intelligibility (Hustad, 2008; Murphy et al., 2006; Nagaraj, 2017; Smith & Pichora-Fuller, 2015). These previous studies have attempted to predict discourse comprehension percent correct performance from speech intelligibility percent correct performance and have found no relationship. The results of these prior studies suggest that traditional measures of speech intelligibility and discourse comprehension do not have a linear relationship. Therefore, a direct comparison between the percent correct performance on a speech intelligibility task and the percent correct performance on a discourse comprehension task may not be appropriate. Consequently, the current study took a different approach to examine the relationship between discourse comprehension and speech intelligibility.

In the current study, statistical modeling with LMER removed speech intelligibility performance from the final models of factors impacting discourse comprehension. However, the fact that speech intelligibility performance was removed from each model should not be interpreted to mean that speech intelligibility was irrelevant to the self-selection process. Speech intelligibility scores in both studies were somewhat limited in range relative to the large range in self-selected time-compression ratios and SNRs. Consequently, there was no significant relationship between the self-selection variables and speech intelligibility performance. This notion could suggest that when listeners self-selected their time-compression ratios and SNRs they also minimized the systematic variation in performance that is required of a significant predictor variable.

If true, this theory may imply that speech intelligibility performance played a role in modulating self-selected time-compression ratio and SNR.

Ultimately, due to the nature of the task in the current study it is difficult to definitively state the true impact of speech intelligibility performance on a discourse comprehension task. More work is needed in this area in order to gain a better understanding of how speech intelligibility is utilized by individuals with and without hearing loss during a discourse comprehension task. In the current study, only measuring speech intelligibility performance at self-selected time-compression ratios and SNRs makes it difficult to generalize the results to a broader understanding of the relationship between speech intelligibility and discourse comprehension. In future studies, it would be beneficial to measure speech intelligibility performance at a variety of SNRs and time-compression ratios to create a psychometric function around the listener's selection during the discourse comprehension task. This paradigm may provide insight into changes in speech intelligibility performance across a range of time-compression ratios and SNRs related to the listeners' selections during a discourse comprehension task.

Impact of Cognitive Abilities on Discourse Comprehension in Degraded Conditions

Studies 1 and 2 examined the impact of multiple cognitive domains on self-selected time-compression ratios and SNRs during a discourse comprehension task. It was hypothesized that cognitive function would predict performance on each of these two dependent variables. The hypotheses were informed by literature examining the relationship between speech intelligibility and cognitive function, which has been well documented across listener groups and masker types (Fullgrabe et al., 2015; Gordon-Salant & Cole, 2016; Pichora-Fuller et al., 2016; Rönnberg et al., 2013; Schurman et al., 2014). However, research on discourse comprehension performance and cognitive function has produced mixed results (Gordon et al., 2009; Smith & Pichora-Fuller, 2015; Ward et al., 2016), and the majority of previous discourse comprehension studies have only measured one cognitive domain (i.e., working memory capacity).

Results from Study 1 showed that processing speed and working memory influenced listeners' self-selected time-compression ratios. Listeners with faster processing speed and a higher working memory capacity selected faster time-compression ratios. This result, in conjunction with the significant impact of the silences intact condition, provides further evidence for the cognitive hypothesis, underscoring the importance of top-down cognitive processes on discourse comprehension rather than bottom-up processing of acoustic information. Results from Study 2 also showed that working memory capacity significantly influenced preferred SNR during a discourse comprehension task. When speech was presented at a normal rate in noise, working memory capacity was the only influential cognitive domain. Higher working memory capacity resulted in a more adverse SNR selection; that is, listeners with higher working memory could achieve the criterion performance score of 80% correct for discourse comprehension at less favorable SNRs than listeners with lower working memory. Both studies therefore reinforce the importance of working memory capacity on discourse comprehension performance over a range of degraded speech conditions (i.e., rapid speech and speech in the presence of background noise) and listener groups. Lastly, listeners' performance on the Flanker task (i.e., measure of attention/inhibition) was not a significant predictor of either self-selection variable. However, the null result should not be interpreted to mean that attention and inhibition are not involved in discourse comprehension under degraded auditory conditions. The result may simply suggest that the Flanker task did not accurately capture the selective attention abilities required to complete the self-selection task.

Clinical Implications

It is difficult to quantify the impact of multiple individual and environmental factors on discourse comprehension performance because discourse comprehension is challenging and time consuming to measure. The current study took a unique approach to quantifying discourse comprehension performance using a technique that has the potential to be used clinically and appears to be sensitive to the unique communication difficulties of individuals from different listener groups. A method of self-selection was used in both of the current studies. Listeners self-selected a preferred time-compression ratio and a preferred SNR across multiple conditions. All listener groups were reliable when selecting a time-compression ratio or an SNR across multiple trials. The self-selection method yielded a wide range of time-compression ratios and SNRs across individuals and conditions with relatively fixed comprehension accuracy performance. This is comparable to a speech intelligibility task where listeners achieve a fixed percent correct performance over a large range of individual SNRs. The self-selection method in both of the current experiments was successful in revealing differences across groups and conditions. Reliability and sensitivity are essential features of a clinical measure.

Older adults, even those with normal hearing, often perform well on standard clinical speech intelligibility tasks, but report difficulty understanding speech in real-world situations. A discourse comprehension task may be more representative of daily communication than a standard speech intelligibility task. The results of the current experiments suggest that the use of a discourse comprehension task may provide insight into the difficulties that older adults with and without hearing loss experience during real-world communication. The task presented in the current experiments could also be used to evaluate hearing aid benefit and the relative utility of different hearing aid algorithms. For example, self-selected SNR during a discourse comprehension task could be assessed with and without hearing aids in order to document hearing aid benefit on a task similar to everyday communication requirements. Alternatively, this discourse comprehension task could be used to document a lack of hearing aid benefit. Often, patients are not satisfied with their hearing aids in real-world listening environments, despite showing significant hearing aid benefit on a traditional speech intelligibility task in the clinic. Evaluating speech comprehension using a realistic task may be more sensitive in showing the true difficulties experienced by older adults with and without hearing loss. Lastly, if a hearing aid algorithm could improve or lower the self-selected SNR on this type of task, it is possible that the algorithm would have the potential to improve comprehension in real-world environments.

References

ANSI (2018). ANSI S3.6-2018, "American National Standard Specification for Audiometers (R2018)," American National Standards Institute, New York.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley A, Logie R, Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. *J Mem Lang, 24,* 490-502.

Baddeley, A. D. & Salame´, P. (1986). The unattended speech effect: Perception or memory? *J Exp Psychol Learn, 12*, 525–529.

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

Bell, R., Buchner, A., & Mund, I. (2008). Age-related differences in irrelevant-speech effects. *Psychol Aging, 23*, 377–391.

Bell, D.W., Kruel, E.J., & Nixon, J.C. (1972) Reliability of the modified rhyme test for hearing. *J Speech Hear Res, 15,* 287–295.

Best, V., Keidser, G., Freeston, K., & Buchholz, J. M. (2016). A Dynamic Speech Comprehension Test for Assessing Real-World Listening Ability. *J Am Acad Audiol, 27*(7), 515-526.

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

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*, 14(5), 377–381.

Brookshire, R, H. & Nicholas, LE. (1993). The Discourse Comprehension Test. Tucson, AZ: Communication Skill Builders, A Division of The Psychological Corporation.

Brown, J. I., Carlsen, G. R., Carstens, L. (1995). Brown-Carlsen-Carstens Listening Comprehension Test. New York, NY: Harcourt, Brace and World.

Brungart, D.S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am, 109*(3), 1101-1109.

Brungart, D. S., Simpson, D. B., Ericson, A. M., & Scott, R. K. (2001). Informational masking and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am, 110,* 2527–2538.

Boersma, P. & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.34, http://www.praat.org.

Committee on Hearing Bioacoustics and Biomechanics. (1988). Speech understanding and aging. *J Acoust Soc Am, 83,* 859-895.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *J Verb Learn Verb Be, 11,* 671-684.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *J Verb Learn and Verb Be, 19,* 450-466.

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

Eriksen, B, A., & Eriksen C.W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept & Psychophys, 16,* 43–149.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*, **87,** 1725–1736.

Fontan, L., Tardieu, J., Gaillard, P., Woisard, V., & Ruiz, R. (2015). Relationship Between Speech Intelligibility and Speech Comprehension in Babble Noise. *J Speech Lang Hear Res, 58*(3), 977-986.

Fullgrabe, C., Moore, B.C., & Stone, M.A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Front Aging Neurosci, 6,* 347.

Freyman, R.L., Balakrishnan, U., & Helfer, K.S. (2004) Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J Acoust Soc Am, 115,* 2246–2256.

Friedman, P. N. & Miyakem A. (2005). Comparison of four scoring methods for the reading span test. *Behav Res Methods, 37,* 581-590.

Gershon, R.C., Wagster, M.V., Hendrie, H.C., Fox, N.A., Cook, K.F., & Nowinski, C.J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology, 80*(11 Suppl 3), S2-6.

Gordon, M. S., Daneman, M., & Schneider, B. A. (2009). Comprehension of speeded discourse by younger and older listeners. *Exp Aging Res, 35*(3), 277-296.

Gordon-Salant, S. (2005). Hearing loss and aging: New Research findings and clinical implications. *J Rehabil Res Dev, 42(4),* 9-24.

Gordon-Salant, S., & Cole, S.S. (2016). Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear Hear, 37*(5), 593-602.

Gordon-Salant, S. & Fitzgibbons, P.J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *J Speech Hear R, 36,* 1276–85.

Gordon-Salant, S. & Fitzgibbons, P.J. (2001). Sources of age-related recognition for time compressed speech. *J Speech Hear R*, 44, 709-719.

Gordon-Salant S, & Fitzgibbons, P.J. (2004) Effects of stimulus and noise rate variability on speech perception by younger and older adults. *J Acoust Soc Am, 115,* 1808–17.

Goossens, T., Vercammen, C., Wouters, J., & van Wieringen, A. (2017). Masked speech perception across the adult lifespan: Impact of age and hearing impairment. *Hear Res*, *344*, 109–124.

Guerreiro, M. J., Murphy, D. R., & Van Gerven, P. W. (2013). Making sense of age-related distractibility: The critical role of sensory modality. *Acta Psychol*, 142, 184–194

Hasher, L., & Zacks, R. T. (1988). *Working memory, comprehension, and aging: A review and a new view*. In G. H. Bower (Ed.), The Psychology of Learning and Motivation, Vol. 22. (pp. 193–225) San Diego, CA: Academic Press.

Helfer, K. S., & Freyman, R. L. (2008). Aging and speech-on-speech masking. *Ear Hear*, *29*(1), 87-98.

Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using singletalker competition. *J Acoust Soc Am*, 120, 2926–2937.

Hustad, K. C. (2008). The Relationship Between Listener Comprehension and Intelligibility Scores for Speakers With Dysarthria. *J Speech Hear R, 51*, 562–573.

Hygge, S., Rönnberg, J., Larsby, B., & Arlinger, S. (1992). Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speed, reversed speech, and noise backgrounds. *J Speech Hear R, 35(1),* 208-215.

Janse, E. (2009). Processing of fast speech by elderly listeners. *J Acoust Soc Am*, *125*(4), 2361-2373.

Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J. (2009) Articulation rate across dialect, age, and gender. *Lang Var Change*, 21, 233–256.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, NJ: Prentice-Hall, Inc.

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

Kayser, S. J., Ince, R. A., Gross, J., & Kayser, C. (2015). Irregular Speech Rate Dissociates Auditory Cortical Entrainment, Evoked Responses, and Frontal Alpha. *J Neurosci, 35(44),* 14691-14701.

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, *J Acoust Soc Am*, *116(4)*, 2395–2405.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psyc Rev, 95(2),* 163-182.

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a singletalker masker. *Ear Hear*, 33, 291–300.

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2014). The influence of informational masking on speech perception and pupil response in adults with hearing impairment. *J Acoust Soc Am, 135,* 1596–1606.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology*, *25*(6), 763-770.

MacPherson, A., & Akeroyd, M. A. (2014). Variations in the slope of the psychometric functions for speech intelligibility: a systematic survey. *Trends ear*, *18*, 1-26

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones, *Speech Commun*, 9, 453–467.

Murphy, D. R., Daneman, M. & Schneider, B. A. 2006. Why do older adults have difficulty following conversations? *Psychol Aging*, 21**,** 49-61.

Nagaraj, N. K. (2017). Working Memory and Speech Comprehension in Older Adults With Hearing Impairment. *J Speech Lang Hear Res*, *60*(10), 2949-2964.

Nasreddine, S., Phillips, A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc, 53*, 695–699.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J Acoust Soc Am*, *97*, 593–676.

Piquado, T., Benichov, J. I., Brownell, H., & Wingfield, A. (2012). The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening. *Int J Audiol*, *51*(8), 576-583.

Presacco, A., Zimon, J., & Anderson, S. (2016). Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol, 116(5),* 2356-2376.

Quene, H. (2007). On the just noticeable difference for tempo in speech. *J Phon*, 35, 353-362.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Curr Dir Psychol Sci*, *17,* 177–182.

Rönnberg J, Arlinger S, Lyxell B, & Kinnefors C. (1989). Visual evoked potentials: Relations to adult speech reading and cognitive function. *J Speech Hear R*, 32, 725-735.

Rönnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B. . . . Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front Syst Neurosci*, *7*, 31.

Rudner, M., Lunner, T., Behrens, T., Thorén, E. S., & Rönnberg, J. (2012). Working memory capacity may influence perceived effort during aided speech recognition in noise, *J Am Acad Audiol, 23(8),* 577–589.

Salthouse, T. A. (1985). *A theory of cognitive aging*. Amsterdam: North-Holland.

Salthouse, T. A., Toth, J., Daniels, K., Parks, C., Pak, R., Wolbrette, M., & Hocking, K. J. (2000). Effects of aging on efficiency of task switching in a variant of the trail making test. *Neuropsychology, 14*(1), 102-111.

Salthouse, T. A. (2010). Is Flanker-Based Inhibition Related to Age? Identifying specific influences of individual differences on neurocognitive variables. *Brain Cogn*, 73(1), 51-61.

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence*, 39(4), 222-232.

Schneider B.A., Avivi-Reich M. & Daneman M. (2016). How Spoken Language Comprehension is Achieved by Older Listeners in Difficult Listening Situations. *Exp Aging Res*, 42, 31-49.

Schneider, B. A., Daneman, M., & Murphy, D. R. (2005). Speech comprehension difficulties in older adults: cognitive slowing or age-related changes in hearing? *Psychol Aging, 20*(2), 261-271.

Schneider, B.A., Daneman, M., Murphy, D.R., & Kwong See, S. (2000). Listening to discourse in distracting settings: The effects of aging. *Psychol Aging, 15*, 110–125

Schneider, B. A., Pichora-Fuller, M. K., & Daneman, M. (2010). *The effects of senescent changes in audition and cognition on spoken language comprehension.* In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), Springer Handbook of Auditory Research: The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis (pp.167-210). New York: Springer.

Schurman, J. Brungart, D. & Gordon-Salant, S. (2014). Effects of masker type, sentence context, and listener age on speech recognition performance in 1-back listening tasks. *J. Acoust. Soc. Am, 136(6),* 3337–3349.

Smith, S. L. & Pichora-Fuller, M. K. 2015. Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss. *Front Psychol*, 6**,** 1394.

Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of Signal-to-Noise Ratios in Realistic Sound Scenarios. *J Am Acad Audiol, 26*(2), 183-196.

Sommers M., Hale, S., Myerson, J., Rose, N., Tye-Murray, N., & Spehar, B. (2011). Listening Comprehension Across the Adult Lifespan. *Ear Hear*, 32, 775–781.

Souza, P.E. & Turner, C.W. (1994). Masking of speech in young and elderly listeners with hearing loss. *J Speech Hear Res, 37,* 655-661.

Tun, P. A. (1998). Fast noisy speech: Age differences in processing rapid speech with background noise. *Psychol Aging,13,* 424–434.

Tun, P. A., O'Kane, G., & Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychol Aging*, 17, 453–467.

Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., Hale, S., & Rose, N. (2008). Auditory-visual discourse comprehension by older and young adults in favorable and unfavorable conditions. *Int J Audiol*, 47 (Suppl 2), S31–S33.

Vaughan, N., Storzbach, D., & Furukawa, I. (2006). Sequencing vs. non-sequencing working memory in understanding of rapid speech by older listeners. *J Am Acad Audiol, 17,* 506–514.

Voeten, C. C. (2019). Using 'buildmer' to automatically find & compare maximal (mixed) models, *R package version*, *1*(6), 1-7.

Ward, C. M., Rogers, C. S., Van Engen, K. J., & Peelle, J. E. (2016). Effects of Age, Acoustic Challenge, and Verbal Working Memory on Recall of Narrative Speech. *Exp Aging Res*, *42*(1), 97-111.

Wiley, T.L., Cruickshanks, K.J., Nondahl, D.M., Tweed, T.S., Klein, R., & Klein, B.K. (1998). Aging and word recognition in competing message. *J Am Acad Audiol, 9,* 191–198.

Wingfield, A., Amichetti, N., & Lash, A. (2015). Cognitive aging and hearing acuity: modeling spoken language comprehension. *Front Psyhchol, 6*:684.

Wingfield, A. & Ducharme, J. (1999). Effects of Age and Passage Difficulty on Listening-Rate Preferences for Time-Altered Speech. *J Gerontol, 54B***,** 199-202.

Wingfield, A., Peelle, J. E., & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging Neuropsychol Cogn, 10(4),* 310–322.

Wingfield A, Poon LW, Lombardi L, & Lowe D. (1985). Speed of processing in normal aging: effects of speech rate, linguistic structure, and processing time. *J Gerontol*, 40, 579-585

Wingfield A. & Tun P.A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *J Am Acad Audiol, 18,* 548–558.

Wingfield A., Tun P.A., Koh C.K. & Rosen M.J. (1999). Regaining Lost Time: Adult Aging and the Effect of Time Restoration on Recall of Time-Compressed Speech. *Psychol Aging*, 14, 380-389.

Zekveld, A., Rudner, M., Kramer, S. E., Lyzenga, J., & Rönnberg, J. (2014). Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Front Neurosci, 8,* 1–11.