EFFECTS OF AGING ON BEHAVIORAL MEASURES OF LISTENING EFFORT

Kimberly Skinner

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

in the Department of Speech and Hearing Sciences,

Indiana University

April, 2020

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

____

Larry Humes, PhD

____

Jennifer Lentz, PhD

____

Karen Forrest, PhD

____

Jason Gold, PhD

March 11, 2020

Acknowledgements

I would like to express my deep appreciation for the financial, academic, and technical support provided to me and kindness extended to me by the members of my research committee and the Department of Speech and Hearing Sciences at Indiana University. I am eternally grateful for the opportunity to pursue a doctor of philosophy degree in speech and hearing sciences. I would also to express appreciation to the Council of Academic Programs in Communications Sciences and Disorders for awarding me their PhD scholarship.

I wish to thank each member of my research committee: Larry Humes, PhD, Jennifer Lentz, PhD, Karen Forrest, PhD, and Jason Gold, PhD. I appreciate your mentorship, guidance, generous giving of your time, and sharing of your knowledge more than words can express. Each of you provided me with skills, knowledge, and experiences that have allowed me to grow professionally. Thank you. Thank you, Larry, for chairing my research committee and for your many many hours spent giving me feedback as I developed, implemented, and wrote up this research project.

I am also thankful for my PhD cohort and others I met during my time living in Bloomington. You are the icing on the cake of my wonderful experience at Indiana University. Finally, a loud thank you to my daughter, Hannah Skinner, for being part of my academic adventures.

Kimberly Skinner

EFFECTS OF AGING ON BEHAVIORAL MEASURES OF LISTENING EFFORT

Difficulty with speech communication in noise is a common problem among elderly individuals. Older adults often report challenges with understanding speech, particularly in noisy environments. Growing evidence suggests that cognitive effort is a significant factor in speech understanding in noise. Although hearing loss is commonly experienced by older adults, according to prevalence estimates, about 4 in 10 adults age 65 and over will have impaired hearing. Older adults also experience decline in a number of cognitive abilities. The focus here was on aging alone to eliminate hearing loss as a contributing factor. The primary focus of this study was to measure cognitive effort (listening effort) in young and older adults with normal hearing while completing a speech in noise task.

This study also examined some methodological issues for the measurement of listening effort. The most common means of behavioral assessment of listening effort is through use of a dual-task paradigm (DTP), whereby participants perform a “primary” speech-perception task along with a “secondary” task that does not involve speech perception. The two tasks can be administered concurrently or sequentially. It is not known whether DTPs administered sequentially and concurrently in the same person will yield similar results. The primary task in the DTP used here was a speech-identification task with a target talker and two competing talkers; the secondary task was either concurrent or sequential recall of a portion of the target message. Another methodological issue examined was the influence of the gender of the competing talker, either the same as (male) or different from (female) the target talker.

The primary finding was that, when the performance of young and older adults was equated at baseline in the DTP, few effects of age on listening effort were seen. Differences between the concurrent and sequential conditions emerged, however, including a larger dual-task effect on the secondary task, slower response times, and poorer performance overall for the sequential condition. Consistent with previous findings in the literature, performance on the speech segregation portion of the DTP was generally better when the genders of the target and competing talkers differed.

____

Larry Humes, PhD

____

Jennifer Lentz, PhD

____

Karen Forrest, PhD

____

Jason Gold, PhD

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

Difficulty with speech communication in noise is a common and long-observed problem among elderly individuals (e.g., Carhart, 1946; Carhart & Tillman, 1970). Older patients, with and without hearing loss, commonly report that they can hear but have difficulty understanding what is being said, particularly in noise (as discussed in recent reviews by, e.g., Pichora-Fuller et al., 2016 and Gagne, Besser, & Lemke, 2017). Growing evidence suggests that cognitive effort increases as the speech signal is degraded or the coding of that speech signal is impaired (Pichora-Fuller et al., 2016). Recently, a framework for understanding effortful listening (FUEL) was developed whereby the interaction between peripheral and cognitive factors during speech understanding is considered (Pichora-Fuller et al., 2016). Consensus is growing that increased listening effort can be attributable, at least in part, to the allocation of additional cognitive resources to the processing of degraded auditory input. For example, according to a well-established model, the Ease of Language Understanding model (Rönnberg, 2003; Rönnberg, Rudner, Foo, & Lunner, 2008; Rönnberg et al., 2013), increased cognitive resources are employed when a linguistic signal is degraded either through external (e.g., noise) or internal (e.g., hearing loss) factors. This increased cognitive effort to support speech processing takes away from limited cognitive resources that would be otherwise employed for higher-level linguistic processing during conversation and could lead to mental fatigue among other problems. If cognitive resources are directed to understanding degraded auditory input (from hearing loss, background noise, or a combination of both), fewer cognitive resources remain for semantic processing, recall, or formulating a reply during conversation (Hornsby, 2013; Pichora-

Fuller et al., 2016). The concept of listening effort is also promising to clinicians and researchers as a potential measure for hearing aid evaluation (Lunner, Rudner, Rosenbom, Agren, & Ng, 2016).

Whereas a number of individual studies support the hypotheses that a) listening effort is increased with sensorineural hearing loss and b) use of hearing aid amplification reduces this effect, in a recent review of scientific literature on the subject of listening effort, Ohlenforst et al. (2017a) found no conclusive evidence across all studies to strongly support these hypotheses. These authors reviewed studies that measured listening effort through self-report (use of surveys or rating scales), behavioral assessment, and physiologic measures (evoked potentials or pupillometry). Ohlenforst and colleagues found evidence of moderate quality from physiologic measurements to support the hypothesis that hearing loss increases listening effort. On the other hand, studies employing either self-report or behavioral measurements differed greatly in terms of experiment parameters (e.g., participant characteristics, speech stimuli, nature of the secondary task) and were of low statistical power; the authors did not find quality evidence from these types of studies supporting the hypothesis that listening effort increases with hearing loss. They concluded that there is need for listening effort studies to be more consistent with one another and to have sufficient statistical power.

In this study, given the sparse evidence from behavioral paradigms for the impact of both aging and hearing loss on listening effort, together with the likely implementation of such measures clinically in the near future, if valid, we were interested in further evaluation of the behavioral assessment of listening effort. According to prevalence estimates, about 4 in 10 adults age 65 and over will have impaired hearing. The older adult with impaired hearing is both chronologically old and hearing impaired. Older adults also experience decline in processing-related cognitive abilities such as working memory and processing speed. As a result, the older adult with impaired hearing may have deficits in both hearing and cognitive processing that may require increased cognitive effort when listening to speech in noise. To simplify this, the focus here was on the effects of aging alone on listening effort. This was accomplished by comparing the performance of young and older adults with normal hearing.

The most common means of behavioral assessment of listening effort is through use of a dual-task paradigm (DTP), whereby participants perform a “primary” speech-perception task along with a “secondary” task that does not involve speech perception. The amount of listening effort required to perform a speech-perception task is reflected in performance on the secondary task compared to performance on that secondary task alone (baseline performance). Secondary tasks that have been used to evaluate listening effort include probe reaction time tasks (e.g. Downs 1982), memory tasks (e.g. Rakerd, Seitz, & Whearty, 1996; Hornsby 2013), and pursuit tracking tasks (e.g. Tun, McCoy, & Wingfield, 2009; Desjardins & Doherty, 2014; Xia et al., 2015).

The primary speech-perception task can be administered concurrently with the secondary task or in a sequential manner. In a concurrent experimental design, participants complete the speech-perception task while also completing the secondary task at the same time; secondary task characteristics vary as described above. In a *sequential* design, sometimes referred to as a “pre-load” design, participants are presented (visually or auditorily) with linguistic material (e.g. letters or digits) for later recall (e.g. Rakerd et al., 1996;). After the presentation of the pre-load material, a speech-recognition measure is administered. Following response on the speech-recognition task, participants must recall the stimuli originally presented. To date, much remains unknown about the various methodological factors taken into consideration when designing a

DTP; even more so when older adults are considered the target population. For example, it is currently not known whether concurrent and sequential DTPs measure the same underlying cognitive processes taking place during speech perception and understanding (Ohlenforst et al., 2017a). Additionally, the relationship between listening effort measurements obtained in a concurrent paradigm and in a sequential paradigm are unknown for young adults, making it difficult to compare results across studies (Gagne, Besser, & Lemke, 2017). Further, if different underlying processes *are* being assessed with each method, it is not known which measurement paradigm is the most sensitive measure of listening effort both for young and older adults. Consider, for instance, just the selection of sensory modalities involved. On the one hand, if both primary and secondary tasks use the same sensory modality, there may be more competition for cognitive resources (Gagne et al., 2017). Alternatively, given that secondary tasks using visual or haptic modalities also demonstrate increases in listening effort, it may be that multitasking in general is sufficiently cognitively demanding to measure listening effort and the specific parameters, including modality, of the secondary task are less important (Gagne et al., 2017).

Ultimately, we are interested in the impact of age-related hearing loss on listening effort. However, those with age-related hearing loss, by definition, are both old and hearing impaired. Considering aging alone, it is well known that cognitive function in healthy aging declines steadily across the adult lifespan (e.g., Craik & Byrd, 1982; Salthouse, 2004). Age-related declines in auditory and cognitive abilities have already begun in middle age or around an age of 50 years (Helfer & Freyman, 2014). To what extent increased listening effort seen in older adults is attributable to sensory decline rather than cognitive decline may be difficult to pinpoint. Given the sparse information available on the effects of aging alone on listening effort, this was the focus here. The research proposed here will compare performance from the same participants, one group of young adults and one group of older adults, both with normal hearing, on both a concurrent and a sequential dual-task paradigm.

Chapter 2. Background Literature Review

The focus of this dissertation is on age-related changes in listening effort, as measured using a DTP. This chapter provides more general background information about a number of topics relevant to this focus. First, a brief review of the effects of aging on hearing is presented, followed by general overviews of speech communication and measures of cognitive abilities. The effects of aging on cognition are then reviewed briefly. The chapter concludes with a more detailed review of the state of knowledge regarding listening effort.

Audiometric assessment

Even though the focus here is on young and older adults with normal hearing, it is important to recognize that many older adults have measurable hearing loss. This is important to know for procedural purposes, and also when generalizing results from this study to older adults. Hearing sensitivity is measured by assessing hearing thresholds at each ear using headphones, while seated in a sound-treated room. The hearing threshold is defined as the softest sound level a person can hear and is estimated following standard clinical procedures. Hearing thresholds are typically measured at the frequencies 250 Hz through 8,000 Hz at octave intervals. The hearing threshold at each frequency is plotted on an audiogram. Hearing thresholds are considered to be within normal limits when the threshold at each frequency is 20 dB HL or better (Jerger & Jerger, 1980).

Hearing Loss in Older Adults

The prevalence of hearing loss increases with age. Figure 1 shows the prevalence of hearing loss by decade from age 50 to 80 and older based on data from the National Health and Nutritional Examination Surveys between 2001 through 2008 (Lin, Niparko, & Ferucci, 2011).

The incidence of age-related hearing loss (ARHL) increases with age as well. In a large, population-based, longitudinal study, Cruickshanks et al. found an overall 5-year incidence of hearing loss of 21% for adults age 48-92 years. When sorted by decade of life, the incidence increased from 11.6% between the ages of 48-60 to 95.5% between the ages of 80-92 years (Cruickshanks et al., 2003).

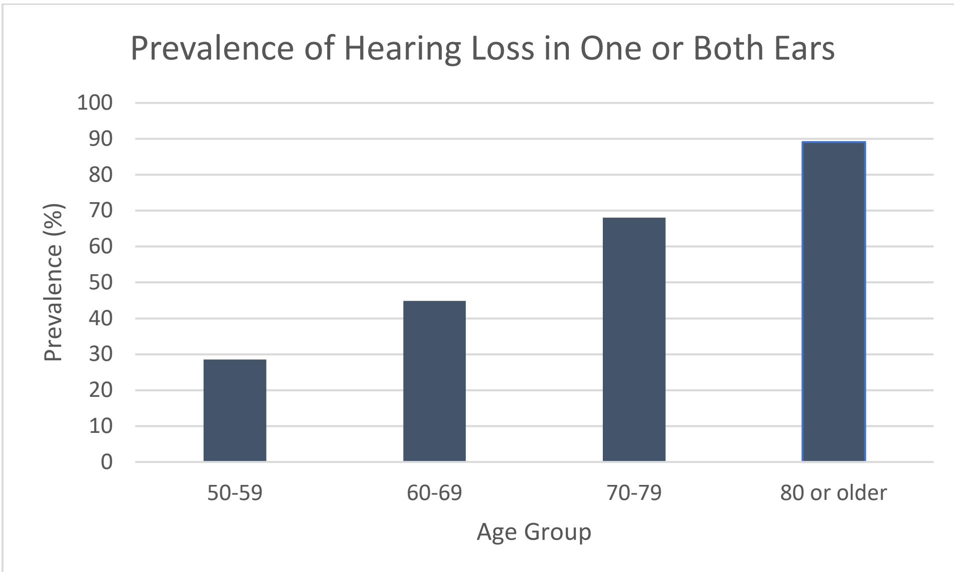

Figure 1. Prevalence of hearing loss in one or both ears by decade of life. Based on the National Health and Nutritional Examination Surveys between 2001 through 2008 (Lin et al., 2011).

Speech communication

The ability to understand speech is affected by ARHL. For most people, recognizing spoken speech is more difficult in the presence of background noise. In the audiology clinic, a patient's word-recognition ability is measured in a quiet setting by delivering a recorded word list and noting the percentage of words the patient is able to correctly recognize. Typically, in ARHL, word-recognition scores in quiet are fairly high. Despite the clinical utility of measuring word-recognition ability in quiet, an individual's ability to communicate in noise is of greater interest in terms of everyday function and eventual aural rehabilitation. As noted previously, difficulty hearing in noisy situations is a common complaint of adults with hearing loss. Speech understanding in noise is measured either using sentences or words presented simultaneously with competing noise. The competing noise can be steady-state noise (white noise) or can be fluctuating noise. Fluctuating noise can be comprised of recorded talkers (ranging from one talker to multiple talkers) or can consist of artificially generated stimuli that sound “speech-like” but do not contain actual speech. Speech- understanding ability in quiet does not perfectly correlate with speech-understanding ability in noise (Pichora-Fuller & Singh, 2006), although there is typically a moderate correlation between these abilities (Humes et al., 1994; Humes, 1999; Humes & Dubno, 2010). In other words, two individuals with similar hearing abilities for understanding speech in quiet can perform somewhat differently on tests of word-recognition ability in noise. As the competing “noise” becomes more speech-like, moreover, the correlations with performance in quiet decrease and large individual differences among older listeners emerge (Humes & Dubno, 2010; Humes, Kidd & Lentz, 2013). To some extent, this difference can be accounted for by looking at the individual’s audiogram: high frequency hearing thresholds tend to be predictive of word recognition ability in noise (Humes & Dubno, 2010; Wilson, 2011). However, a person’s cognitive abilities also factor into speech-communication performance in noise. In a large study evaluating speech understanding in older adults with hearing loss, Humes and colleagues (2013) found that cognitive factors, specifically working memory and global speed of processing, were predictors of speech understanding difficulty in noise after differences in hearing thresholds were accounted for (Humes et al., 2013).

Masking and interference

Experiments show that the masking effects from competing talkers (as opposed to steady-state noise) are largely "informational masking" rather than "energetic masking" (Brungart, 2001; Brungart, Simpson, Ericson, & Scott, 2001). When informational masking occurs, both the target and the competition are audible to the listener, but the listener is not able to differentiate the target from the competition. When energetic masking occurs, overlap of competing signals renders the target inaudible to the listener (Brungart et al., 2001). It is likely that either type of masking could be occurring during a speech task involving competing talkers; what is of interest here is the fact that the competing voices *interfere* in varying degrees with identification of the target (Humes, Kidd, & Fogerty, 2017).

Cognitive abilities and how they are measured

Prior to further discussion how cognitive factors relate to speech communication in noise, a brief overview of cognitive abilities and how they are measured is presented. It should be noted that terms relating to cognitive abilities can sometimes be used differently by different scientists and in some cases do not have consistent definitions. An example of how an ability is measured is given below, but multiple measures for each cognitive ability have been developed.

Working memory refers to the ability to store and process information; the greater the working memory span, the greater number of items can be maintained and manipulated in memory (Baddeley, Eysenck, & Anderson, 2015). Working-memory span is assessed using either simple or complex span tasks (Unsworth & Engle, 2007). In a simple span task, sometimes called a short-term memory task, a participant is given a list of items such as words, digits, or shapes, and is tasked with recalling them in order immediately after the last item is presented. In a complex span task, participants also recall items in the order presented, but are also simultaneously involved in an unrelated cognitive task (Unsworth & Engle, 2007). An example of a complex span task is the Reading Span Test (Daneman & Carpenter, 1980) where participants read a series of sentences and are instructed to recall the final word of each sentence and then to repeat those final words after reading the series of sentences. In a meta-analysis evaluating 22 studies, Unsworth and Engle (2007) concluded that simple and complex span tasks mostly measure the same processes involved in working memory.

Attention can refer to the ability to attend to a desired target or it can refer to allocation of limited cognitive resources to a particular task (Desimone & Duncan, 1995). It is typically measured by presenting a group of items that differ in some dimension and instructing the subject to attend to the item meeting a particular description. For example, a subject may be instructed to report on all the black letters when presented with an image containing black and red letters.

Inhibition refers to the ability to inhibit extraneous information when responding to target stimuli. Inhibition is commonly measured with a variation of the Stroop test (Stroop, 1935), whereby participants are instructed to name text (a color name, for example "red") printed either in the same color as the text (the word "red" printed in red) or in a different or incongruous color (the word "red" printed in yellow). Correctly named items and reactions times on same versus incongruous trials yield a measure of inhibition. Many variations of this test have been created and adapted over the years.

Processing speed is typically measured in reaction time. For behavioral measures, there is often a motor response of some type required and the time required to execute this response may be included or excluded from the measure of processing speed.

Executive control or *executive function* typically refers to a set of cognitive abilities that include planning, adapting to rapidly changing situations, monitoring behavior, and include some or all of the above abilities. In fact, the term can be confusing when not fully defined in a scientific work. The means by which this ability is measured therefore depends on how the researcher defines the term.

Each of these cognitive abilities is thought to play a role in speech understanding in noise. In a survey of 20 experimental studies, no one cognitive measure stood out as a predictor for speech communication difficulty, but measures of working memory were mostly effective in predicting speech recognition in noise ability (Akeroyd, 2008; Humes & Dubno, 2010). Working memory has been shown to be significantly related to speech-recognition performance in noise (Desjardins & Doherty, 2013). Besser and colleagues observed that a larger working-memory span seems to be advantageous in a variety of listening situations (Besser, Koelewijn, Zekveld, Kramer, & Festen, 2013). Recall that Humes et al. (2013), in a battery of tests using aided speech-recognition tests, found that the speech-recognition differences among older listeners were accounted for by cognitive differences rather than auditory or age factors (Humes et al., 2013)

Age-related changes in cognition and individual differences

Speech perception difficulties experienced by older listeners can be attributed not only to hearing loss but also to factors involving the cognitive processes outlined above. Older adults experience declines in these cognitive abilities (Roberts & Allen, 2016). Cognitive abilities in older adults vary quite a bit but some general patterns are seen across studies. Investigators have found a strong association between cognitive measures and speech in noise performance.

Cognitive factors account for aided speech in noise difficulty (e.g. Humes et al., 2013; Füllgrabe, Moore, & Stone, 2014; Houtgast & Festen, 2008; van Rooij, Plomp, & Orlebeke, 1989) once peripheral auditory abilities are taken into consideration, up to 2/3 of systemic variance for aided speech understanding in noise (Humes, 2007; Humes et al., 2013). Cognitive factors involved in speech-understanding difficulties experienced by older adults, especially for speech in a background of competing speech, arise from individual differences in cognitive ability (Humes & Dubno, 2010; Humes et al., 2013).

Moore and colleagues (2014) analyzed data from the UK Biobank on over 40,000 individuals as part of a national longitudinal study in Britain. Speech in noise performance was measured by the digit triplets test (DTT) where each stimulus consisted of three single-syllable digits presented in steady-state spectrally shaped noise. Cognitive measures were the best predictors of speech perception in noise (Moore et al., 2014).

Processing speed, working memory, and inhibition were shown to play a role in speech understanding in noise in a study by Helfer and Freyman (2014). They compared these cognitive measures and the ability to understand speech in a variety of background maskers among young, middle aged, and older listeners. Older groups and younger groups differed significantly on working memory, short term memory, inhibition, and processing speed tasks, as expected; declines were evident in the group of middle-aged listeners as well as in the older group. The individual differences in performance were most strongly associated with cognitive function across the adult lifespan.

Effortful Listening

Clinically, many patients with hearing loss often report that listening to speech in noise is effortful, and many patients with hearing loss report fatigue when listening in noise (Hornsby, 2013; Bess & Hornsby, 2014). Scientific consensus is growing that increased "listening effort" can be attributable, at least in part, to the allocation of additional cognitive resources to the processing of degraded auditory input. According to a well-established model, the Ease of

Language Understanding model (Rönnberg, Rudner, Foo, & Lunner, 2008), increased cognitive resources are employed when a linguistic signal is degraded either through external (e.g., noise) or internal (e.g., hearing loss, decreased cognitive function) factors. This increased cognitive effort to support speech processing takes away from limited cognitive resources that would be otherwise employed during conversation and could lead to mental fatigue among other problems (Hornsby, 2013; Pichora-Fuller et al., 2016). The concept of listening effort is also promising to clinicians and researchers as a potential measure for aural rehabilitation planning (Pichora-Fuller et al., 2016) and hearing aid evaluation (Lunner et al., 2016).

One way to evaluate listening effort is through a dual-task paradigm (DTP). In a DTP, two tasks are administered, whereby one task is designated the "primary task" and the other, the "secondary task". A participant completes each task individually to establish baseline single-task performance, and then simultaneously. The idea is that a drop in performance on the secondary task relative to baseline performance reflects a degree of effort required to complete the primary task. This method is based on the premise that the cognitive system is limited in capacity (Kahneman, 1973). The logic underlying the dual-task paradigm method is that if cognitive resources are being allocated to the primary task, performance on the secondary task will decline. For example, if a secondary task of tracking a moving object on a computer screen is employed, accuracy on that task will be measured at baseline (without performing the primary task), and then will be measured while also completing a primary speech (in the case of speech and hearing research) task; results can be compared within subjects across conditions or between subjects across different subject types. The difference in performance on the secondary task at baseline and while concurrently performing the primary task is considered the "cognitive cost" or "dual task penalty" of performing the primary task. In a dual-task paradigm where the primary task was recall of a word list and the secondary task was tracking a moving object on a computer screen, Tun et al. (2009) found that the cognitive cost of recalling a word list was greater for older adults with and without hearing loss than it was for young adults with and without hearing loss. In other words, even younger adults with hearing loss exhibited a lower cognitive cost to recall the word list than older adults with good hearing. They also found that older adults with hearing loss showed the greatest cognitive cost of all; greater cognitive resources were employed to process degraded auditory input in the older hearing-impaired group.

Dejardins & Doherty (2013) investigated cognitive abilities, listening effort as measured with a dual-task paradigm, and speech recognition in young normally hearing adults and older adults with and without hearing loss. The speech task was to identify speech targets (sentences) in a variety of background maskers (two-talker, six-talker, and speech-shaped noise). The secondary task was a Digital Visual Pursuit Rotor Tracking test. They found that working memory span and processing speed were correlated with speech performance scores, and that older adults demonstrated greater effortful listening than younger adults in both the two-talker condition and the speech-shaped noise condition. In this particular study, there was not a significant difference in listening effort between the older adults with or without hearing loss.

In a recent study, Ward and colleagues (2017) took a slightly different approach and used noise-vocoded speech (rather than speech in noise) and a visual tracking task to measure listening effort in older adults with normal or very mild hearing loss and young normal-hearing controls. The speech stimuli were sentences degraded using the noise-vocoding method, and were presented with 8, 6, or 4 channels of vocoding. Vocoding is often used to simulate the effects of processing by a cochlear implant and the fewer the number of channels of vocoding, the more the speech is distorted and less well it is perceived. The secondary task was a visual monitoring task, where participants were instructed to watch consecutively presented images on a computer monitor and indicate by key press when the same image appeared twice in a row. The performance measure for this task was reaction time to key press. They also measured inhibition, using the Flanker test, where participants viewed a series of arrows that were either congruent or incongruent in direction and were tasked with naming the direction of the arrow (Zelazo, Anderson, Richler, Wallner-Allen, Beaumont, & Weintraub, 2013). Ward and colleagues found age-related differences in executive control accounted for age-related differences in listening effort. They also found that age was a factor more so than executive function ability in the most difficult (4-channel) condition. The authors suggested that these age-related performance differences in the most difficult condition might be attributable to peripheral factors such as temporal envelope sensitivity or to other cognitive factors such as processing speed (Ward, Shen, Souza, & Grieco-Calub, 2017).

Imaging studies can be used to better understand how listeners process speech in noise. During speech in noise tasks, activation of prefontral areas and precueneus regions (regions associated with working memory and attention) are seen (Roberts & Allen, 2016) and are associated with better behavioral performance on speech tasks than adults who do not show this pattern of cranial activity(as discussed in Roberts & Allen, 2016). A model, hemispheric reduction of asymmetry in older adults (HAROLD; Cabeza, 2002), has been observed for a number of processing phenomena and is considered to be a compensatory mechanism (Cabeza, 2002). Briefly, the HAROLD model explains that processes that are seen to elicit asymmetrical brain activation in young adults are seen to be more symmetrical, or reflect recruitment of additional brain regions, in older adults. Whether this is compensatory for decreasing cognitive or sensory ability (or both) is not clear. However, this recruitment of additional brain regions does appear to enable older adults to perform at similar levels to younger adults (Roberts & Allen, 2016). In a functional imaging study, Peelle and colleagues (2011) found that peripheral hearing acuity predicted the neural response to speech. They interpreted this finding as supporting a "resource allocation framework" whereby individual sensory acuity predicts recruitment of particular brain regions and the degree to which these regions are recruited (Peelle et al., 2011). This study provides further evidence that the manner in which cognitive resources are employed during speech-understanding tasks varies with the amount of stimulus degradation, in this case, degradation from hearing loss.

To what extent increased listening effort is attributable to sensory decline rather than cognitive decline may be difficult to pinpoint. It is not known whether aging adds an additional cognitive deficit in addition to sensory loss (Roberts & Allen, 2016). When both sensory decline and cognitive decline are experienced, it is possible that listening effort increases even further to achieve the goal of speech understanding (Humes & Young, 2016). Because of the fairly high prevalence of hearing loss and declines in cognitive function in the older population, many older adults could be experiencing the effects of both cognitive and sensory decline. For this study, we recruited older adults with normal hearing sensitivity to eliminate age-related hearing decline as a variable.

Processing degraded auditory input

For a listener, speech information is often degraded either because of background noise or because of hearing loss, or because of both occurring at the same time. Processing degraded auditory input increases the "cognitive load" of speech understanding and requires recruitment of additional cognitive resources (Pichora-Fuller & Singh, 2006; Fulton et al., 2016), resources which are considered to be limited (Kahneman, 1973) and would otherwise be used for language processing, memory, or attention.

The ability to recall word lists presented in background noise is impaired even in young, normally-hearing adults (Rabbitt, 1968; Pichora-Fuller & Singh, 2006). The introduction of noise increases cognitive load and this effect is measurable as a reduction in correctly recalled words. In an experiment where individual speech-in-noise thresholds were taken into account, the ability to recall materials presented in noise was the same for younger and older adults (Schneider, Daneman, Murphy, & Kwong See, 2000). In other words, when adjusting the background noise to equate those having different peripheral hearing ability, aging influences were not significant. Many studies have reported decline in comprehension in older adults. However, these studies have not controlled for hearing loss. In many studies, apparent age-related declines in comprehension may be mediated by hearing ability rather than comprehension deficits per se. Both young and older adults perform more poorly on a primary task when also performing a secondary task, in other words, when attention is divided.

Listening effort

Measures of listening effort are of interest because individuals report experiences of effortful or fatiguing listening even when speech is audible and clearly understood (Pichora-Fuller et al., 2016). In other words, it is possible to maintain high levels of accuracy in terms of speech understanding while exerting a large amount of cognitive effort to do so. In this way, measures of word recognition give information about the extent to which a person can accurately understand speech but do not give any information about the cognitive effort required to attain that level of performance. A better understanding of the listening effort expended while processing speech may help audiologists to better tailor amplification approaches (Pichora-Fuller & Singh, 2006) and plan aural rehabilitation strategies (Pichora-Fuller et al., 2016).

Listening effort in older adults

As noted above, Tun et al. (2009) and Dejardins & Doherty (2014) included young and older adults with normal hearing in their studies and found that older adults demonstrated increased listening effort on a DTP relative to the younger adults. In addition to these two studies, Meister, Rahlmann, and Walger (2018), using a DTP, found that even low levels of background noise affected the ability of older adults with normal hearing to maintain high levels of performance on a secondary task. Smith, Pichora-Fuller, and Alexander (2016) developed an auditory working memory task that simultaneously assesses word recognition ability and working memory (Word Auditory Recognition and Recall Measure; WARRM). The task requires participants to identify a word presented in quiet, then determine whether the word started with a letter from the first or second half of the alphabet. They were subsequently required to recall the words in varying set sizes. The task of making a judgment about the starting letter of the word adds to the processing complexity of the overall task. In a study with young, normally-hearing adults and older adults with and without hearing loss, the researchers found that the younger adults performed best on the recall task, and the older adults with normal hearing performed better than older adults with hearing loss.

Sequential vs. concurrent task presentation

DTPs have been implemented in research on listening effort in a variety of ways. For a primary speech task, researchers have used both word and sentence recognition in noise (e.g. Picou & Ricketts, 2014; Desjardins & Doherty, 2014), competing phrases (e.g. Xia, Nooraei, Kalluri, & Edwards, 2015), or a task involving recall of auditorily presented words (e.g. Tun, McCoy, & Wingfield, 2009). As previously mentioned, non-speech (secondary) tasks that have been used to evaluate listening effort include probe reaction time tasks (e.g. Downs 1982), memory tasks (e.g. Rakerd et al., 1996; Hornsby 2013), and pursuit tracking tasks (e.g. Tun, McCoy, & Wingfield, 2009; Desjardins & Doherty, 2014; Xia et al., 2015).

The primary speech-perception task can be administered concurrently with the secondary task or in a sequential manner. In a *concurrent* experimental design, participants complete the speech-perception task while also completing the secondary task. In a *sequential* design, the secondary task is a recall task. This design is sometimes referred to as a “pre-load” design, participants are presented (visually or auditorily) with linguistic material (e.g. letters or digits) for later recall. After the presentation of the pre-load material, a speech-recognition measure is administered. Following response on the speech-recognition task, participants must recall the originally presented stimuli. The large majority of DTP experiments in the literature employ a concurrent paradigm. In a recent review of behavioral measures of listening effort, Gagne and colleagues (2017) found only one peer-reviewed study, the work of Rakerd et al., 1996, utilizing a true sequential paradigm. In that study, participants were presented with digits for later recall, followed by a speech in noise task that involved running discourse (Rakerd et al., 1996). The concurrent dual task paradigm allows the researcher to examine processing resources in addition to memory (Gagne et al., 2017). It has been suggested that the concurrent paradigm holds a higher level of ecological validity (Gagne et al., 2017) though situations do arise where listeners desire to recall information while also carrying out a speech task. The extent to which cognitive shifting would differently affect performance between a concurrent or a sequential DTP is unknown. Conceivably, a higher performance cost related to cognitive shifting would be seen in a concurrent DTP, where the study participant would need to switch between the speech perception task and the secondary task at the same time. By comparison, in a sequential DTP, a participant can dedicate full attention to the pre-load material prior to completing the speech perception task.

Listening effort and subjective ratings of effort

In a number of studies, ratings of effort have been used as a subjective outcome measurement to evaluate the impact of hearing loss on speech understanding ability or evaluate hearing instrument features (e.g., Feuerstein, 1992; Humes, Christensen, Bess, & Hedley-Williams, 1997; Bentler & Duve 2000). In a recent work, Picou & Ricketts (2018) explored the relationship between speech recognition, listening effort expended, and subjective ratings of effort. They found that whether or not self-ratings of listening effort correlate with behaviorally measured listening effort depended on the way the question was asked. These researchers asked older adults with symmetrical hearing loss to complete a DTP and provide ratings for the following: mental work; desire to give up; desire to improve the listening situation; and tiredness. They found that high ratings of a desire to improve the listening situation correlated with poor performance on the secondary task. They found that high ratings of mental work expended correlated with poor word recognition (primary task) performance. They found no correlation between task performance and ratings on a desire to give up or ratings of tiredness. In the current study, listeners were asked to rate their desire to improve the listening situation following a clinical speech in noise assessment and also throughout the DTP experiment.

Age differences in listening effort are important to identify in a reliable and valid manner. Although the dual-task paradigm is the consensus paradigm for behavioral measures of listening effort, a number of factors have varied across studies that make it difficult to interpret the results. For example, the similarity of the stimuli and modalities across the primary and secondary tasks has varied. Both concurrent and sequential paradigms have been used, with most researchers using a concurrent paradigm; it is unclear if the two approaches yield similar results in the same individuals. Even when hearing loss has been eliminated as a factor in various ways, observed DTP performance differences between young and old adults may not reflect age per se but could still be due to other age-related degradations impacting performance on the primary task alone.

The research described here compared performance on a DTP where the primary and secondary tasks were completed concurrently and where the primary and secondary tasks are completed sequentially. In order to keep the tasks as similar as possible across conditions, a memory task was selected as the secondary task, whereby the stimuli to be recalled could be presented either concurrently or prior to the speech task. Recorded Coordinate Response Measure (CRM; Bolia, Nelson, Ericson, & Simpson, 2000) materials were chosen for the primary speech task. Each CRM sentence follows the same format and is comprised of the word "ready", followed by a call sign, which identifies the target sentence, the words "go to"', and a color-number combination, followed by the word "now" (example: "*Ready* **Baron** *go to* **green four** *now*") (Bolia et al., 2000). The CRM corpus contains 256 sentences spoken by each of 4 male and 4 female talkers with combinations derived from 8 different call signs, 4 colors, and the numbers 1-8. The listener is instructed to listen for the call sign ("Baron" in the current example) and select the corresponding color-number combination from a set of choices presented visually. These tests assess a listener's ability to identify a two-word target in the presence of similarly-constructed speech competition (Humes et al., 2013). These speech stimuli are desirable for our purposes for a few reasons. First, even though the speech materials are closed-set, the fact that there are 32 possible color-number combinations significantly reduces the guess rate. Second, there are minimal learning effects with these materials (Eddins & Liu, 2012) and they can be used for multiple trials, in contrast with other sentence-like stimuli. Third, because of the unchanging structure of the sentences, there is little or no predictability based on linguistic cues (Eddins & Liu, 2012). Further, the use of multiple talkers allows for comparisons across different types of speech competition, such as 1, 2 or 4 competing talkers having a gender the same or opposite the target talker (Humes et al., 2017). And finally, single digits, a part of the response, are well-suited to a recall task. Additionally, self-report effort ratings were obtained to allow comparisons of DTPs in the same subject, YNH and ONH alike.

In terms of the dual-task paradigm, one methodological concern is that the participant must give priority to the primary task in order for the results to be interpretable (Gagne et al., 2017). However, it is not always known how the participant has prioritized their attention. In an effort to determine whether participants the primary or the secondary task in the current research, scores on both tasks were analyzed.

Speech segregation and fundamental frequency

This experiment explores effects of competing voices on the ability to identify a target, both in terms of accuracy and effort. As previously discussed, competing talkers in the CRM task interfere in varying degrees with identification of the target (Humes et al., 2017). One acoustic cue long shown to aid in speech segregation is fundamental frequency (f_o) (Arehart, King, & McLean-Mudgett, 1997; Humes, Lee & Coughlin, 2006; Lee & Humes, 2012). Sentence combinations were presented in two experimental conditions: male target with male voice competition and male target with female voice competition. We expected to see a relative reduction of listening effort when the male talker is presented with female voice competition for all listeners.

Chapter 3. Methods

The general purpose of this study was to compare performance on a dual-task paradigm (DTP) administered concurrently and sequentially, in young and older normally-hearing adults. The primary task in the DTP was a speech-segregation task; the secondary task was a digit-recall task. Details about participant selection, stimuli, apparatus, and procedures are presented in this section.

Participant Selection

A total of 41 participants were recruited into the study in two groups: young, normally hearing (YNH) adults, aged 18-25 years; and older, normally hearing (ONH) adults, aged 50-69. Three of the ONH participants failed to meet inclusion criteria (two failed based on hearing status; one failed based on cognitive status) and did not complete any tasks beyond initial screenings. Two of the YNH and one of the ONH participants did not complete the study because of illness or scheduling difficulties. This left a total of 35 participants (17 YNH, 1 male, 15 females, 1 undeclared; 18 ONH, 1 male, 17 females) who completed the experimental procedures. The mean age in the YNH group was 20.9 years (sd = 1.4 y). The mean age in the ONH group was 60.4 years (sd = 4.7 y). Participants were recruited from previous studies in the Indiana University Department of Speech and Hearing Sciences (SPHS), from flyers posted in SPHS, from Indiana University's online classified advertisements, and via social media postings. Participants were paid $10/hour for participation.

Inclusion and Exclusion Criteria

Pure-tone air- and bone-conduction thresholds at octave frequencies from 250 through 4000 Hz were required to be 20 dB HL or better at each ear and word recognition scores for a standard clinical measure (CID W-22 monosyllables in quiet presented 40 dB above their speech-recognition threshold) needed to be 80% or better in each ear. Participants were excluded if a conductive hearing loss was evident, if there was suspected retrocochlear pathology in either ear based on the participant's case history or audiometric thresholds, or if hearing thresholds were asymmetric by more than 15 dB at two or more frequencies. Tables 1 and 2 show the mean word recognition scores and hearing thresholds for each group. Note that the only frequencies tested were the octave frequencies from 250 – 4,000 Hz. Although both groups are referred to as having normal hearing, thresholds were significantly poorer for the older adults at 1,000 Hz and 4,000 Hz at the right ear and at 4,000 Hz at the left ear (two-sample t-test; $p < .01$ in all cases) Participants were required to have a score ≥ 26 on a dementia screen, the MMSE Mini Mental State Exam (MMSE V2; Folstein, Folstein, White, & Messer, 2010) (described below), to be included in the study. Potential participants were excluded if they had a history of neurological disorder, speech and language disorder diagnosed within the past five years, constant tinnitus, or were non-native speakers of English. Testing was discontinued and the participant was not enrolled in the study if they failed to meet any of the inclusion criteria.

Group	WRS R	WRS L
YNH	97.41 (3.14)	96.35 (2.67)
ONH	98.11 (2.79)	96.21 (3.12)

Table 1. Mean word recognition scores (WRS) for each ear (R = right ear; L = left ear) for the young (YNH) and older (ONH) groups. (Standard deviations in parentheses.)

	250 R	500 R	1000 R	2000 R	4000 R	250 L	500 L	1000 L	2000 L	4000 L
YNH	9.23 (4.00)	7.35 (3.59)	6.18 (3.32)	8.24 (3.93)	4.41 (3.91)	9.62 (3.80)	8.53 (3.86)	8.53 (6.32)	8.24 (2.46)	6.76 (3.93)
ONH	8.85 (7.12)	9.17 (6.70)	10.28 (5.28)	9.44 (6.39)	11.94 (8.77)	10.77 (8.13)	9.44 (7.45)	10 (5.69)	10.56 (5.91)	14.17 (8.09)

Table 2. Mean hearing threshold values (dB HL) at octave frequencies from 250 through 4,000 Hz and mean word recognition scores (WRS) for each ear (R = right ear; L = left ear) for the young (YNH) and older (ONH) groups. (Standard deviations in parentheses.)

Materials and stimuli

Mini-Mental State Exam

Cognitive status of all potential participants was screened using the MMSE (Folstein et al., 2010). The MMSE is a well-established dementia screening tool used to identify cognitive disability. It is administered by asking participants a series of questions designed to measure general cognitive status.

Connected Speech Test

The Connected Speech Test (CST; Cox, Alexander & Gilmore, 1987) was used to assess word-recognition ability in noise. This test is comprised of ten-sentence passages on a familiar topic, which is made known to the participant ahead of time, administered in the presence of multi-talker babble. Each passage contains 25 key words for scoring. Target sentences and competing babble were presented monaurally to the right ear.

Spatial Short-Term Memory (SSTM)

The Spatial Short-Term Memory (SSTM) test (Lewandowsky, Oberauer, Yang, & Ecker, 2010) was administered to obtain a visually based measure of memory span that does not rely on hearing sensitivity. Because we plan to include participants with hearing loss in future experiments, it was desirable to utilize a measure that would not be confounded by the presence of hearing loss. In addition, even though both the young and older listeners have hearing designated to be normal at and below 4000 Hz, the two groups differ slightly in hearing sensitivity over this same frequency range. The SSTM is a subtest of the Working Memory Capacity battery developed by Lewandowsky and colleagues (Lewandowsky et al., 2010) and is a simple span task. A series of circles is presented on a 10 x 10 grid in set sizes ranging from 2-6 (with five trials of each); the participant is tasked with recalling the spatial relations between the circles by touching the touchscreen for the grid cells that had been occupied by circles after they've been removed from the screen. Figure 2 shows an example screen with a set size of five circles. The participant would have seen each of these circles presented sequentially, and then would have been tasked with the recalling the spatial location of the set of five circles.

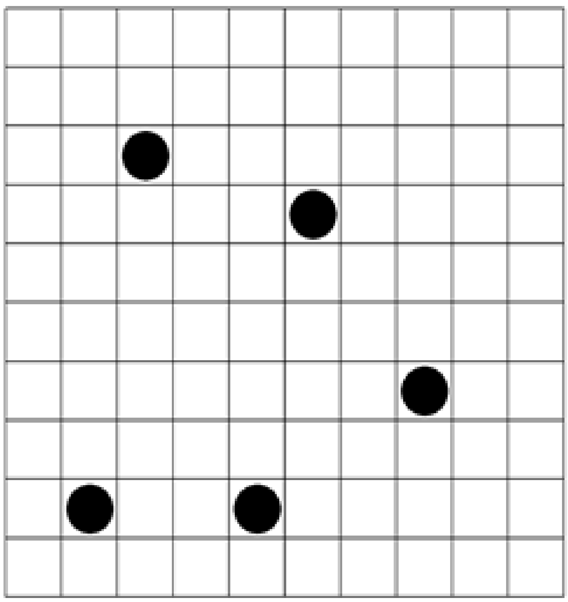

Figure 2. An example screen from the Spatial Short Term Memory task (SSTM) based on Lewandowsky et al., 2010.

CRM- speech segregation task

Recorded Coordinate Response Measure (CRM; Bolia et al., 2000) materials were used for the speech-segregation task. As discussed in the previous chapter, each CRM sentence follows the same format and is comprised of the word "ready", followed by a call sign, which identifies the target sentence, the words “go to”’, and a color-number combination, followed by the word “now” (example: “*Ready* Baron *go to* green four *now*”) (Bolia et al., 2000). The listener was instructed to listen for the call sign (“Baron” in the current example) and select the corresponding color-number combination (“green four” in example) from a set of choices presented visually. Sentences were presented initially in quiet to familiarize the participant with the task. During the experimental conditions, stimuli consisted of a total of three different talkers, one target (always male and identifiable by the Baron call sign) and two competing talkers. The competing talkers were either both male talkers or both female talkers, and always used a call sign other than “Baron”.

The term target-to-competition ratio (TCR) is used here to refer to the level of the target sentence relative to each individual competing sentence (Humes et al., 2017; Brungart et al., 2001). Because there were two competing talkers, and the TCR refers to the relationship between the target and each competing talker *individually*, the sound intensity of the competition is twice that for single-talker competition. In other words, the overall signal-to-noise ratio (SNR) in this study is always 3 dB poorer than the TCR. For example, with all three speech stimuli, one target and two competitors, set to the same sound level individually, the TCR is 0 dB but the SNR is -3 dB.

Listening Effort Rating Scale

After each section of the CST, and at regular intervals throughout the experimental conditions, participants were asked the following question: “On a scale of 1 to 10, please rate your desire to improve the listening situation.” The wording of this question was chosen based on the work of Picou, Moore, & Ricketts (2017) and Picou & Ricketts (2018), who found that a question probing a listener’s desire to control a listening situation correlated with behavioral measures of listening effort. Picou and colleagues posed the question, “How likely would you be to try to do something else to improve the situation (e.g., move to a quiet room, ask the speaker to speak louder)?” Participants in the current study were given examples of improving the listening situation if they asked for clarification. Participants were informed ahead of time that they would be asked to provide this rating and were given written instructions that identified a rating of “1” as being “not at all” and “10” as a lot. Ratings were obtained initially when administering the CST to establish an individual baseline measure of listening effort; the same question and rating scale were used throughout the experimental conditions.

Initial measures

Testing took place over the course of three or four sessions. Consent was obtained and initial screening measures were completed at the beginning of the first session. A brief oral case history was taken, followed by the administration of the MMSE. Otoscopy was performed at each ear. Tympanograms were measured bilaterally using a Grason-Stadler Instruments (Eden Prairie, MN) GSI39 tympanometer. All audiometric screening and assessment measures, as well as experimental protocols, were conducted in a sound-treated room that met or exceeded ANSI guidelines for permissible noise levels for audiometric testing (ANSI, 1999). Pure-tone testing was completed using a calibrated Grason-Stadler Instruments (Eden Prairie, MN) GSI61 audiometer and ER-3A (Etymotic Research, Inc., Elk Grove Village, IL) insert earphones. Word recognition ability in quiet was evaluated using CID W-22 word test (Hirsh, Davis, Silverman, Reynolds, Eldert, & Benson, 1952) recorded word lists (one half-list at each ear). If the participant met all the inclusion criteria, testing continued as described subsequently.

Word recognition in noise

The Connected Speech Test (CST, Cox et al., 1987) was used to assess speech-recognition ability in noise. This test is comprised of ten-sentence passages on a familiar topic, which is made known to the participant ahead of time, administered in the presence of multi-talker babble. Target sentences and competing babble were presented monaurally to the right ear at a +2 dB signal-to-noise ratio, with target sentences presented at 63 dB SPL. Two sets of passages were administered. Scores were obtained for each passage and an average score was computed.

Subjective rating of listening effort

After each passage of the CST, participants rated their listening effort using the rating task described above. They were asked the following question: "On a scale of 1 to 10, please rate your desire to improve the listening situation." Because two passages were used, a geometric mean of the two ratings was calculated.

Visual Short-Term Memory

The Spatial Short-Term Memory (SSTM) test (Lewandowsky et al., 2010) was administered using a program developed by Lewandowsky and colleagues designed to run with PsychToolbox (Brainard, 1997; Pelli, 1997) and MATLAB version 2013a (32-bit). Participant responses were collected using a stylus and a touch screen. Responses were automatically scored based on how closely the response pattern matched with the presented pattern, with partial credit given for responses that deviated by only one cell on the grid. A proportion correct score was calculated based on the total possible number of points.

Speech perception task (primary task)

Recorded Coordinate Response Measure (CRM; Bolia, et al., 2000) materials were used as the primary speech task. For the current study, three male talkers (Talker 1 was the target voice with Talkers 0 and 2 as competition) and two female talkers (Talkers 6 and 7) were used, with the full set of 32 color-number combinations used for each of these speakers. Participants were instructed to listen for the call sign in the target sentence ("Baron") and select the corresponding color-number combination from a set of choices presented visually.

Stimuli were presented by computer using Tucker Davis Technologies System-III hardware (RP2.1 24-bit capable D/A converter, 48828-Hz sampling rate, HB7 headphone buffer) delivered via Etymotic 3-A insert earphones. In all cases, stimuli were delivered to the right ear only. The left insert earphone was disconnected but was placed in the ear canal to attenuate any unwanted sounds. Intensity was normalized to a root-mean-square (RMS) pressure level of 77 dB SPL, which reflects average levels for a noisy conversational setting (Killion, 1997) and was tested at 0 dB TCR. The highest anticipated presentation levels, using +9 dB TCR, were verified as resulting in peaks of 93 dB SPL. These are levels that would be considered safe levels for the estimated duration of presentation. Equipment was checked daily and monthly using noise files generated to match the acoustic characteristics of the speech stimuli. Acoustic calibration was performed on a monthly basis with a Larson Davis (Depew, NY) Model 800 sound level meter and Model 2575 microphone fitted with a Bruel & Kjaer (Nærum, Denmark) DB-0138, 2-cm^3 coupler. Calibration voltage at the output of the HB7 was measured with a Fluke (Everett, WA) Model 45 multimeter and Phillips (Amsterdam, The Netherlands) Model 3335 Oscilloscope. The calibration voltage was verified with the multimeter daily prior to subject testing.

The experiment was administered using a specially designed MATLAB program. For all experimental tasks, participants initiated trials and gave responses by selecting objects on the touch screen using a stylus. The participant initiated each block by tapping on the word "begin" when they were ready. After each sentence presentation, a column containing all four colors and a column containing all eight numbers was shown on the computer screen and the participant selected the color and number corresponding to the target sentence with the call sign "Baron". After making their selection, the participant selected "ok" to begin the next trial. This self-pacing of trials was designed to minimize penalties for either age group with the use of a fixed inter-trial interval. Sentences were initially presented with no competing talkers to orient the participant to the nature of the stimuli and to ensure understanding of the task. Following initial practice, sentences were presented with two competing male talkers at a +12 dB target-to-competition ratio (TCR) to further orient the participant to the nature of the task. The same male voice (Talker 1) and call sign ("Baron") were used as the target sentence for the CRM throughout the experiment.

Baseline target-to-competition ratio

Target-to-competition ratios (TCR) for the speech task performed as a single task were adjusted to estimate 79.4% performance accuracy following procedures described below. This TCR is referred to here as the "baseline TCR" for each participant, meaning that this is the TCR at which the participant completed the task administered as a single task at 79.4% accuracy. This performance level was chosen to yield commonly experienced speech-to-noise ratios (SNRs) (Smeds, Wolters, & Rung, 2015; Wu et al., 2018) and is also a reasonably high level of performance; a desired feature for the DTP measure. It has been shown that listening effort *decreases* at SNRs producing 50% correct performance levels (Ohlenforst et al., 2017b; Wu, Stangl, Zhang, Perkins, & Eilers, 2016), a performance level used commonly for speech-in-noise testing. It is possible that when a listening situation is too effortful, the participant does not try as hard as they would at a more favorable SNR and therefore do not expend large amounts of effort. The TCR yielding 79.4% performance was estimated in two steps using the CRM with the competition comprised of two male talkers. First, a one-up, three-down adaptive procedure (Levitt, 1971) was used to approximate the desired range of TCRs for each listener. Then, a more precise estimate of baseline was obtained using the method of constant stimuli, presenting 20 trials at the TCR estimated in the first step, and 20 trials each at +6, +12, -6, and -12 dB TCR relative to the TCR estimated in step one for a total of 100 trials. A best-fitting 4-parameter Weibull function was then determined for these data and the TCR corresponding to 79.4% performance was used throughout the experiment.

Recall Task (secondary task)

The same CRM sentences as used for the speech perception task were used for the recall task. Here, however, the participants were asked to identify and then later recall the number in the color-number target. After each sentence was presented, the participant selected the number from a column containing all eight digits. Then, after the full set of sentences was presented, a different screen appeared containing the same number of columns as the set size being tested. The participant was tasked with recalling the digits in order, selecting, for example, the first digit in the set from the first column, the second digit in the set from the second column, and so on. The set size varied by block, and the participant was told the set size at the start of each block. As with the speech segregation task, sentences were initially presented with no competing talkers to ensure understanding of the task. Following initial practice, sentences were presented with two competing male talkers at a +12 dB TCR to further orient the participant to the nature of the task.

Estimate baseline memory span

The set size for which a participant was able to correctly recall 80% of sets within a block of ten sets was estimated at the TCR corresponding to 79.4% performance. The individual memory span was measured adaptively. Initial practice trials were presented in sets of two with no competing talkers, and then in sets of four, with the competing talkers set at a high (+12 dB) TCR. After the initial practice, participants began this task with a set size of five. If performance for the five-item task was 80% or better, the set size was increased, making the task more difficult. If performance was below 80%, the set size was decreased, making the task easier. This process continued until the participant completed at least one set size below 80% correct and at least one set size at 80% or better. Once a set size was identified where the participant recalled approximately 80% of the sets correctly, two additional confirmatory blocks were presented at that same set size. When performance on one of the confirmatory blocks caused the average performance (of all blocks at that set size) to be lower than 80%, the set size was then decreased and additional blocks at the new set size were presented. Ultimately, a total of three blocks were presented at the estimated memory span, with the average scores of these three blocks 80% or better. Typically, the procedures described thus far were all completed at the initial session, which lasted up to two hours. In some cases, the protocol to estimate baseline memory span was initiated at the first session and was completed at the beginning of the second session.

Dual-task experimental conditions

For each of the dual-task paradigms described below, sentence combinations were presented in two conditions: male target with two male voice competitors and male target with two female voice competitors. The contrasting competing-sentence conditions manipulate the segregation of the target male talker from the competing talkers such that less effort should be required for the competing female speech due to improved sound segregation of the male target from the competition. Prior to each experimental condition, participants completed practice trials in sets of two with no competition, and then in sets of four with the TCR set at +12 dB.

Concurrent Dual-Task

CRM trials were presented at the baseline TCR with set sizes corresponding to the participant's individual baseline memory span. The participant was instructed to respond by selecting the color and number heard from the target voice after each sentence was presented. After a set of presentations of these sentences, the participant was then prompted to recall the target number from each presentation.

A total of 120 sentences were presented, broken up into either three or four blocks, depending on set size. See Figure 3, below, for the timing of the secondary (recall) task relative to the primary (speech perception) task.

Sequential Dual-Task

As with the concurrent task, CRM trials were presented at the baseline TCR. However, for the sequential dual-task method, a full set was comprised of two sets of sentences. For the first half of the set, the participant was tasked with identifying only the numbers from the target voice for later recall. For the second half of the set, the participant was tasked with identifying both the color and the number from the target voice. Then, the participant was prompted to recall the numbers from the initial group of sentences. A total of 240 sentences were presented in six to eight blocks, depending on set size. Because the nature of this task required double the sentences as the concurrent task, the participant was prompted to stop and take at least a five-minute break outside of the test booth midway through the experiment. See Figure 3 for the timing of the secondary (recall) task relative to the primary (speech perception) task.

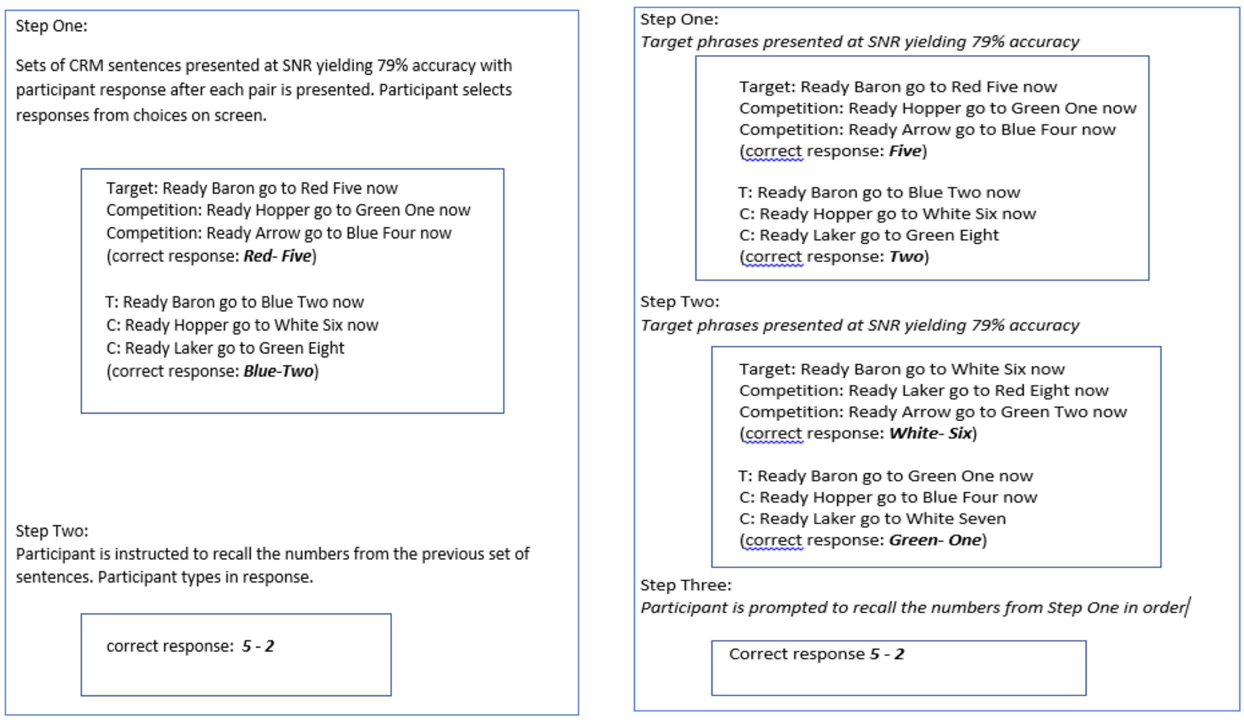

Figure 3. Timing of the secondary task relative to the primary task for the concurrent dual-task (shown left) and the sequential dual-task (shown right). This outline details a two-item task; each participant practiced with two and then four items; the experiment was conducted with set sizes corresponding to the individual memory span.

Rating of listening effort

After each block for the concurrent task and after every other block for the sequential task, the participant was prompted to rate their desire to improve the listening situation. A pen and sheet of paper describing the scale with blanks for each rating were provided. A new sheet was provided at the beginning of each experimental condition; the prompting to provide the rating appeared on the computer screen upon completion of a block of sentences.

Response Times

Response times for the recall dual-task were collected. The response time was defined as the amount of time elapsed between the initial response on the recall dual-task and participant indication of readiness for the next trial. An average response time was calculated by dividing the total response time by the individual set siz**e.**

Counterbalancing order of experimental conditions

Participants were randomly assigned based on order of enrollment into four different experimental groups for purposes of partially counterbalancing the order of presentation of the experimental conditions. Given the difficulty of the dual-task conditions and that trial-to-trial uncertainty in talkers can have larger effects in older adults (Goldinger, Pisoni, & Logan, 1991; Humes, Lee, & Coughlin, 2006; Humes & Coughlin, 2009) it was not desirable to have block-to-block or trial-to-trial variation in gender of the competing talkers which would have resulted from complete randomization or full counterbalancing of order. Instead, half of the groups received the female competition first and the male competition second, and half of the groups received the concurrent task first and the sequential task second. Table 1 describes the order for each experimental group. In most cases, participants completed the first half of the experimental tasks (i.e. the first two) at the second session, and the second half at the third and final session. Thus, the gender of the competing talkers was the same for the entire session in most cases. In some cases, the participant only completed one of the experimental tasks scheduled for the first session at the second session. In this case, the order of presentation was maintained and a fourth session was scheduled to complete the experiment. Participants completed no more than two experimental conditions per session, each lasting between 1.5 and 2 hours.

Group	1st	2nd	3rd	4th
1	CM	SM	CF	SF
2	SM	CM	SF	CF
3	SF	CF	SM	CM
4	CF	SF	CM	SM

Table 3. Order of presentation of the experimental conditions for each experimental group. CM = concurrent dual-task with male competition; SM = sequential dual-task with male competition; CF = concurrent dual-task with female competition; SF = sequential dual-task with female competition.

Second measure of baseline

At the final session, following completion of all experimental conditions, the single-task baseline measures of TCR and memory span were re-evaluated. For the speech-segregation task, the participant completed the same method of constant stimuli outlined above. Twenty trials at the initial adaptive baseline TCR, and 20 trials each at +6, +12, -6, and -12 dB TCR relative to the baseline TCR were presented. A psychometric function was estimated from these data and the TCR corresponding to 79.4% performance was recorded and compared to the baseline value. For the single-task recall baselines, three blocks of ten sets each were presented using the initial single-task baseline TCR for the memory-span task. Percent-correct scores for each block were averaged and compared to the initial baseline values.

Analysis

The variable of ***primary interest*** was the measure of change in performance during the DTP, referred to as the dual-task effect. A mixed-model analysis was applied using the independent groups of ONH and YNH as a between-subjects factor. Repeated-measures variables were concurrent vs. sequential dual-task completion and gender of the competing talkers.

Chapter 4. Results

Results from the statistical analyses are presented in the pages to follow. First, baseline measures and a comparison of differences across age groups are presented. This is followed by a comparison of the baseline TCRs and the baseline performance at individually determined memory spans when performed as a single task, obtained at the initial and final sessions. This analysis was performed to examine the stability of those baselines over the course of the experiment. Comparisons of speech and recall performance, reaction times, and ratings across sequential vs. concurrent condition, gender of the competing talkers (male or female), and age group are then presented. Finally, analyses of "dual-task effect", the amount of change in score from baseline to the experimental tasks, are presented.

Preliminary analyses and data reduction

Baseline measures were obtained to provide an estimate as to whether results from the experimental conditions could generalize to everyday conditions. The data distributions for the baseline measures in many cases were non-normal with unequal variance; non-parametric analyses were used for these variables. In cases of significant differences between groups, an effect size (*r*) was calculated by dividing the z score by the square root of the sample size (Cohen, 1992). The data analysis was conducted using IBM SPSS version 25.0.

Figures 4-6 show medians and interquartile ranges for the SSTM, CST, and CST self-report effort ratings. A Mann-Whitney test was used to evaluate potential differences between the YNH group and the ONH group for each of these measures. Performance on the SSTM was significantly better for the YNH group (Mdn = 85.8%) than for the ONH group (Mdn = 73.8%), *U* = 17.0, *p* < .001. The effect size for this analysis was found to be large following Cohen's convention (z = 4.49; *r* = 0.76) for effect sizes (Cohen, 1992).

Performance on the CST was significantly better for the YNH group (Mdn = 41%) than for the ONH group (Mdn = 35%), *U* = 72.5, *p* = .008. Recall that all participants were tested at a speech level of 63 dB SPL and an SNR of +2 dB. The effect size for this analysis was found to be medium following Cohen's convention (z = 2.67; *r* = 0.45) for effect sizes (Cohen, 1992).

Self-report effort ratings given after each half of the CST test did not significantly differ between the YNH group (Mdn first half = 6; Mdn second half = 8) and the ONH group (Mdn first half = 7; Mdn second half = 9), *U* (first half) = 113.5, p > 0.10; *U* (second half) = 108, p > 0.10. Recall that higher ratings reflect greater effort with the scale running from 0-10.

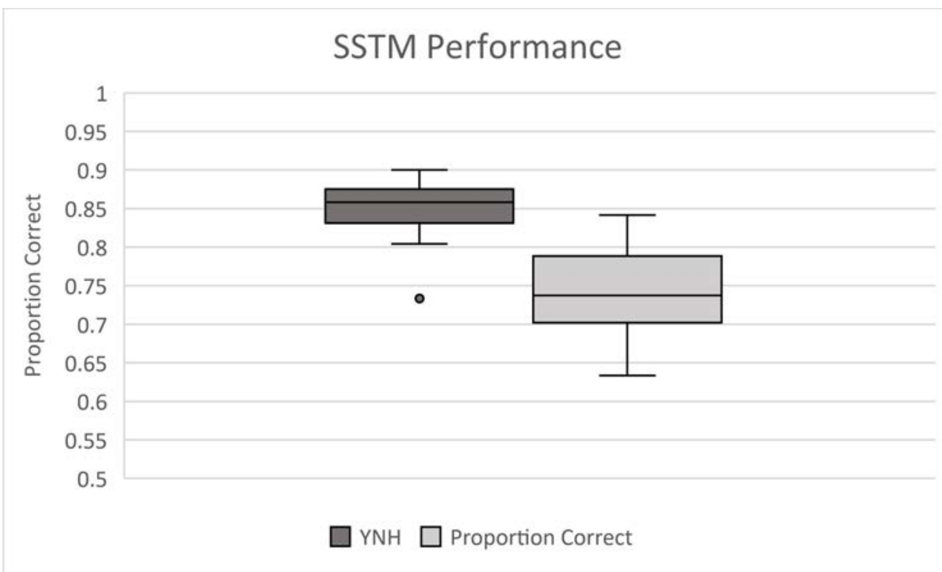

Figure 4. Median and interquartile range for performance on the Spatial Short-Term Memory (SSTM) test for the younger (YNH) and older (ONH) groups.

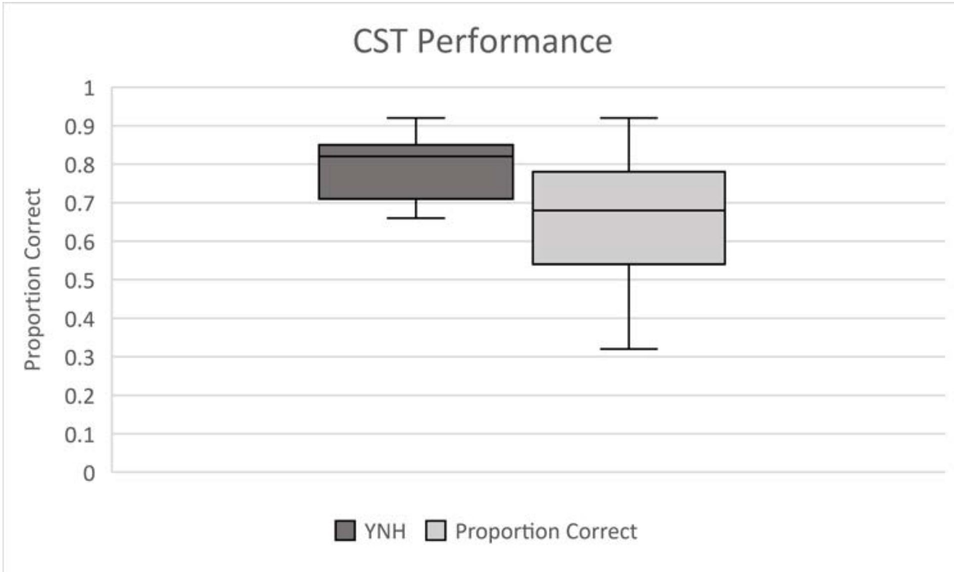

Figure 5. Median and interquartile range for the performance on the Connected Speech Test (CST) for the younger (YNH) and older (ONH) groups.

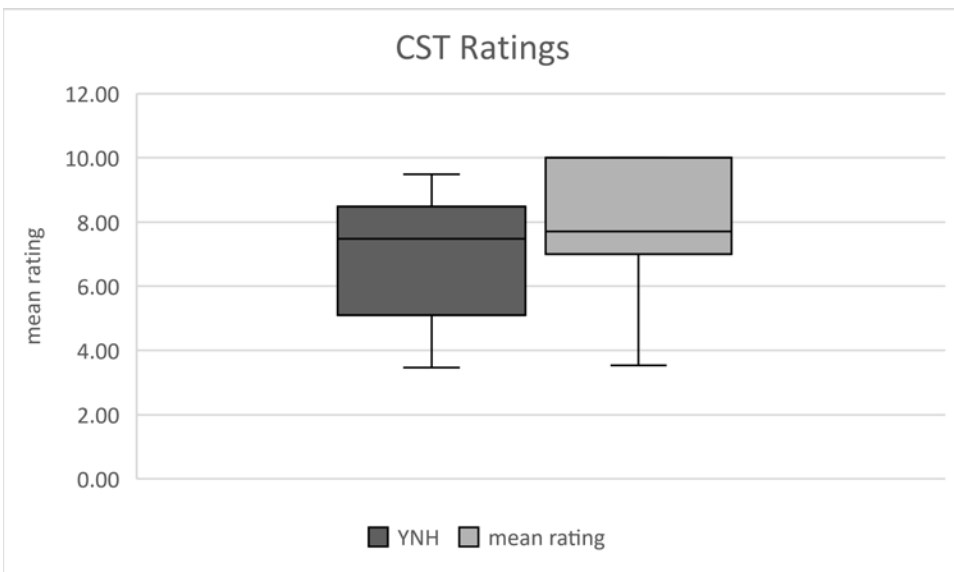

Figure 6. Median and interquartile range for self-reported effort ratings (on a scale of 1-10) of the Connected Speech Test (CST) for the younger (YNH) and older (ONH) groups.

Baseline TCR and memory span values were obtained for each of these tasks when administered as a single task. Median and interquartile ranges for baseline TCR and memory span are shown in Figures 7 and 8. Baseline TCR values for the primary speech-identification task did not differ significantly between the YNH (Mdn = 1.30 dB) and ONH groups (Mdn = 3.75 dB), *U* = 98.0, *p* =0.069. Baseline memory span, or the longest set size a participant was able to recall for the experimental stimuli at least 80% of the time, was significantly larger for the YNH group (Mdn = 5 items) than for the ONH group (Mdn = 4 items), *U* = 91.5, *p* = .034. The effect size for this analysis was found to be medium following Cohen's convention (z = 2.12; *r* = 0.36; Cohen, 1992). Scores ranged between 3 and 6 for the YNH group and between 3 and 5 for the ONH group.

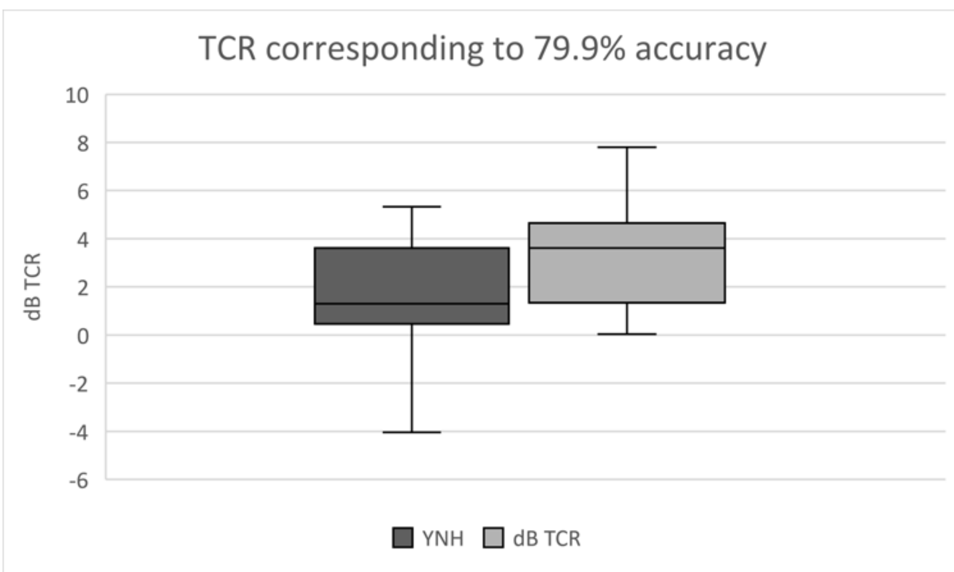

Figure 7. Median and interquartile ranges for the TCR at which 79.9% accuracy was achieved on the speech segregation task, when administered as a single task. YNH = young normal hearing; ONH = older normal hearing.

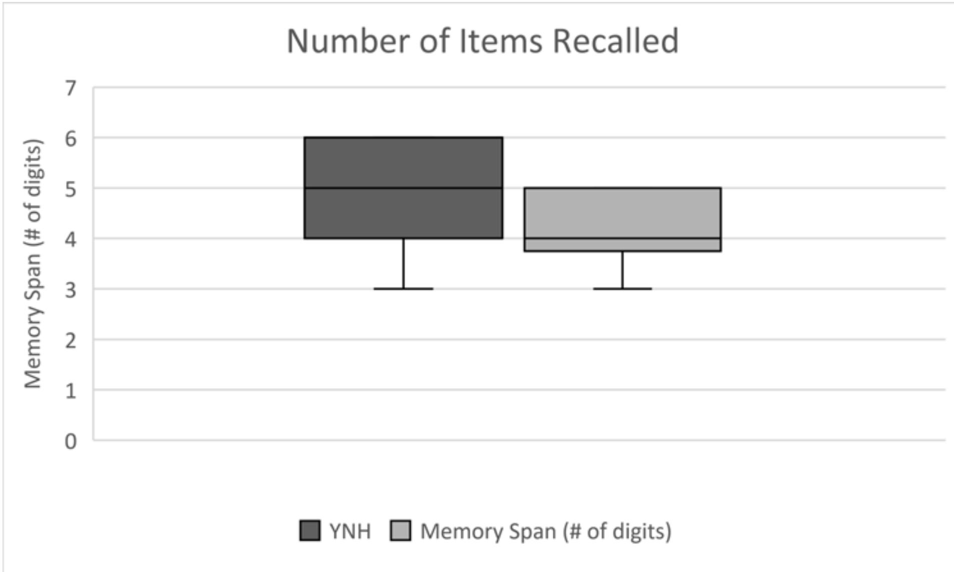

Figure 8. Median and interquartile ranges for the number of items recalled on the recall task, when administered as a single task. YNH = young normal hearing; ONH = older normal hearing. The largest number of items recalled was 6 and 5 for the YNH and ONH groups, respectively.

To summarize, the YNH and ONH groups differed in terms of performance on both the measure of working memory (SSTM) and the measure of speech in noise performance (CST), with the younger group outperforming the older group. Regarding the baseline, single-task measures, the YNH group demonstrated a significantly larger memory span for the test materials. The YNH group was also able to perform the speech segregation task at a less favorable TCR, though this difference did not reach statistical significance. Self-report ratings of effort on the CST did not differ between groups.

Recall that single-task measures of both the primary and secondary measures were obtained twice: once at the beginning and again after completion of all laboratory sessions. This was an attempt to evaluate the stability of these single-task baselines, against which the performance in all dual-task conditions is compared when determining the “dual-task effect”, the primary measure of listening effort in this project. Figure 9 shows median and interquartile ranges for the TCR required to attain 79.9% correct on the speech segregation task when administered as a single task. Initial and final baseline measures of TCR were weakly correlated, ($r = 0.28$), with participants in both groups achieving 79.9% accuracy at lower TCRs at the final session.

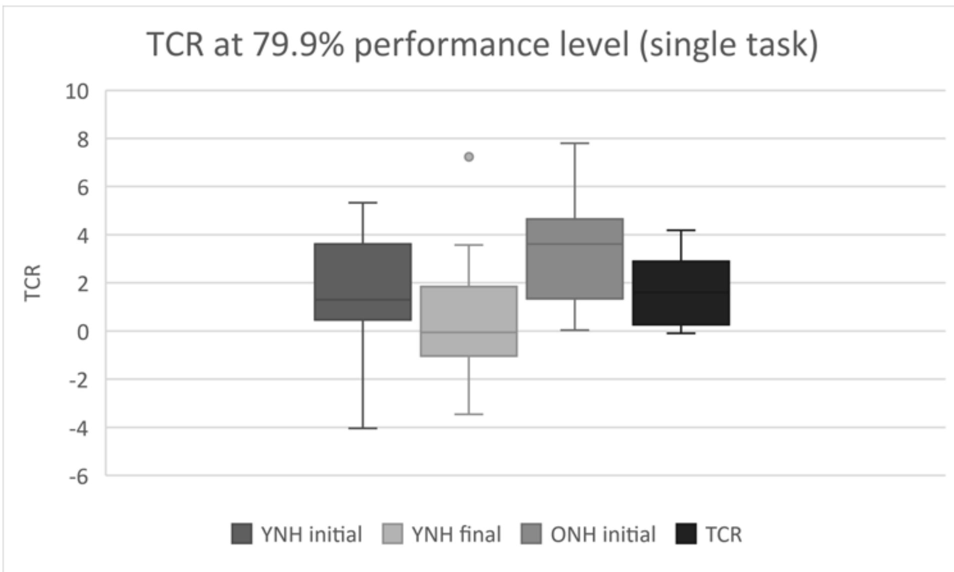

Figure 9. Median and interquartile ranges for the TCR at which 79.9% accuracy was achieved on the speech segregation task, when administered as a single task, at the initial and final sessions. YNH = young normal hearing; ONH = older normal hearing.

Once the participant's individual memory span for the experimental materials was estimated, all future testing, including final measures, was conducted at that span size. The span size was defined as the set size for which the participant could correctly recall at least 80% of trials within a block of ten trials. Proportion-correct scores for the recall task were transformed to rationalized arcsine units (RAU) (Studebaker, 1985) for analysis. Figure 10 shows median and interquartile ranges for performance on the recall task at the initial and final sessions. Initial and final measures of recall performance at individual memory span were found to be moderately correlated for all subjects combined (r = .37, p = .029).

A further analysis probed differences among the young and older groups for initial and final measures for both TCR and memory-span measures. Scores overall were weakly correlated from initial to final session. The older adults showed improved performance at the final session relative to the initial session. Pearson product-moment correlations indicated that initial and final TCR values were weakly correlated for the YNH group ($r(15) = .21$, $p > .05$) and were moderately correlated for the ONH group ($r(16) = .53$, $p < .05$). A Wilcoxon Signed Ranks Test indicated that for the YNH group, a significant difference was not seen between initial TCR values (Mdn = 1.30 dB) and final TCR values (Mdn = -.05 dB), z = -1.82, p = .068. For the ONH group, a significant difference was seen between initial TCR values (Mdn = 3.62 dB) and final TCR values (Mdn = 1.6 dB), z = -2.50, p = .012. In other words, at the end of the experiment, older adults were able to achieve 79.9% accuracy on the speech task at a less favorable TCR. The effect size for this analysis was found to be large following Cohen's convention (r = .59) for a large effect size (Cohen, 1992).

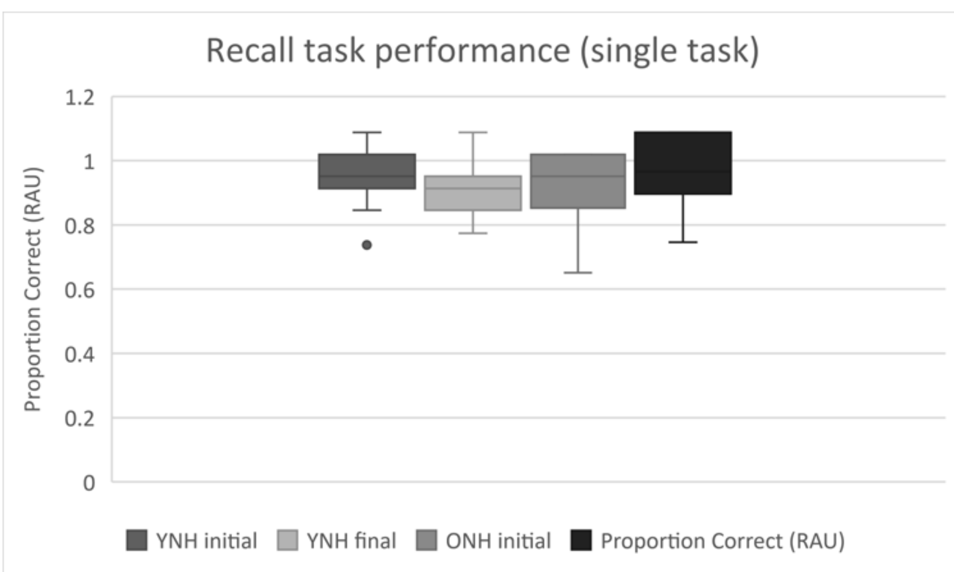

Figure 10. Median and interquartile ranges for performance on the recall task at individual span, when administered as a single task, at the initial and final sessions. YNH = young normal hearing; ONH = older normal hearing.

Overall, across both groups, scores on the recall task were weakly correlated from the initial to final sessions. Pearson product-moment correlation indicated small correlation between recall scores at the initial and final sessions ($r(33) = .29$, $p > .05$). Initial and final recall performance at the individually determined memory span was not significantly correlated for the YNH group ($r(15) = .34$, $p > .05$) but was moderately correlated for the ONH group ($r(16) = .53$, $p = .023$). For the YNH group, a significant difference was not seen between initial performance at the individual memory span (Mdn = 95 RAU) and final performance at the individual memory span (Mdn = 92 RAU; Wilcoxon z = -0.82, $p = .411$). For the ONH group, a significant difference was seen between performance for the initial memory span (Mdn = 91 RAU) and performance at the for the final individual memory span (Mdn = 94 RAU; Wilcoxon z = -2.14, $p = .03$. The effect size for this analysis was found to be large following Cohen's convention ($r = .51$; Cohen, 1992).

To recap, as a group, the TCR at which 79.9 % accuracy was achieved on the speech task was weakly correlated at the initial and final sessions. Further analysis showed a moderate correlation between performance at initial and final sessions for the older group but not the younger group. Additionally, older adults were seen to perform the task at a more difficult TCR at the final session than they did initially. Performance on the recall task followed a similar pattern. As a group, performance on the recall task was moderately correlated at the initial and final sessions. Further analysis showed a moderate correlation between performance and the initial and final sessions on the recall task for the older group but not the younger group.

Performance on the recall task was significantly improved from initial to final sessions for the older adults.

Dual-Task Measures

Results from the dual-task measures are presented below. First, results from the primary task (speech segregation task) is presented, followed by results from the secondary task (recall task). A Generalized Estimating Equation (GEE), which is a mixed model and estimates the parameters of a Generalized Linear Model (GLM), was used for all subsequent analyses. A primary advantage to the GEE for the purpose of this experiment is that it can be used to analyze repeated measures. Further, it can accommodate the "missing" data points from participants who completed fewer blocks due to the individual set size chosen for the experiment (Hardin & Hilbe, 2003).

Proportion-correct scores for both the speech task and the recall task were first calculated and then transformed to rationalized arcsine units (RAU) (Studebaker, 1985) for data analysis. All speech-identification and recall scores reported here are expressed in RAU values. To further normalize the distribution, the GEE analysis employed a log transform. Use of the log transform required elimination of negative RAU values. To do so, a value of 40 was added to all scores. This value was subsequently subtracted to arrive at the means reported here. All *p*-values for multiple paired comparisons reported are Bonferroni-corrected values (p_{adj} where $p_{adj} = p/N$ for N paired comparisons for that dependent variable).

Speech and recall scores were analyzed using a GEE with order as a covariate. Recall that subgroups received different condition orders using a counterbalanced design. No significant effect of order or interaction with order was seen ($p = .60$). Speech and recall scores were analyzed with block number as a covariate given that later blocks could have been impacted by learning or fatigue. No significant effect of block number or interaction with block number was seen (p = .66). Speech and recall scores were analyzed using a GEE with recall set size (based on individual memory span) as a covariate. No significant effect of set size or interaction with set size was seen (p > .05). As a result, results were pooled across these variables and they were not included as variables in further analyses**.**

Primary task (speech segregation task)

Mean speech-identification scores during the dual-task paradigm for concurrent and sequential conditions and male and female talker competition are shown in Figure 11. A main effect indicating better performance with female talker competition (p < .001) was observed. An interaction indicating greatest performance for the concurrent condition with female talker competition (CF dual-task condition; p = .004) was also observed. Significant interactions between age group and gender of the competing talkers (p = .724), age group and condition (p = .654) or age group, gender of the competing talkers, and condition (p = .13) were not seen. See table 2 for a summary of estimated marginal means, standard errors, and confidence intervals.

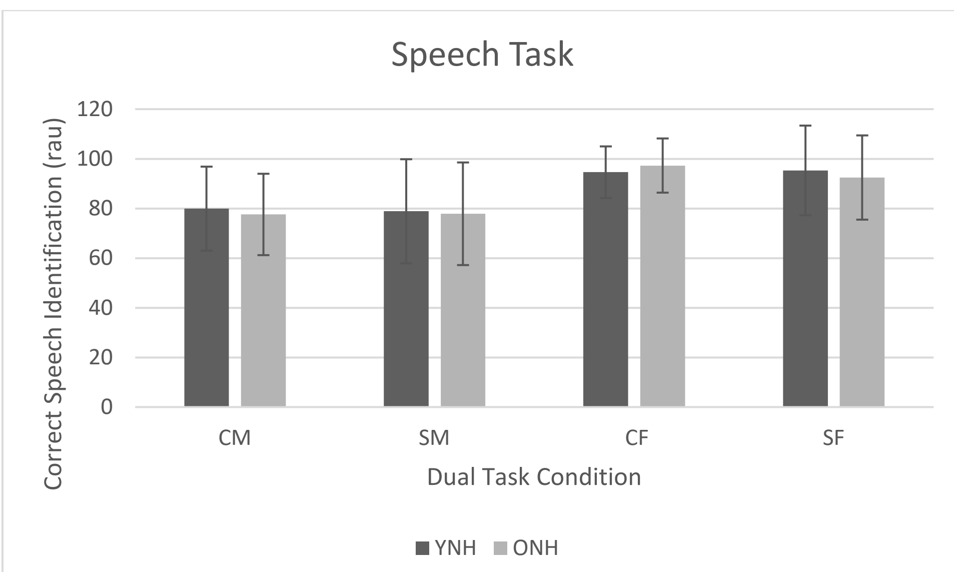

Figure 11. Mean and standard deviation values for the speech-identification task in the dual-task paradigm. YNH = young normal hearing; ONH = older normal hearing; CM =

concurrent condition, male talker competition; SM = sequential condition, male talker competition; CF = concurrent condition, female talker competition; SF = sequential condition, female talker competition.

Variable	EMM (rau)	Std. E	95% CI	95% CI
Male Comp.	82.22	2.62871	77.22	87.53
Female Comp.	94.81	1.77	91.39	98.35
Con	89.28	2.23	85	93.76
Seq	87.34	2.63	82.33	92.63
Young	89.08	2.89	83.58	94.92
Older	87.53	3.07	81.72	93.74
Con * Male	81.34	2.73	71.15	81.85
Con * Female	97.95	2.03	89.04	97.02
Seq * Male	83.11	3.11	72.23	84.42
Seq * Female	91.76	2.58	81.84	91.95
Con * Young	89.5	3.34	83.17	96.28
Seq * Young	88.66	3.73	81.62	96.26
Con * Older	89.06	2.97	83.42	95.05
Seq * Older	86	3.7	79.07	93.58
Young * Male	83.26	3.68	76.33	90.78
Young * Female	95.28	2.33	90.81	99.96
Older * Male	81.2	3.75	74.15	88.87
Older * Female	94.34	2.67	89.24	99.71
Young * Con * Male	82.77	4.22	74.88	91.43
Young * Con * Female	96.75	2.77	91.47	102.32
Young * Seq * Male	83.75	4.07	76.12	92.1
Young * Seq * Female	93.83	4.06	86.19	102.12
Older * Con * Male	79.93	3.49	73.36	87.05
Older * Con * Female	99.17	2.99	93.48	105.19
Older * Seq * Male	82.48	4.68	73.77	92.15
Older * Seq * Female	89.73	3.22	83.63	96.26

Table 2. Estimated marginal means (EMM), standard errors (Std. E), and 95% confidence intervals (CI) for the full factorial analysis of gender of the competing talkers, condition, and age for the primary (speech segregation) task. Comp = competition; con = concurrent condition; seq = sequential condition. Significant main effects and interactions are bolded. * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$

An analysis was performed a second time with SSTM, CST, and baseline single-task TCR as covariates. The covariate of baseline single-task TCR was found to be a significantly predictive of performance on the speech-segregation dual task (b = -0.367, $p < .001$; 95% confidence intervals [-0.521, -0.211]. A smaller baseline single-task TCR was predictive of better speech-identification dual-task scores. In other words, participants who were able to achieve 79.9% accuracy on the speech-segregation task as a single task with a less favorable TCR showed poorer performance on the speech-segregation task as a dual task. The covariates of SSTM score and CST score were not found to be significantly predictive of speech-identification scores.

Secondary task (recall task)

Mean recall scores during the dual-task paradigm for concurrent and sequential conditions and male and female talker competition are presented in Figure 12. A main effect indicating better performance in the concurrent condition was observed ($p = .008$). Significant interactions between age group and gender of the competing talkers ($p = .21$), age group and condition ($p = .21$) or age group, gender of the competing talkers, and condition ($p = .93$) were not seen. See Table 4 for a summary of estimated marginal means, standard errors, and confidence intervals.

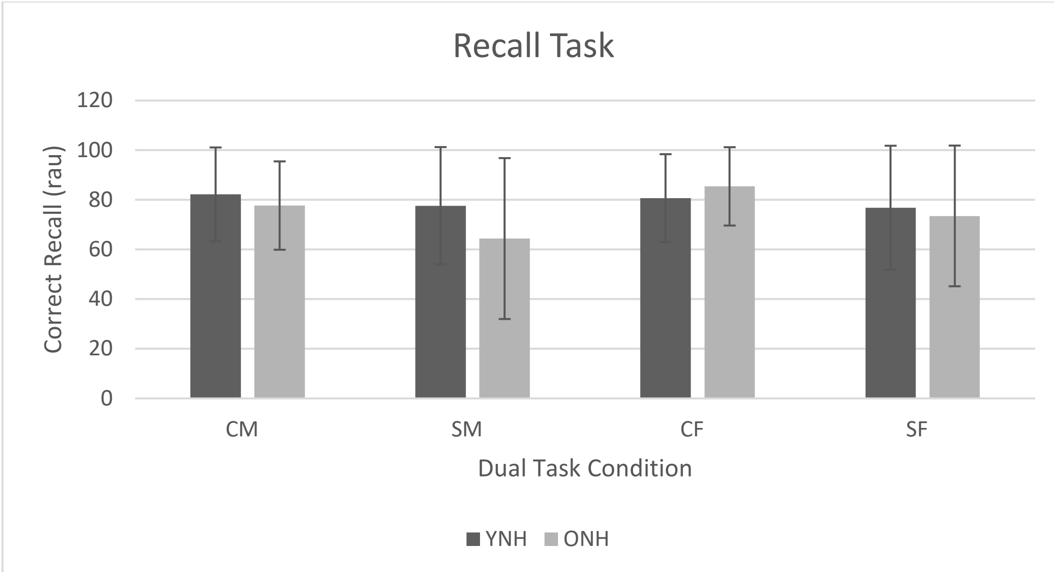

Figure 12. Mean and standard deviation values for the recall task in the dual-task paradigm. YNH = young normal hearing; ONH = older normal hearing; CM = concurrent condition, male talker competition; SM = sequential condition, male talker competition; CF = concurrent condition, female talker competition; SF = sequential condition, female talker competition.

Variable	EMM (rau)	Std. E	95% CI	95% CI
Male Comp.	75.87	2.4	71.3	80.72
Female Comp.	79.27	2.68	74.18	84.68
Con	81.22	1.89	77.6	85.01
Seq	74.04	3.04	68.3	80.22
Young	79.73	2.58	74.82	84.95
Older	75.43	3.52	68.81	82.63
Con * Male	79.94	2.19	75.75	84.35
Con * Female	82.52	2.36	78.02	87.27
Seq * Male	71.99	3.56	65.32	79.3
Seq * Female	76.13	3.45	69.64	83.19
Con * Young	89.5	3.34	83.17	96.28
Seq * Young	88.66	3.73	81.62	96.26
Con * Older	89.06	2.97	83.42	95.05
Seq * Older	86.04	3.7	79.07	93.58
Young * Male	83.26	3.68	76.33	90.78
Young * Female	95.28	2.33	90.81	99.96
Older * Male	81.2	3.75	74.15	88.87
Older * Female	94.34	2.67	89.24	99.71

Young * Con * Male	82.77	4.22	74.88	91.43
Young * Con * Female	96.75	2.77	91.47	102.32
Young * Seq * Male	83.75	4.07	76.12	92.1
Young * Seq * Female	93.83	4.06	86.19	102.12
Older * Con * Male	79.93	3.49	73.36	87.05
Older * Con * Female	99.17	2.99	93.48	105.19
Older * Seq * Male	82.48	4.68	73.77	92.15
Older * Seq * Female	89.73	3.22	83.63	96.26

Table 4. Estimated marginal means (EMM), standard errors (Std. E), and 95% confidence intervals (CI) for the full factorial analysis of gender of the competing talkers, condition, and age for the secondary (recall) condition. Comp = competition; con = concurrent condition; seq = sequential condition. Significant main effects and interactions are bolded. * = $p < .05$; ** = $p < .01$; *** = $p < .001$

An analysis was performed a second time with SSTM, CST, and baseline single-task TCR as covariates. None of these covariates were found to be significantly predictive of recall scores. Because the baseline percent-correct recall scores ranged from 80-100%, the full factorial GEE analysis was performed with baseline percent correct recall scores as a covariate. The baseline percent correct recall score was found to be significantly predictive of higher percent recall scores in the experimental conditions (b = .010, $p = .001$; 95% confidence intervals [0.004, 0.017], with higher percent-correct recall scores predicting better performance on the recall task in the experimental conditions. Further analysis of this finding showed that baseline percent-correct recall score was more predictive of performance on the sequential condition (b = 1.16) than the concurrent condition (b = 0.25).

Response time

Mean response times on the recall task during the dual-task paradigm are shown in Figure 13. Recall that average response times per item were calculated by dividing the total response time, from the first to the final recalled item within a set by the set size. A main effect for faster response time was seen for the concurrent condition (*p* = .008). A significant interaction between age group and condition was seen (*p* = .027), with substantially shorter response times for the YNH group for both (male and female talker competition) concurrent conditions. See Table 5 for a summary of estimated marginal means, standard errors, and confidence intervals.

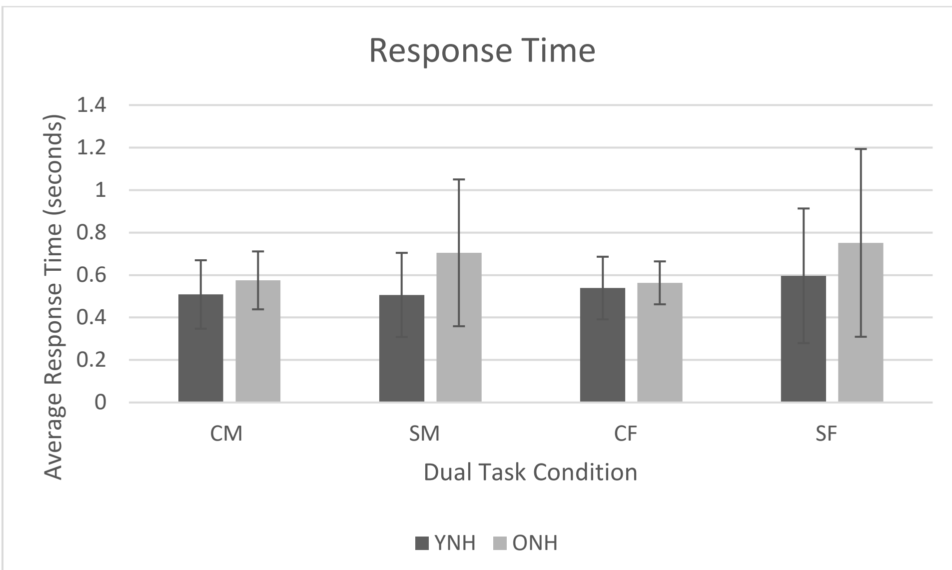

Figure 13. Mean and standard deviation response times during the recall task in the dual-task paradigm. YNH = young normal hearing; ONH = older normal hearing; CM = concurrent condition, male talker competition; SM = sequential condition, male talker competition; CF = concurrent condition, female talker competition; SF = sequential condition, female talker competition.

Variable	EMM (seconds)	Std. E	95% CI	95% CI	
Male Comp.	2.59	0.13	2.34	2.87	
Female Comp.	2.78	0.15	2.5	3.08	
Con	2.5	0.11	2.29	2.73	**
Seq	2.88	0.18	2.54	3.26	

Young	2.63	0.17	2.32	2.97	
Older	2.74	0.2	2.37	3.17	
Con * Male	2.47	0.13	2.23	2.74	
Con * Female	2.54	0.12	2.32	2.77	
Seq * Male	2.72	0.17	2.41	3.07	
Seq * Female	3.04	0.24	2.6	3.56	
Con * Young	2.6	0.18	2.26	2.98	*
Seq * Young	2.66	0.17	2.34	3.02	
Con * Older	2.41	0.14	2.16	2.7	
Seq * Older	3.12	0.34	2.52	3.85	
Young * Male	2.47	0.17	2.16	2.83	
Young * Female	2.79	0.19	2.44	3.19	
Older * Male	2.72	0.21	2.33	3.17	
Older * Female	2.76	0.22	2.37	3.23	
Young * Con * Male	2.46	0.19	2.12	2.86	
Young * Con * Female	2.74	0.2	2.36	3.17	
Young * Seq * Male	2.48	0.17	2.18	2.83	
Young * Seq * Female	2.85	0.27	2.36	3.44	
Older * Con * Male	2.48	0.18	2.16	2.85	
Older * Con * Female	2.35	0.13	2.12	2.61	
Older * Seq * Male	2.98	0.3	2.44	3.64	
Older * Seq * Female	3.25	0.41	2.53	4.18	

Table 5. Estimated marginal means (EMM), standard errors (Std. E), and 95% confidence intervals (CI) for the full factorial analysis of gender of the competing talkers, condition, and age for the response time measures. Comp = competition; con = concurrent condition; seq = sequential condition. Significant main effects and interactions are bolded. * = $p < .05$; ** = $p < .01$; *** = $p < .001$

Self-report ratings of effort during the dual-task paradigms

Figure 14 shows median and interquartile range participant ratings following experimental blocks. Recall that participants were asked to rate effort as follows: “on a scale of 1 to 10, please rate your desire to improve the listening situation”, with 1 indicating “not at all” and 10 indicating = “a lot”.” A main effect for condition was observed, with higher ratings given for the concurrent condition (*p* = .007). A main effect of gender of the competing talkers (*p* < .001) was also observed, with higher ratings given for male talker competition. There were no significant interactions. The covariate of baseline TCR was found to be a significantly predictive of higher ratings (b = .738, *p* = .009; 95% confidence intervals [0.183, 1.294], with a less favorable baseline TCR predicting greater self-report effort. The covariates of SSTM score and CST score were not found to be significantly predictive of higher ratings. See Table 6 for a summary of estimated marginal means, standard errors, and confidence intervals.

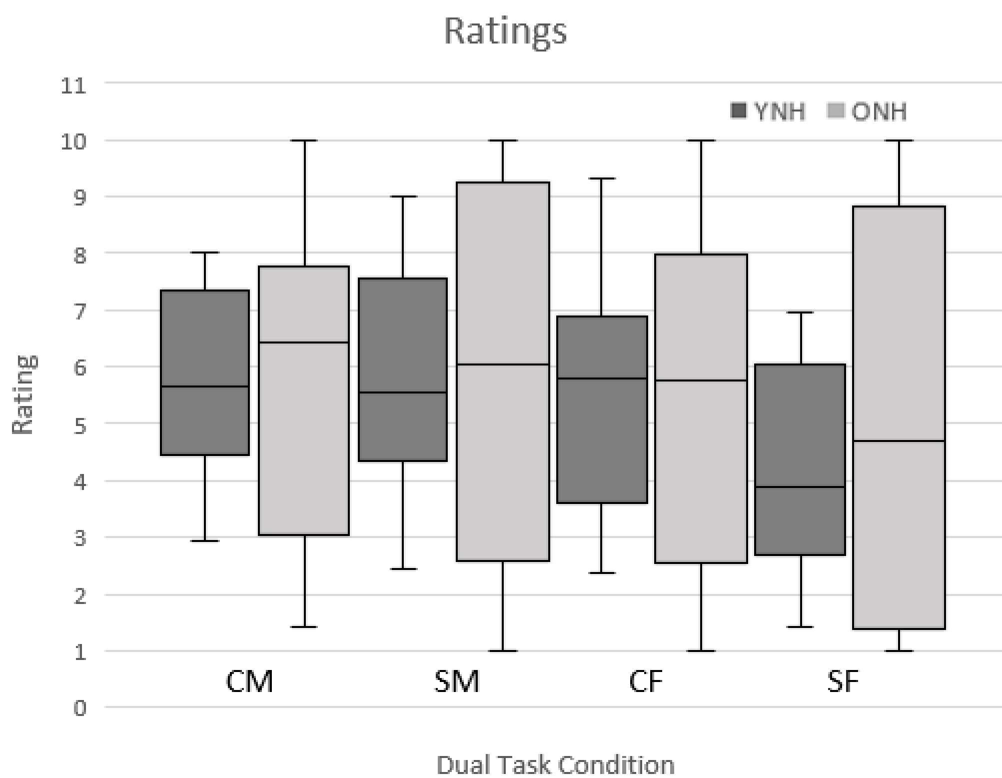

Figure 14. Median and interquartile range ratings following the experimental blocks. Participants were asked "on a scale of 1 to 10, please rate your desire to improve the listening situation". 1= "not at all"; 10 = "a lot"; YNH = young normal hearing; ONH = older normal hearing; CM = concurrent condition, male talker competition; SM = sequential condition, male talker competition; CF = concurrent condition, female talker competition; SF = sequential condition, female talker competition.

Variable	EMM (rating)	Std. E	95% CI	95% CI	
Male Comp.	6.06	0.41	5.3	6.9	***
Female Comp.	5	0.41	4.26	5.9	
Con	6.03	0.44	5.22	7	**
Seq	5.03	0.42	4.26	5.9	
Young	5.27	0.43	4.5	6.2	
Older	5.75	0.68	4.56	7.2	
Con * Male	6.4	0.44	5.59	7.3	
Con * Female	5.68	0.49	4.79	6.7	
Seq * Male	5.74	0.47	4.9	6.7	
Seq * Female	4.4	0.44	3.61	5.4	
Con * Young	5.94	0.55	4.96	7.1	
Seq * Young	4.68	0.48	3.83	5.7	
Con * Older	6.11	0.7	4.88	7.7	
Seq * Older	5.4	0.73	4.15	7	
Young * Male	5.79	0.45	4.98	6.7	
Young * Female	4.8	0.45	3.99	5.8	
Older * Male	6.34	0.71	5.09	7.9	
Older * Female	5.2	0.69	4.01	6.8	
Young * Con * Male	6.22	0.5	5.31	7.3	
Young * Con * Female	5.67	0.68	4.49	7.2	
Young * Seq * Male	5.39	0.55	4.41	6.6	
Young * Seq * Female	4.06	0.48	3.22	5.1	
Older * Con * Male	6.58	0.73	5.3	8.2	
Older * Con * Female	5.68	0.71	4.45	7.3	
Older * Seq * Male	6.11	0.77	4.78	7.8	
Older * Seq * Female	4.77	0.78	3.47	6.6	

Table 6. Estimated marginal means (EMM), standard errors (Std. E), and 95% confidence intervals (CI) for the full factorial analysis of gender of the competing talkers, condition, and age for ratings obtained after experimental blocks. Comp = competition; con = concurrent condition; seq = sequential condition. Significant main effects and interactions are bolded. * = $p < .05$; ** = $p < .01$; *** = $p < .001$

Dual-Task Effect

The dependent variable of primary interest in this study was the "dual-task effect". Recall that the measure of cognitive cost in a dual-task paradigm can be viewed as a change in performance on a task when performed alongside another task, compared to performance on that task when performed as a single task. The primary task in this experiment was the speech-segregation task. Therefore, it was expected that speech scores would remain relatively unchanged throughout the experiment. On the other hand, performance on the secondary task (the recall task) was expected to decrease in the dual-task paradigm.

The dual-task effect for the speech task was calculated by subtracting the dual-task rau-transformed percent correct scores from the single-task rau-transformed baseline score. Recall that the single-task baseline score was the same (~80% correct) for all participants. Dual-task effects reflecting greater effort are positive values whereas gains in performance, reflecting less effort in the dual-task relative to the single task, are negative values. Means and standard deviations for concurrent and sequential conditions, with male and female talker competition are shown in Figure 15. A main effect of gender of the competing talkers was the only significant effect observed (p <.001) with GEE analysis with a relatively greater effect seen with male talker competition than for the female talker competition. The mean dual-task effect for the male talker competition was close to zero for both tasks; when the competing talkers were female, participants did better on the speech segregation task. An interaction between condition and gender of the competing talkers was seen after removing the factor of age group from the analysis (p = .029), with greater effect seen for male talker competition in the sequential condition. See Table 7 for a summary of estimated marginal means, standard errors, and confidence intervals for the full factorial analysis.

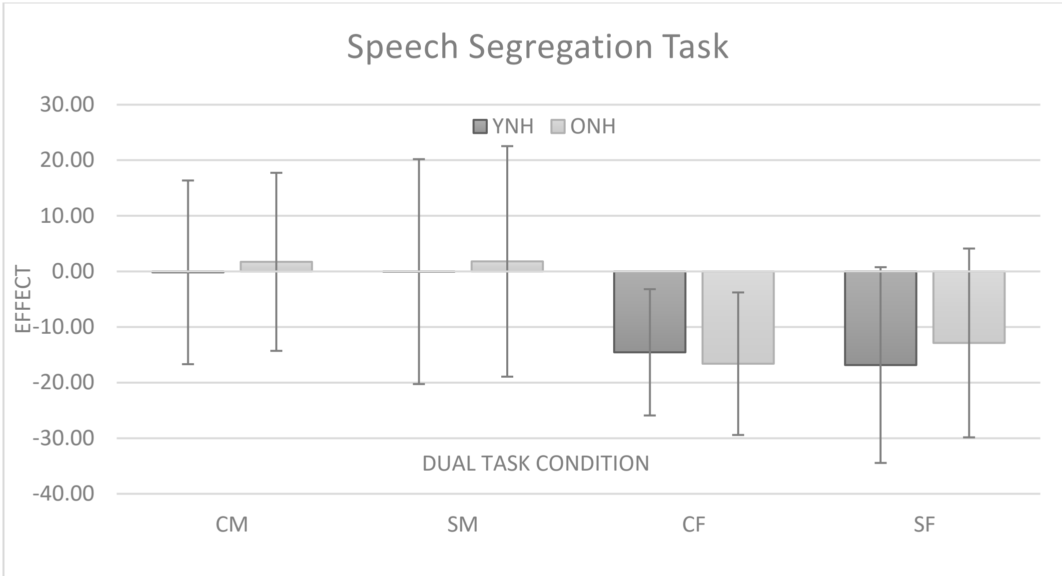

Figure 15. Means and standard deviations for the dual-task effect, in RAU, calculated based on performance on the speech-segregation task as the primary task. YNH = young normal hearing; ONH = older normal hearing; CM = concurrent dual task, male competition; SM = sequential dual task, male competition; CF = concurrent dual task, female competition; SF = sequential dual task, female competition.

Variable	EMM (effect)	Std. E	95% CI	95% CI	
Male Comp.	-0.91	-37	-5.56	4.38	***
Female Comp.	-15.46	-38	-18.72	-11.7	
Con	-8.4	-38	-11.98	-4.37	
Seq	-9.64	-37	-14.54	-3.79	
Young	-9.8	-37	-14.83	-3.78	
Older	-8.23	-37	-13.51	-1.89	
Con * Male	0.04	-37	-4.68	5.4	
Con * Female	-15.07	-38	-18.18	-11.51	
Seq * Male	-1.84	-37	-7.25	4.47	
Seq * Female	-15.84	-37	-20.48	-10.11	
Con * Young	-8.18	-37	-13.19	-2.23	
Seq * Young	-11.35	-36	-18.13	-2.46	
Con * Older	-8.63	-37	-13.5	-2.87	
Seq * Older	-7.83	-36	-14.35	0.35	
Young * Male	-1.96	-36	-8.33	5.69	
Young * Female	-16.03	-38	-20.38	-10.71	
Older * Male	0.17	-36	-6.29	7.88	

Older * Female	-14.88	-37	-19.49	-9.22
Young * Con * Male	-0.89	-36	-7.69	7.34
Young * Con * Female	-14.1	-38	-18.22	-9.2
Young * Seq * Male	-3	-36	-10.22	5.98
Young * Seq * Female	-17.81	-36	-24.26	-8.71
Older * Con * Male	1	-37	-5.16	8.26
Older * Con * Female	-16	-38	-20.41	-10.6
Older * Seq * Male	-0.64	-36	-8.26	8.81
Older * Seq * Female	-13.7	-37	-19.56	-6.17

Table 7. Estimated marginal means (EMM), standard errors (Std. E), and 95% confidence intervals (CI) for the full factorial analysis of gender of the competing talkers, condition, and age for ratings obtained after experimental blocks. Negative values indicate average performance was better on the dual-task. Comp = competition; con = concurrent condition; seq = sequential condition. Significant main effects and interactions are bolded. * = $p < .05$; ** = $p < .01$; *** = $p < .001$

The dual-task effect for the recall task was calculated by subtracting the rau-transformed percent-correct scores for the dual-task conditions from the average of rau-transformed percent-correct scores for the single-task baseline obtained at the initial and final sessions. Figure 16 shows means and standard deviations for concurrent and sequential conditions and male and female talker competition. A main effect of condition ($p = .002$) was observed, which was the only statistically significant effect, with a greater effect seen for the sequential condition. See Table 8 for a summary of estimated marginal means, standard errors, and confidence intervals.

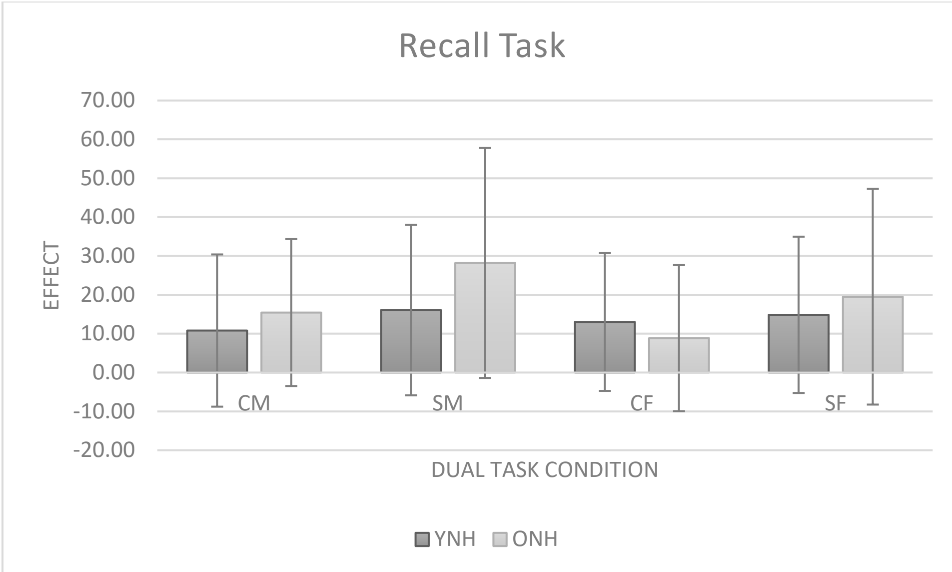

Figure 16. Means and standard deviations for the dual-task effect calculated based on performance on the recall task. YNH = young normal hearing; ONH = older normal hearing; CM = concurrent dual task, male competition; SM = sequential dual task, male competition; CF = concurrent dual task, female competition; SF = sequential dual task, female competition.

Variable	EMM (effect)	Std. E	95% CI	95% CI	
Male Comp.	17.08	2.1	13.22	21.28	
Female Comp.	13.9	2.1	9.95	18.23	
Con	12.04	1.8	8.61	15.77	
Seq	19.16	2.4	14.73	24.02	**
Young	13.59	1.9	10.1	17.38	
Older	17.41	3	11.94	23.6	
Con * Male	13.28	2.4	8.82	18.25	
Con * Female	10.83	1.9	7.23	14.79	
Seq * Male	21.21	2.8	16	27.01	
Seq * Female	17.19	2.9	11.82	23.26	
Con * Young	11.76	2.6	6.96	17.17	
Seq * Young	15.5	1.9	11.86	19.47	
Con * Older	12.32	2.6	7.6	17.64	
Seq * Older	23.11	4.6	14.84	32.9	
Young * Male	13.59	2.6	8.83	18.93	
Young * Female	13.59	2.3	9.39	18.24	
Older * Male	20.84	3.3	14.82	27.68	
Older * Female	14.21	3.6	7.72	21.82	

Young * Con * Male	11.13	3.4	5	18.32
Young * Con * Female	12.39	2.6	7.56	17.84
Young * Seq * Male	16.2	2.7	11.26	21.73
Young * Seq * Female	14.82	3	9.27	21.15
Older * Con * Male	15.54	3.4	9.36	22.7
Older * Con * Female	9.33	2.8	4.19	15.24
Older * Seq * Male	26.76	5.3	17.29	38.13
Older * Seq * Female	19.69	5.1	10.58	30.84

Table 8. Estimated marginal means (EMM), standard errors (Std. E), and 95% confidence intervals (CI) for the full factorial analysis of gender of the competing talkers, condition, and age for ratings obtained after experimental blocks. Negative values indicate average performance was better on the dual-task. Comp = competition; con = concurrent condition; seq = sequential condition. Significant main effects and interactions are bolded. * = $p < .05$; ** = $p < .01$; *** = $p < .001$

A dual-task effect for the recall task was also calculated by two additional means: by subtracting the rau-transformed percent-correct scores from the scores obtained at the initial session, and by subtracting the rau-transformed percent correct scores from the scores obtained at the final session. A paired-sample t-test indicated that there was not a significant difference between the effect obtained using scores from the initial session (M = 16.3, SD = 24.3) and the effect obtained using scores from the final session (M = 17.9, SD = 24.7); t (749) = -3.9, p < .001). Potential differences by group were also explored. Paired-sample t-tests indicated that for the YNH group, there was not a significant difference between the effect obtained using scores from the initial session (M = 15.4, SD = 21.3) and the effect obtained using scores from the final session (M = 12.8, SD= 20.6); t (351) = 4.408, $p < .001$). For the ONH group, there also was not a significant difference between the effect obtained using scores from the initial session (M

=17.1, SD = 26.7) and the effect obtained using scores from the final session (M = 22.3, SD = 27.2); t (397) = -11.3, $p < .001$).

These additional calculations were performed for the recall task, and not the speech task, because of the way the single task performance was calculated. For the recall task, the span size was determined at the initial session and was defined as the single task performance at which the participant obtained 80% correct; scores ranged from 80%-100%. At the final session, the recall task was re-administered at the set size that was determined at the initial session. In this way, performance on that task, in terms of percent correct, varied at both sessions. For the speech task, the single task performance was fixed at 79.9% at both sessions, with TCR to obtain that percent correct score varying for each participant at each session.

Correlations among measures

The associations among baseline measures of the SSTM, CST, and self-report ratings of effort on the CST and various performance measures from the DTP were examined. Pearson product-moment correlations among these measures were computed across all participants and for each group individually.

Correlation analysis indicated that SSTM scores and single-task baseline performance on the recall task were weakly correlated and non-significantly for both groups as a whole ($r(33) = .14$, $p > .05$), for the YNH group ($r(15) = .16$, $p > .05$), or for the ONH group ($r(16) = -0.20$, $p > .05$). Correlation analysis indicated that CST scores and single-task baseline performance on the speech-segregation task were weakly to moderately, but not significantly, correlated for both groups as a whole ($r(33) = 0.33$, $p = .05$), the YNH group ($r(15) = 0.30$, $p > .05$), and the ONH group ($r(16) = 0.25$, $p > .05$).

The associations between self-report ratings of effort obtained following the CST and the ratings obtained during the dual-task, and the DTE measured during the experiment, were also examined. Correlations were computed for each of the four dual-task conditions: concurrent condition with male competition (CM); sequential condition with male competition (SM); concurrent condition with female competition (CF); and sequential condition with female competition (SF). Across all subjects, CST effort ratings were moderately and significantly ($p < .05$) correlated with ratings for each of the conditions (CM, $r(33) = .51$; SM, $r(33) = .63$; CF, $r(33) = .53$; SF, $r(33) = .57$). For the YNH group, CST ratings were weakly and non-significantly correlated with the effort ratings for the CM condition ($r(15) = .24$, $p > .05$), strongly and significantly correlated for the SM condition ($r(15) = .70$, $p < .05$) and moderately correlated for the CF ($r(15) = .39$, $p >.05$) and SF ($r(15) = .55$, $p < .05$) conditions with only the latter of the two yielding a significant correlation. For the ONH group, CST effort ratings were strongly correlated with the DTE ratings for the CM ($r(16) = .70$, $p < .05$), SM ($r(16) = .63$, $p < .05$), and CF ($r(16) = .69$, $p < .05$) conditions, and moderately correlated for the SF ($r(16) = .55$, $p < .05$) condition. In general, the self-report effort ratings were correlated across those obtained initially for the CST and those obtained during the DTP.

The CST effort ratings were weakly correlated and non-significant ($p < .05$) with the DTEs measured for the speech dual-task for both age groups. For the YNH group, the correlation coefficients were as follows: CM, $r(15) = .21$; SM, $r(15) = .04$; CF, $r(15) = .11$; SF, $r(15) = -0.14$. For the ONH group, the correlation coefficients were as follows CM, $r(16) = -.14$; SM, $r(16) = .12$; CF, $r(16) = .00$; SF, $r(16) = .09$. A different pattern emerged for the recall dual-task. For the YNH group, CST effort ratings were moderately and negatively correlated with recall performance on the dual-task for the CM ($r(15) = -.43$, $p > .05$) and CF ($r(15) = -.54$, $p <$

.05) conditions, and were weakly correlated for the SM ($r(15) = -.18$, $p > .05$) and SF ($r(15) = -.09$, $p > .05$) conditions. For the ONH group, CST effort ratings were moderately correlated with recall performance on the dual-task for the SM ($r(16) = .34$, $p > .05$) condition and were weakly correlated for the CM ($r(16)= -.13$, $p > .05$), CM ($r(16) = -.04$, $p > .05$) and SF ($r(16) = .05$, $p > .05$) conditions. In general, CST effort ratings obtained initially were not correlated with the recall-based DTE effort measures.

Correlation analyses were performed on the dual task effects for the sequential and the concurrent condition for each group individually and for the group as a whole. For the speech-segregation task, dual task effects were strongly correlated and significant ($p < .05$) across the sequential and concurrent paradigms for both groups as a whole ($r(33) = .67$), for the YNH group ($r(15) = .60$), and for the ONH group ($r(16) = .75$). For the recall task, dual task effects were moderately correlated for both groups as a whole ($r(33) = .45$, $p > .05$), for the YNH group ($r(33) = .40$, $p > .05$), and for the ONH group ($r(33) = .44$, $p > .05$).

To summarize, the pattern that emerges is that the two self-report measures of effort, on the CST and during the dual-task paradigm, are generally moderately and significantly correlated, but neither rating is correlated with the DTE-based measures of effort. On the other hand, the two DTE-based approaches, sequential and concurrent, are moderately to strongly correlated, but only for the speech-segregation measures and not the recall measures. This confirms, from measures obtained from the same participant, whether young or old, that not all effort measures are necessarily measuring the same aspect of listening effort.

For the ONH group only, a correlation was performed between age and high-frequency pure-tone average (HFPTA; average of hearing thresholds at 1,000, 2,000, and 4,000 Hz). Age and HFPTA were moderately and significantly correlated ($r(16) = .50$, $p < .05$). Given this correlation, partial correlations were computed between age and each of the baseline measures and the dual task effects, controlling for HFPTA. Partial correlations were also computed between HFPTA and each of the baseline measures and the dual task effects, controlling for age.

Controlling for HFPTA, age was moderately and significantly correlated with the SSTM ($r(16) = .56$, $p < .05$) and CST rating ($r(16) = -.57$, $p < .05$), but was weakly and non-significantly correlated with baseline TMR ($r(16) = -.09$, $p > .05$), CST scores ($r(16) = -.25$, $p > .05$), memory span ($r(16) = .09$, $p > .05$), and baseline recall performance ($r(16) = -.21$, $p > .05$). For the dual task effects measured on the DTP, partial correlation analysis indicated that age and the dual task effects for the speech segregation task were not correlated for the CM ($r(16) = .01$, $p > .05$), SM ($r(16) = -.08$, $p > .05$), CF ($r(16) = -.06$, $p > .05$), and SF ($r(16) = -.05$, $p > .05$) conditions. Partial correlation analysis indicated that age and the self-report ratings of effort given after each experimental condition were moderately and significantly correlated only for the CM ($r(16) = -.52$, $p < .05$) [SM ($r(16) = -.34$, $p > .05$), CF ($r(16) = -.36$, $p > .05$) SF ($r(16) = -.11$, $p > .05$)].

Controlling for age, the HFPTA was weakly and non-significantly correlated with the baseline measures of SSTM ($r(16) = -.19$, $p > .05$), baseline TCR ($r(16) = -.01$, $p > .05$), baseline memory span ($r(16) = .08$, $p > .05$), and the CST score ($r(16) = -.25$, $p > .05$) and was moderately correlated with the CST rating ($r(16) = .55$, $p < .05$). For the dual task effects measured on the DTP, partial correlation analysis indicated that HFPTA and the dual task effects for the speech segregation task were weakly and non-significantly correlated for the CM ($r(16) = -.08$, $p > .05$), SM ($r(16) = -.04$, $p > .05$), CF ($r(16) = -.25$, $p > .05$), and SF ($r(16) = -.28$, $p > .05$) conditions. Partial correlation analysis indicated that HFPTA and the dual task effects for the recall task were also non-significantly correlated [CM ($r(16) = .33$, $p > .05$), SM ($r(16) =$

.22, $p > .05$),, CF ($r(16) = -.06$, $p > .05$), SF ($r(16) = -.03$, $p > .05$)]. Partial correlation analysis indicated that the same was true for self-report ratings of effort given after each experimental condition [CM ($r(16) = .40$, $p > .05$), SM ($r(16) = .17$, $p > .05$), CF ($r(16) = .24$, $p > .05$), SF ($r(16) = .19$, $p > .05$)]

Chapter 5. Discussion

The purpose of this study was to measure listening effort during speech-in-speech tasks for young and older adults. The main measure of listening effort made use of a dual-task paradigm (DTP). The DTP was administered concurrently and sequentially to determine if the approaches differed in the measured listening effort. As previously detailed, the primary task in the DTP was a speech- segregation task; the secondary task was a digit-recall task. Dual-task paradigm procedures have been used to measure listening effort following the theoretical assumption that cognitive processing resources required to complete language processing tasks are limited in capacity (Broadbent, 1958; Kahneman, 1973). If the resources required to perform both tasks are greater than available cognitive resources, and the participant is instructed to prioritize a primary task, then a decrease in performance on the secondary task will be observed (Gagne et al., 2017).

The variables of interest in this study were age group of the participant (younger vs. older), experimental condition (concurrent vs. sequential) and gender of the competing talkers in the speech task (male vs. female with male talker as target in all cases). Differences were seen between the younger and the older group in terms of initial measures. Younger adults scored higher than younger adults on both the CST and the SSTM. A statistically significant larger memory span was observed for the recall task for the younger group. Additionally, the older group required a more favorable TCR to complete the speech segregation task, though this difference was not statistically significant. With regard to the dual-task effects measured for each age group, age-group differences were observed as longer response times for the older adults, in particular for the sequential condition. As expected, and consistent with findings from other researchers (e.g., Arehart, et al., 1997; Humes et al., 2006; Lee & Humes, 2012), speech-identification performance was better for both age groups when the gender of the competing talkers differed from that of the target talker, in this case a male target talker and two competing female talkers. This relative increase in performance was seen when the DTP was administered both concurrently and sequentially. In terms of a dual-task effect on the recall measure, the primary outcome measure in this experiment, a greater effect was seen in the sequential condition than in the concurrent condition. In the following pages, patterns observed in the data analysis are reviewed in more detail. Possible factors explaining these patterns are also discussed.

Baseline measures

Younger adults outperformed older adults on both the CST and the SSTM. This would suggest that the individual tasks that comprised this experiment (speech-segregation task and a recall task) would potentially be more difficult for the ONH group. Recall that prior to the DTP measures, both the speech-segregation task and the recall task were administered as single tasks. This served two purposes. First, this provided a comparison of performance when completed alone as opposed to while competing with an additional task. Second, the baseline performance allowed for individual calibration of difficulty level for each of the tasks. Not surprisingly, given the group differences in CST and SSTM performance, the YNH group outperformed the ONH group on both the speech segregation task and the recall task. Even though a statistically significant initial TCR was not seen across age groups ($p = .06$), median values suggest that, overall, the older adults needed a more favorable TCR to achieve the same speech-segregation performance (~80% correct) as the younger adults. The baseline memory spans for the test materials, on the other hand, were significantly lower for the older group. It should be kept in mind that the older group completed the experiment under conditions designed to equate performance with the younger adults at baseline. To do so, the older adults were tested with slightly more favorable TCRs and significantly shorter recall sets in the DTP. This was particularly important given the fact that, although the older adults were considered to have "clinically normal" hearing, hearing thresholds at 1,000 and 4,000 Hz were significantly poorer than those of the young adults in the test ear. In many everyday situations, it may not be possible to present the older listeners with a better TCR or a shorter memory set (e.g., recall of a phone number at a noise restaurant). The focus here, though, was on the extra effort required to perform both tasks together and it was important to equate performance, and presumably the underlying difficulty, for each task separately, compensating for age differences for the single-task measures in the process.

Initial vs. final measures

Recall that single-task measures of both the primary and secondary measures were obtained twice: once at the initial session and again after completion of all laboratory sessions. This was an attempt to evaluate the stability of these single-task baselines, against which the performance in all dual-task conditions is compared when determining the "dual-task effect", the primary measure of listening effort in this experiment. For the younger group, scores did not differ significantly between the initial and final sessions for the speech-segregation task or the recall task. The older group, but not the younger group, demonstrated improved performance on both tasks by the end of the experiment. Given this possible training effect, it is possible that the older adults required a longer training period before initiating the experimental conditions. It was also found that scores on both tasks for the older group were moderately correlated. Even though scores improved from the initial to the final session, the correlation suggests that those who performed poorly initially also performed poorly at the final session and those who performed well initially also performed well at the final session. Given the similarity across sessions for the younger group, and in light of the moderate correlation between sessions and the modest improvements seen in the older adults, the baseline measures were considered to be stable in subsequent analysis of the dual-task effects.

Performance on the dual task

Both speech and recall performance were poorer in the sequential dual-task condition than in the concurrent condition. Response times were slower on the sequential task as well. One possible explanation for the poorer speech-segregation performance in the sequential condition is that participants would have begun the speech-segregation portion of the dual task with the full set of digits to be recalled, interfering with the ability to separate target items from competition. Hunter and Pisoni (2018) evaluated a similar sequential paradigm where they presented digits for later recall and participants were tasked with identifying spectrally degraded sentences. They found reduced speech-recognition performance with the addition of the sequential recall task (Hunter & Pisoni, 2018).

There are at least two possible explanations for the poorer recall performance on the sequential task. First, the delay between presentation and recall of items was greater in the sequential condition. In the sequential condition, the participant was prompted for recall after a full set of sentences, equaling the participant's individual set size, was presented for speech segregation. In the concurrent task, the participant was prompted for recall immediately after the final sentence of the full set of sentences was presented for speech segregation. Retroactive interference is a second possible explanation for poorer recall performance. In both the sequential and the concurrent task, items presented later would interfere with the recall of items presented earlier. However, in the sequential task, participants were tasked with identifying numbers in the speech segregation component of the task that were not to be stored in memory for later recall. These digits presented during the speech-segregation task could certainly have interfered with the participant's ability to recall the initially presented digits.

Performance on the speech-segregation dual-task was better when the gender of the competing talkers was female. As previously discussed, it was expected that participants would take advantage of the cue provided by separation of voice fundamental frequency to aid in speech segregation (e.g., Arehart, et al., 1997; Humes et al., 2006; Lee & Humes, 2012). There was no difference in recall dual-task performance when the gender of the competing talkers was male or female. Given that overall speech-identification performance was about 80% for the male competition and 90% for the female competition, equivalent recall performance for the dual task for these two conditions suggests that the measured dual-task effects were robust across this 80-90% range of performance.

Participants with a lower baseline TCR (those who were able to achieve 79.9% accuracy on the speech-segregation task at lower TCRs) showed poorer performance on the speech-segregation task as a dual task. Even though all participants completed the speech segregation task at the same initial 79.9% performance level, those completing the experiment at less favorable TCRs (i.e., those who did better on the speech task as a single task) performed more poorly on the speech-segregation dual-task. Hunter & Pisoni (2019) explored the impact of cognitive load on speech perception by varying cognitive load, sentence predictability, and spectral degradation. They found a complex interaction between these factors and suggested that varying levels of speech intelligibility are a factor when considering the relationship between cognitive effort and word recognition (Hunter & Pisoni, 2019). In the current experiment, even though baseline performance was the same for all participants, those who completed the DTP with a less favorable TCR demonstrated poorer performance. Thus, although measured speech-identification performance was equivalent at 79.9% the effort expended to get to that level of performance may have been greater for those who did so at lower TCRs than those who did so at more favorable TCRs.

As with baseline speech-identification performance, baseline recall performance was also predictive of dual-task recall scores. Recall that baseline recall performance ranged from 80-100%. Participants whose average baseline recall scores were closer to 80% performed poorly relative to those with baseline recall scores closer to 100% on dual-task recall. This effect was larger for the sequential condition relative to the concurrent condition. This may be related to the slope of the psychometric function for the recall task. As mentioned, it was our intent to equate baseline performance on the speech-identification and recall tasks for the dual task, using performance level as an indirect metric of the effort required. A performance level of approximately 80% was chosen as a level that would elicit measurable changes in effort (i.e. at a level that was not too easy or too difficult) on the dual task. When measuring the psychometric function for memory span, we found that most participants had a very steep slope from poor performance (less than 60% correct) to 100% performance. As a result, more than half of the participants (23 total) completed dual-task paradigm at a set size that yielded 90% accuracy or greater on the recall task, rather than the target of 80%. It is possible that combining the speech and recall task was easier for those completing that task at a set size that was closer to the upper asymptote of their psychometric function for the recall task. That this effect was stronger in the sequential dual task condition is consistent with findings that performance in general was poorer for that condition. For the female-talker competition, however, the overall performance was about 90% for speech-identification, which would have been more closely matched to recall performance. Nonetheless, the recall performance on the dual task was the same for both the male and female competition. Thus, a simple explanation of the association between baseline recall performance and dual-task recall scores based on the imbalance between recall and speech-identification performance may not be plausible. Rather, it may once again reveal that there were variations in effort required to achieve the measured level of performance at baseline such that more reserve remained available in the dual-task conditions for those with higher baseline single-task recall scores.

Response Time

Average response times were faster for the concurrent dual-task condition. An interaction between age group and condition was seen, with substantially shorter response times for the YNH group for both male and female talker competition sequential conditions. Response times varied greatly in the older group for the sequential condition and were less varied for the concurrent condition. It was expected that older adults would exhibit slower response times (Cerella & Hale, 1994). However, this was only evident for the sequential condition. Several factors, together or in isolation, could explain slower response times. A slower response time could arise from slower motor movements in selecting the responses. On the other hand, if this were the sole reason for slower response times, then there would be no reason to expect that responses times would be slower and more varied for the sequential versus the concurrent condition. It is also possible that older adults took longer to "scan" the contents of short-term memory for recall of the digits, and that this process was additionally lengthened for the sequential condition.

Dual-task effects

The rationale for using a recall task to measure listening effort was that as listening effort increased when attempting to identify a speech target while also maintaining digits for later recall in memory, fewer cognitive resources were available for encoding the digits into memory (Picou & Ricketts, 2014). We hypothesized that if the amount of listening effort, as measured by the dual-task effect on the recall task, was significantly different between the concurrent and sequential tasks then this would support the hypothesis that different cognitive processes are being evaluated in each dual-task paradigm (Ohlenforst et al., 2017a; Gagne et al., 2017). Recall that the dual-task effect was calculated by subtracting percent-correct scores during the dual task from percent-correct scores established at baseline (single task). A greater dual-task effect on the secondary task was seen for the sequential condition. However, further analysis showed that dual-task effects for the speech task and the recall task were moderately correlated for the older group (Pearson correlation = 0.35). This suggests that there may be at least some, about 10%, shared variance between these two DTP measures.

Consider the differing nature of the concurrent vs. the sequential task in terms of working memory and verbal rehearsal. The multicomponent model of working memory (Baddeley & Hitch, 1974) proposes three components that collectively comprise working memory, the phonological loop, the visuo-spatial sketchpad, and the central executive. The model proposes that the working memory system temporarily maintains acoustic (including speech) and visual/spatial sequences and configurations in the phonological loop or visuo-spatial sketchpad, respectively. Both of these are controlled by a central executive, which manipulates the information in each subsystem. Speech information is maintained in the phonological loop is through continued attention or the activity of verbal rehearsal. Verbal rehearsal refers to the repetition of speech information, either silently or out loud, as a strategy for maintaining a list of items in working memory (Baddeley, Eysenck, & Anderson, 2015). Different temporal patterns of dual-task presentation (specifically, concurrent vs. sequential) may require different verbal rehearsal strategies.

In the concurrent task, participants were presented with sets of sentences and tasked with identifying a color and number while also remembering that same number for later recall. This required that the participant identify the target number while simultaneously updating the set of digits held in short term for later recall. This would require verbally rehearsing the digits for later recall while also segregating the target color and number in each new trial, updating the set of digits for later recall after each trial. In the sequential task, participants were presented with the full set of digits to be recalled prior to completing the speech segregation task. This would also require verbally rehearsing the digits for later recall while also segregating target color and number. As previously mentioned, in the sequential task, the potential for retroactive interference is a potential explanation for the poorer performance on this task. To complete the sequential DTP, the participant would have had to encode the full set of digits, perform the speech segregation task with a different set of digits, and then recall the initial set of digits. This, along with the fact that the delay between presentation and recall prompt was greater for the sequential task, provides a possible explanation for differences seen between the two tasks.

A greater effect was seen for male over female talker competition on the speech-segregation task. This should not be interpreted as a general advantage for conditions with competing female speech. Rather, it is just one possible instantiation of "same gender" versus "different gender" conditions for the target and competing talkers. That is, had the target talker been female instead of male, the female competition would have yielded lower scores than the male competition (Humes et al., 2006: Lee & Humes, 2012). It was expected that overall performance would be better when the gender of the competing talkers was female rather than male, given that differences in fundamental frequency aid in segregating a speech target from competing speech (e.g., Arehart, et al., 1997; Humes et al., 2006; Lee & Humes, 2012).

In some cases, even with male talker competition, superior performance was observed in the dual-task condition for either the speech-identification or the recall dual-task. It is possible that the tasks themselves were not difficult enough for some participants and combining the two as a dual task was not particularly challenging. It is also possible that this effect reflects the improvement that was seen between initial and final measures. A third possibility is that some participants tried harder in the dual-task paradigm because it was a more demanding task.

No significant difference in recall effect was seen for male vs. female competition. It was hypothesized that a smaller effect would be seen when the speech task was made relatively easier (female talker competition). However, this effect was not seen. The differing difficulty levels of the speech task did not affect performance on the recall task.

A main finding of this study was that the analysis of the dual-task effects showed no significant effect of age. Recall, however, that the mean memory span for the experimental stimuli was significantly lower for older adults. Further, a trend was seen where older adults needed a more favorable TCR at baseline to achieve the target 79.9% performance level. Other researchers have also found that age-related differences are not significant when materials are presented at difficulty levels individually measured for the participant (e.g., Schneider et al., 2000). Speech communication difficulty experienced by older adults is affected by “cognitive stressors” (e.g. memory load) and “perceptual stressors” (e.g. hearing loss and background noise) (Pichora-Fuller, 2003). As noted previously, for conditions in which it is not possible to individually tailor the acoustics or the set size, such as listening to and recalling a phone number in a noisy restaurant, older adults with normal hearing may well perform worse than young adults tested under equivalent conditions. Here, conditions were manipulated on an individual basis in an effort to equate performance, and presumably effort, for each task. Thus, the findings of no age effects on the dual-task effect should not be generalized to other conditions and studies in which the performance levels on each task have not been equated on an individual basis prior to the dual-task conditions.

The size and “direction” of the dual-task effect was examined in more detail. Since the dual-task effect was calculated by subtracting performance scores on the dual-task from baseline scores, it would have been possible for the dual-task effect to show either decreased or improved performance during the dual-task. For each participant, the percentage of time where the dual-task effect indicated poorer performance, improved performance, or no effect on the dual-task relative to the single task was calculated. First, for each individual, there were multiple estimates of dual-task performance obtained. These repeated-measures estimates were averaged with the mean and standard deviation established for each subject. Mean effects that were within one standard deviation of zero were considered measurement variability and were tallied as “no effect”. Those individual means that differed by more than one standard deviation from zero, in either direction, were then tallied across the groups. Figure 17 shows the number of instances per condition where performance on the dual-task paradigm was flagged as being poorer than baseline (bottom segment), no different from baseline (middle segment), and better than baseline (top segment).

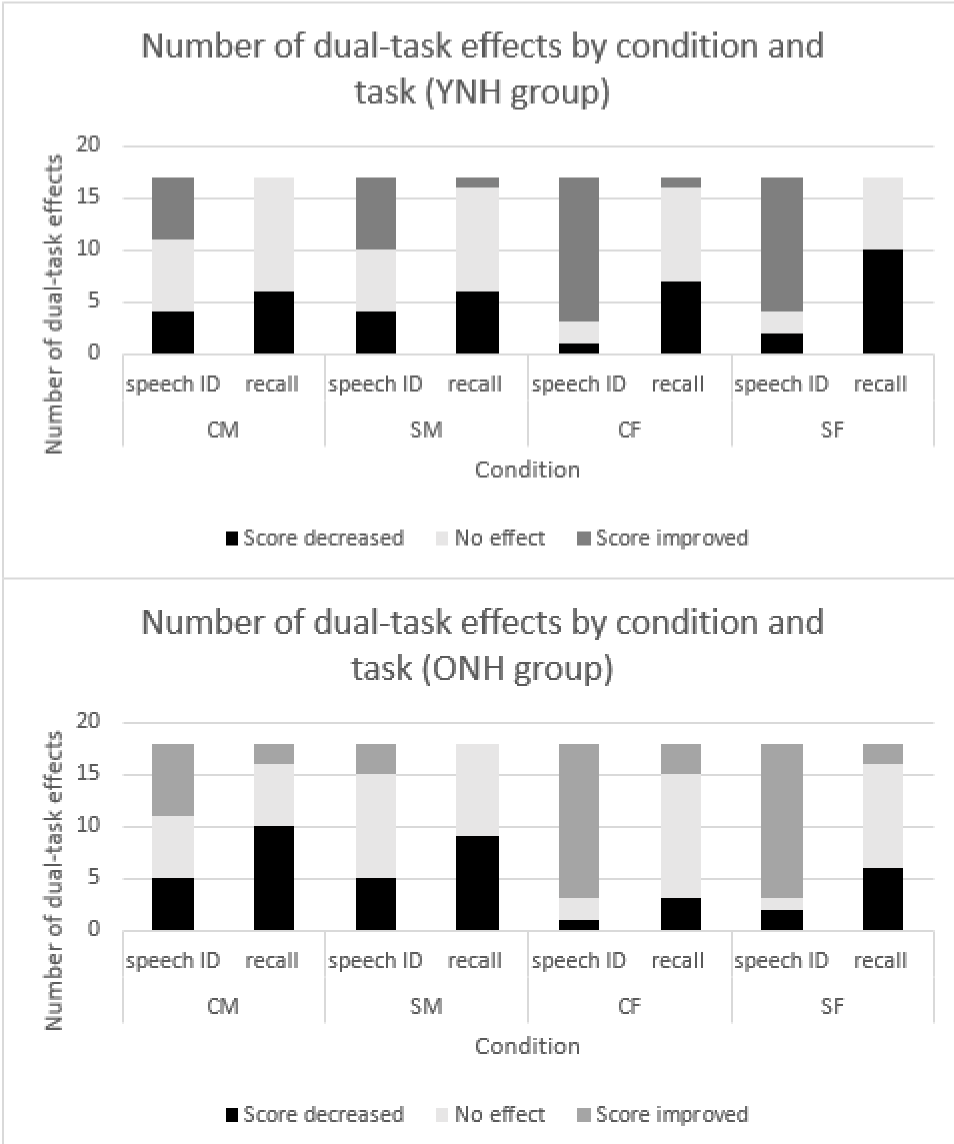

Figure 17. Number of dual-task effects reflecting decreased performance, no effect, and improved performance, by condition, for each task and participant group. YNH (top panel) = young normal hearing (N=17); ONH (bottom panel) = older normal hearing (N=18); CM = concurrent dual task, male competition; SM = sequential dual task, male competition; CF = concurrent dual task, female competition; SF = sequential dual task, female competition.

For the group as a whole, an effect of poorer performance on the dual-task paradigm was seen 29% of the time across all conditions. This effect was seen 36% of the time for the speech-identification task and 79% of the time for the recall task. This effect was seen 53% of the time in the concurrent condition and 61% of the time in the sequential condition. This effect was seen 70% of the time with male talker competition and 44% of the time with female talker competition. The converse of this effect, seen 39% of the time across all conditions, was that performance improved during the dual-task paradigm. Across all conditions, no effect in performance on the dual-task paradigm was seen 32% of the time.

As discussed, it was not unexpected for performance to improve on the speech task when the gender of the competing talkers was female. In all cases, the target voice was the same male talker; it was expected that the speech segregation task would have been easier with female competition (e.g., Arehart, et al., 1997; Humes et al., 2006; Lee & Humes, 2012). When examining correlations between both components of the dual task, a pattern emerged where the majority of participants exhibited an improvement in performance on the speech-segregation task while also showing a decrease in performance on the recall task. This suggests that most participants were trying to focus on the speech segregation task as instructed. When doing so, their performance improved over baseline. Because participants were focusing their effort on the speech segregation dual-task, performance on the recall task suffered compared to baseline.

Ratings of Effort

Self-report ratings of listening effort were obtained following the CST testing and also multiple times per experimental condition. Despite the relatively poorer performance on the CST, self-report effort ratings did not reflect expenditure of greater effort by the older group. On the other hand, individual baseline TCR was predictive of how the participant rated the experimental conditions. Participants who achieved 79.9% accuracy on the speech-identification task as a single task at a less favorable TCR gave ratings across conditions reflecting greater effort was expended than desired. This is noteworthy, given that performance on the speech task as a single task was the same for everyone at baseline (79.9%). It appears that those who were able to achieve 79.9% accuracy on the speech task alone at less favorable TCRs noted greater effort to complete the task when paired with the secondary recall task. This finding of increased self-report rating of effort is consistent with the fact that participants with smaller individual TCR also demonstrated poorer performance on dual-task speech segregation.

Participants' self-report ratings of effort for the concurrent task suggested that this condition was more difficult than the sequential task. This is contrary to the measured dual-task effects which indicated greater effort for the sequential paradigm. Recall that the wording of the self-report rating of listening effort was based on the work of Picou and colleagues (Picou et al., 2017, and Picou & Ricketts, 2018). These researchers evaluated relationships between self-report questions and behavioral measures of listening effort. They found that their question probing a listener's desire to control a listening situation correlated with behavioral measures of listening effort. There are several possible explanations for why the findings of the current study are only in partial agreement with this previous work. First, the wording of the question in the current study was not exactly the same, with the current study asking "On a scale of 1 to 10, please rate your desire to improve the listening situation." and the Picou et al. work asking "How likely would you be to try to do something else to improve the situation (e.g., move to a quiet room, ask the speaker to speak louder)?" (Picou et al., 2017, and Picou & Ricketts, 2018). Second, in the previous work, participants were asked to rate not only their desire to improve the listening situation, but also their mental work, tiredness, and desire to give up. In the current work, just the one question was asked. It is unknown whether the correlation between the desire to improve a listening situation and listening effort expended as measured on a behavioral task is less robust when the question is asked in isolation. It is also possible that the way the question was presented in the current experiment simply did not fully capture the participant's listening effort. Listening effort is conceptualized as complex and multidimensional (Pichora-Fuller et al., 2016; Moore & Picou, 2018). Moore and Picou (2018) recently reported findings where participants likely substitute a simpler question when asked to provide a rating probing listening effort. In other words, the question the participant appears to be answering is not the same question as the one the researcher is asking. In the current study, increased self-report ratings of effort appeared to follow performance on the DTP when considering the TCR at which participants performed the DTP but were not in agreement with the dual-task effects when considering the concurrent vs. the sequential condition. It is possible that asking participants to rate their desire to change the listening situation was not the right question to tap into differences in listening effort among the concurrent and sequential conditions.

A further possible explanation for the relatively higher self-report ratings of effort for the concurrent condition is that the concurrent condition required greater amounts of cognitive flexibility. As previously noted, a greater degree of cognitive shifting could conceivably be required in a concurrent DTP, where the study participant would need to switch between the speech perception task and the recall task at the same time. By comparison, in a sequential DTP, a participant could dedicate full attention to the pre-load material prior to completing the speech perception task. Perhaps the increased self-report ratings of difficulty for the concurrent condition reflect a greater need for cognitive shifting. On the other hand, a cost in terms of behavioral performance would be expected for a task requiring a great deal of cognitive shifting (e.g. Hirsch, Nolden, Declerck, & Koch, 2018). In this study a greater performance cost was seen in the sequential condition. Perhaps the rating question asked in the current study taps into cognitive processes not measured by performance costs.

Correlations among measures

Baseline measures of speech performance in noise (CST) and working memory (SSTM) were weakly correlated with baseline performance on the tasks used in the DTP; with a moderate correlation seen between CST and baseline speech-segregation performance for the young group alone. One possible explanation is that the tasks differed in key factors influencing performance. The SSTM measured working memory using the visual modality whereas the recall task used here was an auditory task. The CST differed from the current speech segregation task in terms of syntactic structure of the speech materials as well as predictability of target words.

Baseline self-report ratings of effort were moderately correlated with ratings given during the DTP. Baseline self-report ratings of effort were not strongly correlated, or were even negatively correlated, with observed DTEs. As discussed, it is possible that the rating prompt used here evaluated cognitive processes not measured by the DTE.

Dual-task effects observed for the concurrent speech task were strongly correlated with those seen for the sequential speech task for the older group and moderately correlated for the younger group. Dual-task effects observed for the concurrent recall task were moderately correlated with those seen for the sequential recall task for both groups. The moderate correlation seen on the recall task could be an indication that participants were changing how they prioritized attention to the recall task across the concurrent and sequential conditions. To the extent that these sequential and concurrent estimates of DTE are strongly correlated, the impact of this measurement parameter is less important. That is, those who did best using one method tended to also do best on the other. Thus, the relative ordering of participants within a group was fairly constant across both the sequential and concurrent methods.

For the older group, hearing thresholds were not strongly or significantly correlated with baseline or outcome measures. Given that each participant completed the study at an individualized difficulty level, strong correlations between hearing loss and dual-task performance would not have been expected. Hearing thresholds were moderately correlated with participant age. Controlling for hearing levels, greater age was positively correlated (r = .56) with SSTM performance. It is not clear why this would be the case, considering that memory abilities tend to decline with age (Salthouse, 2004). On the other hand, consistent with the literature, older adults performed significantly worse than younger adults on the SSTM in this study (Figure 4). Age was also negatively correlated with self-report ratings of effort on the CST (r = -.57) and on the CM condition of the dual-task paradigm (r = -.52), when controlling for hearing levels. Thus, for these self-report measures of effort, the older the ONH participant, the lower the self-reported listening effort. Perhaps, this reflects the surprisingly better cognitive processing (SSTM) in these same older ONH participants. Controlling for age, hearing levels were correlated with self-report ratings of effort on the CST (r=.55). Here, the greater the high-frequency hearing loss among the ONH participants, the greater the self-reported effort, at least for the CST.

Study Limitations

Even though the speech task was the primary task in the DTP, participants exhibited a greater dual-task effect for the speech task over the recall task 11% of the time. Participants were instructed to do their best on the speech dual-task, but perhaps instructions could have been worded in a way to ensure greater focus on the primary task. In the end, we can’t know or predict where listeners will focus their attention during the experiment. As discussed in Gagne et al., (2017), it may be difficult to interpret data from experiments using the dual-task paradigm if participants are not consistently prioritizing the primary task over the secondary task. (Gagne et al., 2017). Another factor that is difficult to control is individual motivation, an important aspect of listening effort (Pichora-Fuller et al., 2016).

Another limitation of the study was the sample size. Although a total 35 were able to complete the study, the sample was somewhat small considering the number of variables under investigation. An efficient, mixed design was used in which multiple measures per subject were taken. The sample size of 35 can be considered small but adequate for the research design.

Although we have characterized both groups as normal hearing and differing in age alone, the hearing thresholds of the ONH group still were significantly worse than the YNH group at 1,000 and 4,000 Hz, and likely at higher frequencies as well. Related to this, as noted earlier, many older adults, about 40%, have hearing loss and the results obtained here may not generalize to older adults with impaired hearing.

The speech-identification task used here is challenging. The target sentence and the competing sentences have the same temporal structure and grammatical structure, differing only in talker gender and call sign as cues for segregating target from competition. This was intentional to minimize variability in stimuli and in the nature of the segregation cues available. Further, they had no real semantic information to aid in speech identification as the content was virtually identical for the target and competing sentences. The materials, however, are certainly not representative of everyday speech and results obtained may not generalize to everyday listening as a result. In addition, the use of two competing talkers is among the more difficult competing-speech conditions in general and for the CRM in particular (Humes, et al., 2017). Competing speech comprised of 6 or more talkers, for example, is somewhat more difficult acoustically as the gaps in the fluctuations of individual competing talkers are filled by other competing talkers, but it is less distracting in that the listener can't decipher individual messages or content that might interfere cognitively with the target. At the other end of the continuum, having only one competing talker is acoustically the easiest with lots of gaps in the competing signal through which the target could be heard, but cognitively may be the most distracting as the content of the competing speech is as easy to decipher as that of the target. Two-talker competition is somewhere in-between single-talker and babble (> 6 talkers) competition and the results obtained here may not generalize to those other listening conditions.

A *limited-capacity cognitive resource model* (Kahneman 1973) is assumed here as a way of explaining dual task effects observed in the dual-task paradigm, whereby cognitive resources are allocated for task processing which would otherwise be available for processing of additional tasks. This type of model assumes that the tasks are processed in parallel, with cognitive resources allocated disproportionally to the primary task. Other models have been proposed to explain behavioral performance costs in a general multitasking paradigm. *Capacity sharing models* (e.g. Tombu & Jolicoeur, 2003) have also been proposed, where cognitive resources are allocated to each task in a graded manner. This type of model assumes that tasks are processed serially, with task allocation adjusting alongside difficulty of the secondary task. Fischer & Plessow (2015) argue that individuals shift between parallel and serial processing strategies to maximize efficiency of processing. Considering this, it is reasonable to hypothesize that the participants in this study adapted their processing strategies differently to the concurrent and the sequential dual tasks.

Finally, as stated several times in this chapter, care was taken to equate the baseline performance levels for each single task assuming that this would make each about equally difficult in isolation. To do so, TCRs and set sizes were adjusted on an individual basis rather than fixing both for all participants. The results obtained here regarding the effects of age group and paradigm may not generalize to DTP studies in which baseline performance levels are unequal.

Chapter 6. Conclusions

The purpose of this study was to examine listening effort in young and old adults when listening to speech in a background of competing speech. A dual-task paradigm (DTP) was used to measure listening effort and it was administered concurrently and sequentially. The primary task in the DTP was a speech-segregation task; the secondary task was a digit-recall task. The rationale for using a recall task to measure listening effort was that as listening effort increases while attempting to identify a speech target in a background of competing speech, fewer cognitive resources are available for encoding items in memory for later recall (Picou & Ricketts, 2014). The variables of interest in this study were age group of the participant (younger vs. older), experimental condition (concurrent vs. sequential) and gender of the competing talkers in the speech task (male vs. female with male talker as target in all cases).

Each task in the dual task condition was first measured as a single task. For the speech segregation task, the target-to-competition ratio (TCR) yielding a performance level of 79.9% was used. For the digit recall task, a set size yielding at least 80% accuracy was used. There was a trend for older adults to need a more favorable (larger) TCR in order to achieve 79.9% accuracy. The older adults needed a significantly smaller set size on the recall task in order to achieve at least 80% accuracy. Thus, the older adults completed the experiment under relatively easier parameters. Given this, one main finding from the study was that there were no effects of age on listening effort as measured by changes in performance on the DTP.

We hypothesized that if the amount of listening effort, as measured by the dual-task effect on the recall task, was significantly different between the concurrent and sequential tasks then this would support the hypothesis that different cognitive processes are being evaluated in each dual-task paradigm (Ohlenforst et al., 2017a; Gagne et al., 2017). A number of differences between the concurrent and sequential conditions emerged. A larger dual-task effect on the secondary task was seen for the sequential condition. Additionally, performance on both components of the dual task were poorer in the sequential condition. Response times were slower for the sequential condition for both age groups. An effect of age was seen in slower response times on the sequential task but not the concurrent task.

Regarding general measurement methods for listening effort, this study found that the two self-report measures of effort, on the CST and during the dual-task paradigm, were generally moderately and significantly correlated. Neither, however, was correlated with the DTE-based measures of effort. On the other hand, the two DTE-based approaches, sequential and concurrent, are moderately to strongly correlated with one another, but only for the speech-segregation measures and not the recall measures. This confirms, from measures obtained from the same participant, whether young or old, that not all listening- effort measures are necessarily measuring the same aspect of listening effort.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, *47*(sup2), S53–S71.

Arehart, K. H., King, C. A., & McLean-Mudgett, K. S. (1997). Role of fundamental frequency differences in the perceptual separation of competing vowel sounds by listeners with normal hearing and listeners with hearing loss. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1434–1444.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47-90). New York, NY: Academic Press.

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2015). *Memory*. New York: Psychology Press.

Bentler, R. A., & Duve, M. R. (2000). Comparison of hearing aids over the 20th century. *Ear and Hearing*, *21*(6), 625–639.

Bess, F. H., & Hornsby, B. W. (2014). Commentary: Listening can be exhausting-Fatigue in children and adults with hearing loss. *Ear and Hearing*, *35*(6), 592.

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—A review. *Trends in Amplification*, *17*(2), 75–93.

Bolia, R. S., Nelson, W. T., Ericson, M. A., & Simpson, B. D. (2000). A speech corpus for multitalker communications research. *The Journal of the Acoustical Society of America*, *107*(2), 1065–1066.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Broadbent, D. E. (1958). *Perception and communication.* Elmsford, NY: Pergamon Press. http://dx. doi. org/10.1037/10037-000.

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

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, *110*(5), 2527–2538.

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging*, *17*(1), 85.

Carhart, R. (1946). Tests for selection of hearing aids. *The Laryngoscope*, *56*(12), 780–794.

Carhart, R., & Tillman, T. W. (1970). Interaction of competing speech signals with hearing losses. *Archives of Otolaryngology*, *91*(3), 273–279.

Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life span. *Acta Psychologica*, *86*(2–3), 109–197.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Cox, R. M., Alexander, G. C., & Gilmore, C. (1987). Development of the connected speech test (CST). *Ear and Hearing*, *8*(5 Suppl), 119S-126S.

Craik, F. I., & Byrd, M. (1982). Aging and cognitive deficits. In *Aging and cognitive processes* (pp. 191–211). New York: Springer.

Cruickshanks, K. J., Tweed, T. S., Wiley, T. L., Klein, B. E., Klein, R., Chappell, R., Nondahl, D.M., & Dalton, D. S. (2003). The 5-year incidence and progression of hearing loss:

the epidemiology of hearing loss study. *Archives of Otolaryngology–Head & Neck Surgery, 129*(10), 1041-1046.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.

Desjardins, J. L., & Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear and Hearing*, *34*(3), 261–272.

Downs, D. W. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders*, *47*(2), 189–193.

Eddins, D. A., & Liu, C. (2012). Psychometric properties of the coordinate response measure corpus with various types of background interference. *The Journal of the Acoustical Society of America*, *131*(2), EL177–EL183.

Feuerstein, J. F. (1992). Monaural versus binaural hearing: Ease of listening, word recognition, and attentional effort. *Ear and Hearing*, *13*(2), 80–86.

Fischer, R., & Plessow, F. (2015). Efficient multitasking: parallel versus serial processing of multiple tasks. *Frontiers in psychology*, *6*, 1366.

Folstein, M. F., Folstein, S. E., White, T., & Messer, M. A. (2010). *Mini Mental State Examination,* 2nd Edn. Lutz: Psychological Assessment Resources. Inc.

Füllgrabe, C., Moore, B. C., & Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience, 6*.

Fulton, S. E., Lister, J. J., Bush, A. L. H., Edwards, J. D., & Andel, R. (2015). Mechanisms of the hearing–cognition relationship. *Seminars in Hearing*, *36*, 140–149.

Gagne, J.-P., Besser, J., & Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in Hearing*, *21*, 2331216516687287,

Hardin, J. W., & Hilbe, J. M. (2012). *Generalized estimating equations*. London: Chapman & Hall/CRC.

Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America*, *136*(2), 748–759.

Hirsh, I. J., Davis, H., Silverman, S. R., Reynolds, E. G., Eldert, E., & Benson, R. W. (1952). Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders*, *17*(3), 321–337.

Hirsch, P., Nolden, S., Declerck, M., & Koch, I. (2018). Common cognitive control processes underlying performance in task-switching and dual-task contexts. *Advances in Cognitive Psychology*, *14*(3), 62-74.

Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing*, *34*(5), 523–534.

Houtgast, T., & Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *International Journal of Audiology*, *47*(6), 287–295.

Humes, L. E. (1999). Dimensions of hearing aid outcome. *Journal of the American Academy of Audiology*, *10*(1), 26–39.

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology, 18*(7), 590–603.

Humes, L. E., Christensen, L. A., Bess, F. H., & Hedley-Williams, A. (1997). A comparison of the benefit provided by well-fit linear hearing aids and instruments with automatic reductions of low-frequency gain. *Journal of Speech, Language, and Hearing Research*, *40*(3), 666–685.

Humes, L. E., & Coughlin, M. (2009). Aided speech-identification performance in single-talker competition by older adults with impaired hearing. *Scandinavian Journal of Psychology, 50*(5), 485–494.

Humes, L. E., & Dubno, J. R. (2010). Factors affecting speech understanding in older adults. In *The aging auditory system* (pp. 211–257). NewYork: Springer.

Humes, L. E., Kidd, G. R., & Fogerty, D. (2017). Exploring use of the coordinate response measure in a multitalker babble paradigm. *Journal of Speech, Language, and Hearing Research*, *60*(3), 741–754.

Humes, L. E., Kidd, G. R., & Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front Syst Neurosci2013b755*.

Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *The Journal of the Acoustical Society of America*, *120*(5), 2926–2937.

Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., & Lee, L. (1994). Factors associated with individual differences in clinical measures of speech recognition among the elderly. *Journal of Speech, Language, and Hearing Research*, *37*(2), 465–474.

Humes, L. E., & Young, L. A. (2016). Sensory–Cognitive Interactions in Older Adults. *Ear and Hearing*, *37*, 52S-61S.

Hunter, C. R., & Pisoni, D. B. (2018). Extrinsic cognitive load impairs spoken word recognition in high-and low-predictability sentences. *Ear and Hearing*, *39*(2), 378.

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Jerger, J., & Jerger, S. (1980). Measurement of hearing in adults. *Otolaryngology*, *2*, 1225–1250.

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

Killion, M. C. (1997). Hearing aids: Past, present, future: Moving toward normal conversations in noise. British Journal of Audiology, 31, 141-148.

Lee, J. H., & Humes, L. E. (2012). Effect of fundamental-frequency and sentence-onset differences on speech-identification performance of young and older adults in a competing-talker background. *The Journal of the Acoustical Society of America*, *132*(3), 1700–1717.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477.

Lewandowsky, S., Oberauer, K., Yang, L.-X., & Ecker, U. K. (2010). A working memory test battery for MATLAB. *Behavior Research Methods*, *42*(2), 571–585.

Lin, F. R., Niparko, J. K., & Ferrucci, L. (2011). Hearing loss prevalence in the United States. *Archives of Internal Medicine*, *171*(20), 1851–1853.

Lunner, T., Rudner, M., & Rönnberg, J. (2009). Cognition and hearing aids. *Scandinavian Journal of Psychology*, *50*(5), 395–403.

MATLAB Release 2013a, The MathWorks, Inc., Natick, Massachusetts, United States

Meister, H., Rählmann, S., & Walger, M. (2018). Low background noise increases cognitive load in older adults listening to competing speech. *The Journal of the Acoustical Society of America*, *144*(5), EL417–EL422.

Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H., & Munro, K. J. (2014). Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PloS One*, *9*(9).

Moore, T. M., & Picou, E. M. (2018). A potential bias in subjective ratings of mental effort. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2405–2421.

Ohlenforst, B., Zekveld, A. A., Jansma, E. P., Wang, Y., Naylor, G., Lorens, A., Lunner, T., & Kramer, S. E. (2017). Effects of Hearing Impairment and Hearing Aid Amplification on Listening Effort: A Systematic Review. *Ear and Hearing*, *38*(3), 267.

Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., Versfeld, N. J., & Kramer, S. E. (2017). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hearing Research*, *351*, 68–79.

Peelle, J. E., Troiani, V., Grossman, M., & Wingfield, A. (2011). Hearing loss in older adults affects neural systems supporting speech comprehension. *Journal of Neuroscience*, *31*(35), 12638–12643.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology*, *42*(sup2), 26–32.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., & Mackersie, C. L. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing*, *37*, 5S-27S.

Pichora-Fuller, M. K., & Singh, G. (2006). Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification*, *10*(1), 29–59.

Picou, E. M., Moore, T. M., & Ricketts, T. A. (2017). The effects of directional processing on objective and subjective listening effort. *Journal of Speech, Language, and Hearing Research*, *60*(1), 199–211.

Picou, E. M., & Ricketts, T. A. (2014). The effect of changing the secondary task in dual-task paradigms for measuring listening effort. *Ear and Hearing*, *35*(6), 611–622.

Picou, E. M., & Ricketts, T. A. (2018). The relationship between speech recognition, behavioural listening effort, and subjective ratings. *International Journal of Audiology*, *57*(6), 457–467.

Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *The Quarterly Journal of Experimental Psychology*, *20*(3), 241–248.

Rakerd, B., Seitz, P. F., & Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear and hearing*, *17*(2), 97-106.

Roberts, K. L., & Allen, H. A. (2016). Perception and cognition in the ageing brain: A brief review of the short-and long-term links between perceptual and cognitive decline. *Frontiers in Aging Neuroscience*, *8*.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., Signoret, C., Stenfelt, S., & Pichora-Fuller, M. K. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*, 31.

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, *47*(sup2), S99–S105.

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140–144.

Schneider, B. A., Daneman, M., Murphy, D. R., & See, S. K. (2000). Listening to discourse in distracting settings: The effects of aging. *Psychology and Aging*, *15*(1), 110.

Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology*, *26*(2), 183–196.

Smith, S. L., Pichora-Fuller, M. K., & Alexander, G. (2016). Development of the word auditory recognition and recall measure: A working memory test for use in rehabilitative audiology. *Ear and Hearing*, *37*(6), e360–e376.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, *28*(3), 455–462.

Tombu, M., & Jolicœur, P. (2002). All-or-none bottleneck versus capacity sharing accounts of the psychological refractory period phenomenon. *Psychological research*, *66*(4), 274-286.

Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*(6), 1038.

Van Rooij, J., Plomp, R., & Orlebeke, J. F. (1989). Auditive and cognitive factors in speech perception by elderly listeners. I: Development of test battery. *The Journal of the Acoustical Society of America*, *86*(4), 1294–1309.

Ward, K. M., Shen, J., Souza, P. E., & Grieco-Calub, T. M. (2017). Age-related differences in listening effort during degraded speech recognition. *Ear and Hearing*, *38*(1), 74–84.

Wilson, R. H. (2011). Clinical experience with the words-in-noise test on 3430 veterans: Comparisons with pure-tone thresholds and word recognition in quiet. *Journal of the American Academy of Audiology*, *22*(7), 405–423.

Wu, Y.-H., Stangl, E., Chipara, O., Hasan, S. S., Welhaven, A., & Oleson, J. (2018). Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss. *Ear and Hearing*, *39*(2), 293–304.

Wu, Y.-H., Stangl, E., Zhang, X., Perkins, J., & Eilers, E. (2016). Psychometric functions of dual-task paradigms for measuring listening effort. *Ear and Hearing*, *37*(6), 660.

Xia, J., Nooraei, N., Kalluri, S., & Edwards, B. (2015). Spatial release of cognitive load measured in a dual-task paradigm in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *137*(4), 1888–1898.

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33.

Curriculum Vitae

KIMBERLY SKINNER, AuD, PhD

kimberlyskinner@atsu.edu

Education

2014-2020	**Doctor of Philosophy** Major: Speech and Hearing Sciences Minor: Psychology Co-Mentors: Larry Humes, PhD, Jennifer Lentz, PhD Indiana University
2002	**Doctor of Audiology** Arizona School of Health Sciences, A. T. Still University
1997	**Master of Science** Major: Speech Pathology and Audiology California State University, Sacramento
1992	**Bachelor of Arts** Major: Linguistics University of California at Davis

Professional Experience

June 2019 – present	**Assistant Professor** Audiology Department A.T. Still University
Spring 2018-Spring 2019	**Associate Instructor** Department of Speech and Hearing Sciences, Indiana University
Fall 2017	**Educational Audiologist** Indianapolis Public Schools, Indianapolis, IN
2015-2019	**Research Assistant/Research Audiologist** Audiology Research Lab, Indiana University, Larry Humes, PI
2015-2016	**AuD Student Clinical Supervisor** Indiana University Speech and Hearing Clinic
2003-2009	**Adjunct Instructor** A.T. Still University

1998-2014	**Audiologist** **CFY supervisor** Kimberly G. Skinner, Au.D. (Practice Owner) San Luis Obispo, California Paso Robles, California
1996-1998	**Audiologist** San Luis Hearing Aid and Audiology San Luis Obispo, California (Patti Johnstone, Director)

Teaching Experience

June 2019-present	**Assistant Professor** Audiology Department A.T. Still University AUDE 5140 "Auditory Science" AUDE 5240 "Essentials of Audiology I" AUDE 5340 "Essentials of Audiology II" AUDE 5400 "Speech Perception" AUDE 7340 "Hearing Loss and Healthy Aging"
Spring 2019	Certificate earned Graduate Teaching Apprenticeship Program Center for Innovative Teaching and Learning Indiana University
Spring 2018-Spring 2019	**Instructor**, Indiana University Department of Speech and Hearing Sciences SPHS 676 "Advanced Amplification Seminar" SPHS 516 "Introduction to Audiologic Assessment" SPHS 779 "Business Practices"
2015-2019	**Guest Lecturer**, Indiana University Speech and Hearing and Psychological and Brain Sciences Departments SPHS 201 "Speech Anatomy and Physiology" Topic: Introduction to Auditory Physiology SPHS 106 "Overview of Hearing Science" Topics: Introduction to Clinical Audiologic Assessment; Aging and Hearing Loss PSY 329 "Sensation and Perception" Topic: Introduction to Cochlear

Implants (taught to three different classes)

SPHS 516 “Introduction to Audiologic Assessment”

Topics: Otoscopy, case history, bone conduction/tuning fork testing, use of self-report measures

SPHS 561 “Clinical Methods and Practices” Topic: Otoscopy

2003-2009 **Adjunct Instructor**, Arizona School of Health Sciences

Distance Education instructor for transitional and residential AuD students

Courses taught: AUD852 “Professionalism: Leadership and Development”; AUD853 “Ethics in Audiology” (assisted primary instructor)

Credentials/Licensure

2019-present	Licensed audiologist, Arizona
2015-present	Licensed audiologist, Indiana
1996-2015	Licensed audiologist, California
Earned 1997	Certificate of Clinical Competence in Audiology American Speech-Language-Hearing Association

Publications

1. Humes, L. E., **Skinner, K. G.**, Kinney, D. L., Rogers, S. E., Main, A. K., & Quigley, T. M. (2019). Clinical Effectiveness of an At-Home Auditory Training Program: A Randomized Controlled Trial. *Ear and hearing*., *40*(5), 1043-1060.

2. Lentz, J. J., Walker, M. A., Short, C. E., & **Skinner, K. G.** (2017). Audiometric Testing With Pulsed, Steady, and Warble Tones in Listeners With Tinnitus and Hearing Loss. *American Journal of Audiology*, *26*(3), 328-337.

3. **Skinner, K.** & Lentz, J. (2017). Macroscopic versus Microscopic Spectral Weighting of Speech Stimuli. *Proceedings of Meetings of the Acoustical Society of America*. *Volume 26*, 050005. DOI: 10.1121/2.0000369.

Research Presented

1. (2020) **Skinner, K.,** Maxwell, B., Baskerville, A., Milanović, J. Audiology Students Learn Interprofessional Collaboration. Poster presented at the American Academy of Audiology virtual poster session, New Orleans, Louisiana.

2. (2019) Shen, Y., He, Y., **Skinner, K.,** Yun, D. The Effect of Reverberation on Listening Effort: A Dual-Task Study. Poster presented at the Acoustical Society of America meeting, Louisville, Kentucky.

3. (2018) **Skinner, K.,** Humes, L., Kinney, D., Rogers, S., Quigley, T. Hearing Aid Maintenance in Older Adults: Time Course and Intervention. Poster presented at the International Hearing Aid Research Conference, Tahoe City, California.

4. (2017) **Skinner, K.** The Lombard Effect in Speakers with Hearing Loss. Research colloquium, department of Speech and Hearing Sciences, Indiana University.

5. (2017) **Skinner, K.**, Humes, L., Kinney, D., Rogers, S., Main, A., & Quigley, T. Hearing Aid Maintenance in Older Adults. Poster presented at the American Academy of Audiology Conference, Indianapolis, Indiana.

6. (2016) Main, A., Humes, L., Kinney, D., Rogers, S., Quigley, T., & **Skinner, K.** A Randomized Controlled Trial of the Clinical Effectiveness of an At-Home Auditory Training Program. Poster presented at the International Hearing Aid Research Conferences, Tahoe City, California.

7. (2016) Lentz, J. & **Skinner, K.** Comparison of Different Procedures to Assess Spectral Weights for Speech-Like Sounds. Poster presented at the Acoustical Society of America meeting, Salt Lake City, Utah.

8. (2016) **Skinner, K.**, Forrest, K., & Dutta, M. The Effects of Age and Hearing Loss on the Lombard Effect. Poster presented at the American Academy of Audiology Conference, Phoenix, Arizona.

9. (2015) Short, C., Lentz, J., Walker, M., & **Skinner, K.** Preference of tone types for audiometric testing by patients with tinnitus. Poster presented at the International Tinnitus Research Initiative Conference, Ann Arbor, Michigan.

10. (2015) **Skinner, K.** & Forrest, K. The effects of age and hearing loss on the Lombard Effect. Poster presented at the Aging and Speech Communication Conference, Bloomington, Indiana.

11. (2015) **Skinner, K.** & Lentz, J. Spectral weighting of voiceless fricatives in young and elderly listeners. Poster presented at the Aging and Speech Communication Conference, Bloomington, Indiana.

Competitive Awards

2018	Ph.D Scholarship Council of Academic Programs in Communication Sciences and Disorders

Grants Awarded

2019	Department of Speech and Hearing Sciences research grant Indiana University

	Award amount: $500.00
2018	College of Arts and Sciences Travel Award Indiana University Award amount: $300.00
2017	Provost's Travel Award for Women in Science Indiana University Award amount: $500.00
2017	Travel Award Aging and Speech Communication Conference 2017 Award amount: $500.00
2016	Provost's Travel Award for Women in Science Indiana University Award amount: $500.00
2015	Provost's Travel Award for Women in Science Indiana University Award amount: $300.00
2015	Lion's Club of Indiana Speech and Hearing Research Award amount: $500.00
2015	Student Scholarship, Aging and Speech Communication Conference 2015 Award amount: $150.00

Awards

2019	Judith Gierut Outstanding PhD Presentation Award Department of Speech and Hearing Sciences Indiana University

Lab Rotations

Indiana University Department of Speech and Hearing Sciences

Audiology Research Lab	Larry Humes, PI
Auditory Perception Lab	Jenny Lentz, PI
Speech Acoustics Lab	Karen Forrest, PI

Indiana University Department of Psychological and Brain Sciences

Vision Lab	Jason Gold, PI (attended lab meetings)

Research Tools- Programming, Hardware, and Software

Proficient in use of MATLAB, SPSS, SigmaPlot, and Excel
Experience with use of Tucker-Davis Technologies System III RP2.1 Real-Time-Processor, Praat, and R

Leadership/Service Experience- Professional

Spring 2019	Co-chair, Diane Kewley-Port lecture and mentorship event, Indiana University
Spring 2018	Peer reviewer, Journal of the Acoustical Society of America
Spring 2018	Review Committee, Undergraduate Research Grants, SPHS Department, Indiana University
2015-2017	PhD subcommittee, American Academy of Audiology
2015-2016	Co-chair, Diane Kewley-Port lecture and mentorship event, Indiana University
2005-2006	State Leadership Network member, American Academy of Audiology
2004-2005	President, California Academy of Audiology
2004	Contributing Editor, *Feedback* (Official Publication of the Academy of Dispensing Audiologists)
2004	Author, "Distance Education: The Doctors who Teach Future Doctors", *Feedback*, Vol. 15 No. 2, 2004
2003-2005	Conference Committee, California Academy of Audiology
2002-2004	Central California Representative, California Academy of Audiology

Professional Association Membership

American Auditory Society
American Academy of Audiology
Academy of Doctors of Audiology
Audiology Practice Standards Organization
American Speech-Language-Hearing Association