5-2015

Age, Cognition, and Listening Effort

Sapna Mehta
Graduate Center, City University of New York

AGE, COGNITION, AND LISTENING EFFORT

By

Sapna Mehta

A capstone research project submitted to the Graduate Faculty in Audiology in partial fulfillment of the requirements for the degree of Doctor of Audiology, The City University of New York

2015

This manuscript has been read and accepted for the Graduate Faculty in Audiology in satisfaction of the capstone project requirement for the degree of Au.D.

Date

Adrienne Rubinstein, Ph.D.
Faculty Advisor

Date

John Preece, Ph.D.
Executive Officer

THE CITY UNIVERSITY OF NEW YORK

Abstract

AGE, COGNITION, AND LISTENING EFFORT

By

Sapna Mehta

Faculty Advisor: Adrienne Rubinstein, Ph.D.

Second Advisor: Arlene Neuman, Ph.D., New York University Langone Medical Center

Older adults are reported to have more difficulty understanding degraded speech than younger adults. This may be due to greater recruitment of cognitive resources in adverse listening conditions. The purpose of this study was to examine the relationship between listening effort as measured by a dual task paradigm and cognitive abilities, specifically, working memory and selective attention, in normal-hearing older and younger adults in various background noise conditions. Results revealed that speech recognition scores were poorer for older adults, and speech recognition scores declined with decreasing signal-to-noise ratio. Stroop test scores suggested better selective attention ability in the younger participants, with no significant correlation to listening effort. No other significant results were found. Suggestions are made for future studies to continue investigating the effects of age and cognitive ability on listening effort.

Acknowledgements

I would like to acknowledge my advisors Dr. Adrienne Rubinstein and Dr. Arlene Neuman, who gave me the opportunity to pursue this line of study and spent time in guiding me throughout this project. I have learned a great deal from them and am grateful for their instruction and insights. I would also like to give my thanks to the Department of Otolaryngology at the New York University Langone Medical Center for their contributions to and guidance in this project.

I also extend my appreciation to the professors, staff members, and my fellow students in the CUNY Audiology program, who provided support to me throughout my time in the program and from whom I also learned a great deal.

Finally, I give my gratitude to my family and friends, who have always given me love and support, and without whom I would have not been able to complete this project.

Table of Contents

List of Tables

List of Graphs

Introduction

Older listeners often report the perception of increased difficulty in understanding speech in background noise. While the effects of age on speech recognition ability in adverse conditions have been well-documented in the literature, it is not yet clear why these effects occur, although there is considerable evidence that age-related changes in hearing sensitivity account for a significant portion of variation in performance. Nevertheless, even when hearing sensitivity is accounted for, differences in speech understanding related to other individual factors, including age, persist. While there is continued interested in examining these differences in speech understanding ability, there has also recently been increasing interest in studying the effects of various individual factors on cognitive effort expended during difficult listening tasks.

The information degradation theory and effortful theory of listening suggest that humans have limited cognitive capacity for processing incoming stimuli, and when recruitment of additional cognitive resources is required for the processing of degraded inputs, performance on parallel processing or later, higher-order processing, is affected (Pichora-Fuller, 2003; Pichora-Fuller & Singh, 2006; Tun, McCoy, & Wingfield, 2009). It is expected that when exogenous or endogenous factors lead to the degradation of auditory signals, individuals will variously require increased utilization of cognitive resources to process and interpret the stimuli, and will expend greater effort in understanding speech in these instances. Some of the factors that may contribute to degraded input include hearing loss and the loss of spectral and temporal discrimination abilities. In addition, available cognitive resources may be affected by factors such as age and individual variability in cognitive ability. Furthermore, aural rehabilitation efforts that address one or more of these potential contributing factors may or may not alleviate listening effort, regardless of whether they improve speech understanding performance. It is therefore important to understand how factors other than audiometric thresholds interact to affect processing of degraded auditory stimuli, and how such processing may differentially place demands on cognitive resources of individuals. The following review will focus on the effects of age and age-related factors, including cognitive abilities, on speech understanding and effortful listening.

I. Age and Speech Perception

Aging has been shown to affect the peripheral and central processing of auditory information in ways that may impact speech understanding. While research has supported that hearing thresholds account for a significant amount of the variability in speech understanding between older and younger listeners, other factors, such as temporal processing, spectral processing, and cognitive abilities, also may have an effect on performance in speech recognition tasks.

Age-related changes in hearing sensitivity and speech recognition

Studies that have examined the effects of aging on speech recognition abilities have found peripheral hearing loss to be a principle factor in performance in quiet and in noise (Humes, 1996), and in fact, there have been discrepant findings about the effect of age on performance when hearing loss has been taken into account. For example, Souza and Turner (1994) compared younger listeners with normal hearing, younger listeners with sensorineural hearing loss, and older listeners with sensorineural hearing loss, on a word recognition performance in various background noise conditions; the effect of pure-tone thresholds above 4000 Hz were minimized by utilizing a high-pass masking noise in all conditions. Souza and Turner (1994) found that hearing loss demonstrated an effect on performance, while age was not found to have any significant effect. Similarly, Takahashi and Bacon (1992) examined differences in performance on a speech recognition task between younger participants and older participants using low-probability sentences from the Speech in Noise Test with modulated and un-modulated background noise. Takahashi and Bacon (1992) found that there was a significant difference in pure-tone thresholds between the younger and older groups, and that this difference in hearing sensitivity accounted for most of the variation in performance between the groups. A study by Abel, Sass-Kortsak, and Naugler (2000) also found an effect of age on performance in noise. Their study showed that younger participants with normal hearing performed slightly better than older participants with normal hearing on a consonant discrimination task in noise. However, Abel et al., (2000) found that variation in performance on the discrimination task for final consonants was negatively correlated to high-frequency pure-tone thresholds as well as frequency selectivity (measured by critical masker levels of a 3150 Hz 1/3 octave band noise for a 4000 Hz probe tone) only for the older group, and the authors suggested that any age-related difference in performance on the discrimination task was primarily attributed to high-frequency thresholds and frequency selectivity ability (Abel et al., 2000).

Age-related changes to the auditory system other than peripheral sensitivity

Despite findings that hearing loss accounts for a great amount of the variation seen in speech recognition performance between younger and older listeners, other studies have found evidence that other age-related changes may occur that also contribute to the decline in speech recognition performance in many older adults (Pichora-Fuller & Souza, 2003). For example, Dubno, Dirks, and Morgan (1984) examined the relationship between age, hearing loss, and speech recognition ability in quiet and in noise, and found that in both, quiet and noise, there were main effects of hearing loss on performance. They also found that for speech recognition tasks in quiet, the articulation index and the predicted audibility of speech stimuli at given presentation levels could predict performance across age groups (Dubno et al., 1984). However, the results of the authors' study (Dubno et al., 1984) also revealed that predicted audibility of stimuli and the articulation index could not predict performance in noise for older participants, and that while there was no significant effect of age on performance in quiet, age did have an effect on performance in noise. Therefore the results of the authors' study suggest that factors related to age other than hearing sensitivity may contribute to a decline in speech recognition performance in noise.

A study by Smith, Pichora-Fuller, Wilson, and MacDonald (2012) suggest that such additional age-related factors may be temporal and spectral processing abilities. The authors found that when younger, normal-hearing participants performed a word recognition task with speech stimuli that were altered by adding spectral and temporal distortions that were hypothesized to mimic the age-related effects of neural dys-syncrony and spectral smearing, their performance was similar to the performance of older normal-hearing listeners who performed a similar word recognition task with speech stimuli that were not artificially distorted (Smith et al., 2012). Smith et al. (2012) also found that the older normal-hearing participants' performance with spectrally and temporally altered speech stimuli was similar to the performance of older participants with hearing loss with unaltered speech stimuli. Their results suggest that while peripheral hearing loss does account for some spectral and temporal distortions to speech inputs, other age-related changes may also contribute to the degradation of speech as it passes through the auditory system. A study by Gordon-Salant and Fitzgibbons (1993) also found that for younger and older listeners with normal hearing and with hearing loss, age and hearing loss both independently had an effect on performance on a speech recognition task of low-predictability sentences in noise with the following distortions: time compression, reverberation, and interruption. Gordon-Salant and Fitzgibbons (1993) additionally found that age and gap duration discrimination ability were related to performance on the speech recognition task in the reverberant condition, consistent with the idea that temporal processing abilities that affect speech recognition may be impacted by age.

The effects of age and hearing loss on temporal and spectral processing abilities were also examined by Fitzgibbons and Gordon-Salant (2004). The authors measured difference limens for tonal inter-onset intervals for trains of tone bursts in conditions of spectral complexity (varying tone-burst frequencies), temporal complexity (varying duration of inter-onset intervals), and in spectrally and temporally complex conditions for groups of younger and older listeners with normal hearing and hearing loss (Fitzgibbons & Gordon-Salant, 2004). Fitzgibbons and Gordon-Salant (2004) found no effect of hearing loss on difference limens, but they did find that older adults had larger difference limens compared to younger adults, and that this difference was greater for the temporally complex condition, and the combined spectrally and temporally complex condition, suggesting that changes in duration of intervals between sequenced stimuli may be harder to detect with age-related changes other than hearing loss.

Age-related differences in the processing of time-varying stimuli were also examined by Tremblay, Piskosz, and Souza (2003); the authors recorded P1, N1, and P2 cortical auditory evoked potentials and behavioral discrimination performance in younger normal-hearing listeners, older normal-hearing listeners, and older listeners with hearing loss in response to synthetic speech stimuli (consonant-vowel syllables) with varying voice-onset times. Tremblay et al. (2003) found that older adults performed worse on the discrimination task than younger adults, and that their N1 and P2 latencies were prolonged compared to the younger group, regardless of hearing thresholds. A study by Tremblay, Billings, and Rohila (2004) examined the effect of presentation rate of tonal and speech syllable stimuli trains on cortical N1 and P2 auditory evoked potentials in younger and older normal-hearing listeners, and found that fast presentation rates resulted in prolonged latencies of the evoked potentials in older adults for both stimuli types, while medium presentation rates resulted in prolonged latencies for older adults for speech stimuli; they found no effect of age when stimuli were presented at a slow rate, or for a medium rate of presentation of tonal stimuli. Anderson, Parbery-Clark, White-Schwoch, and Kraus (2012) also found an effect of age on neural representations of speech stimuli, recorded as speech-evoked auditory brainstem potentials; in particular, the authors found that older adults with normal hearing had delayed latencies, decreased phase-locking, and smaller magnitude responses compared to younger adults with normal hearing. These studies support the hypothesis that changes in temporal processing and changes in the neural representation of timing cues for complex stimuli are, in part, related to age.

Relationship between cognitive abilities and speech understanding

In addition to hearing sensitivity, temporal processing, and spectral processing abilities that may change with age, age-related cognitive changes have also been thought to be related to difficulties in speech understanding that many older adults report. Studies that have examined both the effects of hearing loss and cognitive ability on speech recognition performance have generally found that hearing loss primarily accounts for differences in speech recognition performance; nevertheless, studies have also shown that some cognitive abilities, such as working memory span, processing speed, and selective attention do explain some of the performance variation (e.g., Akeroyd, 2008; Humes, Lee, & Coughlin, 2006; van Rooij & Plomp, 1990).

Working memory is used for the temporary storage, manipulation, and integration of information, and is necessary for language processing (Baddeley, 2003). It has also been suggested that processing of language through working memory is relatively quick unless there is a mismatch between an incoming signal and a stored phonological representation of the symbol, as can occur with distorted speech, in which case working memory requires the use of additional cognitive resources for more explicit processing (Rudner, Foo, Rönnberg, & Lunner, 2009). Processing speed refers to the speed at which a cognitive activity is carried out, as well as the speed with which multiple simultaneous cognitive processes can be carried out; it is suggested that slowing of processing speed may impact many aspects of cognition and performance because slower processing of information may lead to the recruitment of additional cognitive resources in order to complete a task, and information that is processed early on may no longer be available for integration and use during later stages of processing (Desjardins, 2011; Salthouse, 1996). Selective attention refers to the ability to attend to target stimuli while inhibiting response to distracting stimuli, such as when an individual is listening to speech in background noise (Pichora-Fuller & Singh, 2006). It has been suggested that older adults demonstrate less ability in inhibiting responses and performing selective attention tasks compared to younger adults, and that this lack of selective attention may lead to increasing load on working memory, since a greater range of stimuli are processed and given consideration (Hasher, Stoltzfus, Zachs, & Rypma, 1991). The below studies have variously found relationships between these cognitive abilities and speech recognition performance.

Schvartz, Chatterjee, and Gordon-Salant (2008) examined the effects of age and cognitive abilities on difficult phoneme perception tasks. Normal-hearing younger, middle-age, and older adults were asked to repeat vocoded and spectrally shifted phonemes that simulated cochlear implant processing. The authors found that the younger group performed significantly better on phoneme recognition compared to the older groups; they also found that measures of verbal memory were predictors of performance on the difficult speech task regardless of age, and that within the middle-age and older groups, verbal memory and processing speed were both predictors of performance (Schvartz et al., 2008). Tun and Wingfield (1999) found that while hearing sensitivity did account for variation in performance for speech in noise recognition tasks, processing speed, as measured by the Wechsler's Adult Intelligence Scale digit symbol substitution test and by the recall speed during the speech recognition task, also accounted for the variation in performance, and that both hearing loss and processing speed together accounted for any differences in speech recognition ability between younger and older participants. Finally, while Harris, Eckert, Ahlstrom, and Dubno (2010) did not look at speech recognition performance directly, they did find that attention and processing speed, as measured by the Purdue Pegboard and Connections tests, predicted random gap detection abilities for both younger and older adults, and that older adults performed significantly poorer on the random gap detection task compared to younger adults, as well as demonstrated slower processing speeds on the Purdue Pegboard and Connections tests.

The relationship between age and cognitive processes in listening tasks is also supported by studies that utilized functional magnetic resonance imaging (fMRI) and anatomical magnetic resonance imaging (MRI) to examine cortical activation and gray matter volume in younger and older adults (Wong, Ettlinger, Sheppard, Gunasekera, & Dhar, 2010; Wong, Uppunda, Parrish, &

Dhar, 2008). Wong et al. (2010; 2008) found that while older adults were generally able to maintain performance on a speech recognition in noise task similar to that of younger adults, their performance was significantly worse in poor signal-to-noise ratio (SNR) conditions (0 dB SNR and -5 dB SNR); they also found that in these conditions, the performance of older adults was significantly correlated to the volume of gray matter in the left pars triangularis and the thickness of left superior frontal gyrus, as well as to activation of the pre-frontal cortex, the superior temporal gyrus, and the precuneus (Wong et al., 2010; Wong et al., 2008). These results suggest that older adults may need to recruit more cognitive resources to process difficult listening tasks, and that adults with better capacity for certain cognitive tasks may perform better on difficult listening tasks.

II. Listening Effort

The results of the above studies demonstrate some relationship between cognitive abilities and speech recognition understanding in difficult listening tasks (Akeroyd, 2008; Harris et al., 2010; Humes et al., 2006; Schvartz et al., 2008; Tun & Wingfield, 1999). As cognitive abilities may decline with age and are necessary for the processing of incoming stimuli in demanding listening conditions, it has been suggested that in order to perform difficult listening tasks, compensatory cognitive processes and resources may be recruited (Desjardins, 2011; Pichora-Fuller & Singh, 2006). The attention, cognitive resources, and effort involved in understanding speech are generally referred to as listening effort (Downs, 1982; Feuerstein, 1992; McGarrigle et al., 2014). While the studies mentioned above generally assessed performance on speech recognition tasks directly, speech recognition performance may not always reflect the degree to which cognitive resources are being recruited and effort is being expended on listening. Maintenance of good performance on speech recognition tasks by utilizing additional finite and limited cognitive resources may come at the expense of parallel processing or performance on simultaneous tasks. It is therefore necessary to be able to measure the expenditure of cognitive resources and mental effort in listening tasks without relying solely on speech understanding performance.

Assessing Listening Effort

Because listening effort may vary despite performance on a speech recognition task, assessing listening effort has not been as straightforward as measuring changes in speech recognition ability. However, various studies have been successful in measuring the effect of difficult listening tasks on physiologic and behavioral variables other than performance on the primary listening task. Listening effort may be assessed behaviorally, either through performance measures or through self-assessment scales. Self-assessment scales, such as the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) and the Dartmouth Primary Care Cooperative Project Scales (COOP; Wasson, Kairys, Nelson, Kalishman, & Baribeau, 1994) have been used in studies to compare variability in self-assessment of listening effort to variability in other physiologic and behavioral measures, as well as to determine the sensitivity of these instruments to changes in cognitive load (Bess, Dodd-Murphy, & Parker, 1998; Hicks & Tharpe, 2002; Mackersie & Cones, 2011). The Abbreviated Profile of Hearing Aid Benefit (APHAB; Cox & Alexander, 1995), the Speech, Spatial and Qualities of Hearing Scale (SSQ; Gatehouse & Noble, 2004), Borg's CR-10 scale of physical exertion (Borg, 1990), and other similar scales have also been used to assess self-perception of listening effort and ease of communication (e.g., Desjardins, 2011; Desjardins & Doherty, 2013; Getzmann, 2012; Hallgren, Larsby, Lyxell, & Arlinger, 2005; Hol, Kunst, Snik, Bosman, Mylanus, & Cremers, 2010a; Hol, Kunst, Snik, & Cremers, 2010b; Larsby, Hallgren, Lyxell, & Arlinger, 2005; Zekveld, Kramer, & Festen, 2010).

While self-assessment scales of ease of communication, listening effort, and exertion have been shown to be related to hearing loss, aural rehabilitation, and difficulty of listening conditions (e.g., Bess et al., 1998; Hallgren et al., 2005; Hol et al., 2010a; Hol et al., 2010b; Larsby et al., 2005), these scales have not consistently been shown to correlate to other objective measures of cognitive load and effort. For example, Hicks and Tharpe (2002) found that the COOP did not demonstrate significant differences between children with and without hearing loss, even while objective measures of exertion did. Similarly, Mackersie and Cones (2011) found only a modest correlation between ratings on the NASA-TLX and changes in a physiologic measure of exertion during listening tasks in adults with normal hearing, although this result may have been affected by the ease of the task used in the study and the ceiling performance of participants. Desjardins (2011) and Desjardins and Doherty (2013) also did not find a relationship between self-ratings of effort on a scale of 0 to 100 and listening effort as measured by decline in performance on a visual tracking task while listening to speech in different noise conditions. It would therefore appear that self-assessment tools may not be suited for investigating subtle changes in listening effort.

Physiologic measurements have been used as assessment for listening effort, and include measures of heart rate, skin temperature, electromyographic (EMG) responses, skin conductance, pupil dilation, salivary cortisol levels, and event-related cortical potentials. These measures rely on a baseline recording against which experimental measures can be compared within study participants. Skin conductance (O’Gorman & Lloyd, 1988) and pupil dilation (Zekveld, et al.,

2010) have been found to be the most sensitive of physiologic measures to increasing listening demand.

Event-related cortical potentials (ERPs) have received increasing attention recently for their potential in measuring listening effort and effects of cognitive load. Bertoli and Bodmer (2014) have identified the novelty-related P3 potential and the late positive potential (LPP) that follows it as possible markers of listening effort. Bertoli and Bodmer (2014) also found an effect of task difficulty on the amplitudes of the N1 and N2 potentials recorded in response to novel stimuli, which they suggest may reflect a variation in attention and processing of the novel stimuli. Another potential ERP marker of listening effort may be found in the phase synchronization of ERPs across sweeps in the latency range of the N1 response (Bernarding, Strauss, Hannemann, Seidlers, & Corona-Strauss, 2013). Bernarding et al. (2013) used a syllable discrimination task with two levels of difficulty, and measured the ERP recordings when participants identified a target syllable. The authors reported that for the middle-age participants in their study, the difficulty of the task resulted in significant differences in phase synchronization index of the N1 response. This difference was not seen for younger adults in their study, and this may be a result of the level of difficulty of the tasks (Bernarding et al., 2013). Based on the aforementioned studies, while event-related potentials may prove to be useful in identifying and quantifying listening effort, further research is needed in order to determine which task designs and stimuli types provide the most reliable and valid representations of listening effort.

Yet another way in which listening effort may be measured is through behavioral assessments. Dual task paradigms have commonly been employed in research to explore the effects of difficult listening tasks on mental effort. Dual task paradigms involve a primary listening task, on which participant performance is monitored, and a secondary task, in which either accuracy or reaction time are measured. The concept of using a dual task paradigm to measure listening effort stems from the theory of limited cognitive capacity, in which resources available for cognitive processing are finite, and when a difficult task demands the use of increased cognitive resources in order to maintain a certain level of performance on the task, the lack of resources available for another task will result in a decrease in performance on that second task (Pichora-Fuller, 2003; Ninio & Kahneman, 1974; Rabbitt, 1968; Wingfield, Tun & McCoy, 2005). A variety of secondary tasks have been used, including recall tasks and visual tasks. In an early study, Downs (1982) used a dual task paradigm with a primary speech recognition task, and a secondary task in which participants had to monitor and respond to a light. Downs (1982) examined the reaction time for the secondary task as a measure of listening effort in participants with hearing loss who tried to perform the speech recognition task with and without hearing aids. He found that when the speech recognition task was performed with hearing aids, reaction times on the secondary task were improved, and he suggested that this reflected improvement in listening effort (Downs, 1982). Many other studies have also utilized a dual task paradigm to measure listening effort in adults (e.g., Rakerd, Seitz, & Whearty, 1996; Sarampalis, Kalluri, Edwards, & Hafter, 2009; Tun et al., 2009), as well as in children (Howard, Munro, & Plack, 2010; McCreery & Stelmachowicz, 2013). As many of these studies found some change in the secondary task measures due to independent variables related to listening conditions, it appears that the dual task paradigm has good validity as a measure of listening effort.

Review of the findings of studies of listening effort

Studies in listening effort have focused mainly on the effects of hearing loss. Rakerd et al. (1996), Hicks and Tharpe (2002), Bertoli and Bodmer (2014), and Bernarding et al. (2013) found that hearing loss had an effect on listening effort, measured variously through performance on secondary tasks in dual task paradigms, or through cortical auditory ERPs (specifically amplitude of the LLP and phase synchronization of the N1). Tun et al. (2009) also used a dual task paradigm, in which older and younger participants with and without hearing loss were asked to perform a visual tracking task, and simultaneously recall words they had recently listened to and repeated. Tun et al. (2009) found an effect of age and hearing loss on listening effort as measured by decline in performance in the visual tracking task from the single to dual task conditions, attributing the effect mostly to older listeners with hearing loss. Interestingly, Desjardins (2011) and Desjardins and Doherty (2013) found, in a study that examined the effects of age, hearing loss, and cognitive ability on listening effort as measured with a dual task paradigm, that hearing loss did not have a measurable effect on listening effort in various noise conditions. However, the authors also found that older adults, regardless of hearing status, demonstrated greater listening effort when listening to speech in certain background noise conditions (two-talker babble and speech-shaped noise) compared to younger adults (Desjardins, 2011; Desjardins & Doherty, 2013).

Other studies examining listening effort have also found an effect of age. Bernarding et al. (2013), whose study, as mentioned above, found an effect of hearing loss on listening effort, also found an effect of age on listening effort. In a different study, Getzmann (2012) used a dual task paradigm in which older and younger normal-hearing participants had to perform a word target detection task while listening to dichotically presented stories; the participants also had to recall content of the attended story. Getzmann (2012) found that older listeners performed better than younger listeners at detecting targets, but that better target detection was correlated to lower performance on the secondary recall task. He also found that the P3b component of the auditory ERP was of smaller amplitude in older adults than younger adults, while the P3a component amplitude was greater, suggesting that older adults may be using greater attentional resources in the listening task (indicated by P3a) while demonstrating an overall decrease in later processing activity (Getzmann, 2012). Gosselin and Gagne (2011) also found that age affected listening effort. In a dual task paradigm in which the primary task was a speech in noise recognition task, and the secondary task was a pattern recognition task for tactile stimuli, Gosselin and Gagne (2011) found that older adults with normal hearing (through 3000 Hz) were less accurate and had longer response times for the secondary task compared to younger adults with normal hearing, indicating greater listening effort for older adults. They also found that when the noise level was reduced for the primary task, the decline in performance in the secondary task was also reduced, reflecting improvement in listening effort (Gosselin & Gagne, 2011).

Other studies have investigated the use of hearing assistive devices and strategies in alleviating listening effort. Sarampalis et al. (2009) found that the application of a noise reduction algorithm used in hearing aid processing to the presentation of speech in noise reduced listening effort in a poor signal-to-noise ratio condition. A study conducted by Downs (1982), mentioned above, found that the use of hearing aids for a speech recognition task improved performance on a secondary task, indicating improvement in listening effort. Hughes and Galvin (2013) examined the effect of cochlear implant use on listening effort in adolescents using a dual task paradigm. They found that when performance by adolescent cochlear implant users on a word recognition in noise task was matched to performance by peers with normal hearing by individually adapting SNR, listening effort, measured by change from baseline in response time to a secondary concurrent visual matching task, was also similar. The authors also found that when they compared bilateral cochlear implant use to unilateral implant use within the same participants, three out of the eight participants demonstrated improvement in listening effort with bilateral cochlear implant use. These results suggest that cochlear implant use may allow improvement in listening effort for adolescents with hearing loss, and that for some individuals, bilateral processor use improves listening effort compared to unilateral use (Hughes & Galvin, 2013).

III. Rationale and Objectives

As the above literature review has shown, audiometric thresholds alone do not predict all the variation in performance in understanding degraded speech signals, and other individual factors must be taken into account. Furthermore, performance on a difficult speech recognition task does not predict the cognitive resources and effort necessary to maintain a certain level of performance. While it has been shown that age-related differences in performance on speech recognition tasks in difficult listening conditions exist, few studies have examined the effects of age on listening effort. In addition, while studies have suggested a relationship between speech recognition ability and cognitive abilities such as working memory, selective attention, and processing speed, it is not yet clear to what degree cognitive ability affects speech understanding, and the effects of individual variation in cognitive ability on listening effort are not yet fully understood.

This study seeks to further explore the effects of age on listening effort during difficult speech recognition tasks, as well as examine the relationship between cognitive abilities, specifically working memory and selective attention, and listening effort. A dual task paradigm was used to measure listening effort in younger and older normal-hearing adults. The primary task was a speech recognition task that utilized vocoded spectrally degraded sentences in quiet, and two levels of background noise; the secondary task consisted of a digit recall task. Working memory ability was measured with forward and backward digit span tests based on the digit span test of the Wechsler's Adult Intelligence Scale (Wechsler, 2008), and selective memory was measured with a Stroop test based on Golden's version of the test (Chafetz & Matthews, 2004; Golden, 1978; Lansbergen, Kenemans, & van Engeland, 2007).

The hypotheses tested in this study were as follows:

1) There is a significant effect of age on listening effort.

2) There is a significant effect of listening condition (quiet and different levels of background noise) on listening effort.

3) There is a significant interaction of age and background noise condition on listening effort.

4) There is a correlation between working memory and listening effort, and selective attention and listening effort.

Methods

Approval for this study was obtained from the City University of New York Brooklyn College Internal Review Board (IRB). The study was conducted at the Graduate Center Audiology Lab and Brooklyn College Diana Rogovin Davidow Hearing Center of the City University of New York. Funding for this study was provided by the City University of New York's Department of Audiology.

I. Participants

Participants in the age groups of 18-40 years and 55 years and older were recruited by posting flyers at the Graduate Center and Brooklyn College of the City University of New York, and through email or face-to-face recruitment of individuals known to the investigators. The inclusion criteria were having hearing within normal limits at octave intervals from 250 to 4000 Hz in both ears, passing a cognitive screening, and having been a monolingual English speaker until the age of 5 years. Ultimately, 10 individuals ages 18-40 years and 6 individuals 55 years and older participated in the study. The ages of the participants in the younger group ranged from 18 to 33 years with a mean age of 24.9 years (*SD* = 3.81 years). The ages of the participants in the older group ranged from 58 to 64 years with a mean age of 60.7 years (*SD* = 2.16 years). All participants were provided compensation for their time and transportation to the testing site.

II. Screening Measures

Screening measures were conducted upon the initial meeting with the potential participant at one of two test sites: the CUNY Graduate Center Audiology Lab or the CUNY Brooklyn College Diana Rogovin Davidow Hearing Center. Cognitive screening was conducted using verbal and written administration of the Mini Mental State Exam (MMSE) (Folstein, Folstein, & McHugh, 1975). A score of 27 or higher out of 30 on the MMSE was required to pass the cognitive screening. This cut-off value was selected as per findings from O’Bryant et al. (2008) that a cut-off score of 27 provided greater sensitivity in detecting dementia in a college-educated sample of elderly subjects compared to a score of 24.

Audiologic screening was conducted in one of the sites’ audiologic double-booth suites utilizing a Grason-Stradler GSI-61 audiometer and Telephonics TDH-49 supra-aural headphones. Otoscopy was performed prior to the screening and individuals’ ears were checked for collapsing canals. Hearing thresholds were measured in each ear for the octave and inter-octave frequencies of 250, 500, 1000, 2000, 3000, and 4000 Hz. A hearing threshold was defined as the lowest presentation level at which an individual detected and responded to tonal stimuli in two out of four presentations with no visual cues. Only individuals whose thresholds for tested frequencies were less than or equal to 25 dB HL were included in the study. Participants who did not meet the inclusion criteria based on the cognitive or hearing screenings were referred to the appropriate health care professional for further follow up.

III. Cognitive Measures

Working memory was assessed in participants by means of forwards and backwards digit span tests, based on the digit span test of the Wechsler’s Adult Intelligence Scale (Wechsler, 2008). For both digit span tests, strings of digits were presented on a computer screen in large bold black typeface for two seconds, the screen then went blank, and a participant was asked to recall the digit string either in order for the forwards condition, or in reverse order for the backwards condition. For the forwards digit span test, the strings ranged from two to nine digits, and for the backwards digit span test, the strings ranged from two to eight digits. There were two presentations of each length of digit string, and strings consecutively got longer. Items continued to be presented until the participant incorrectly recalled two items of a certain string length or the test list was completed. Each digit span test was scored by giving one point for each string recalled correctly, and the longest forwards and backwards digit span was obtained by counting the number of digits in the longest string each participant could recall correctly in each of the tests.

Selective attention and response inhibition were assessed with a Stroop test based on Golden's version of the test (Chafetz & Matthews, 2004; Golden, 1978; Lansbergen, et al., 2007). Stimuli were created by printing 100 words in bold typeface on three sheets of 8x11.5 inch paper each; one sheet contained color names printed in black ink, another sheet contained color names printed in the same color as the word described, and the final sheet contained color names printed in colors that differed from what was described. Five different colors were use (red, yellow, blue, brown, black), and each participant was shown the colors on a computer screen prior to the test to ensure ability in identifying and differentiating the colors. Participants were given each paper with stimuli in the following order: words printed in black (W condition), words printed in the same color the word described (C condition), and finally, words printed in colors different than described (CW condition). The participants were asked to read aloud the words in the W condition, to identify the color of the ink in the C condition, and to identify the color of the ink in the CW (incongruent) condition, while ignoring the word. Participants were scored on how many items they correctly read or identified in each condition in the allotted time. As described by Golden (1978), a predicted CW score was calculated from each participant's C and W scores (CW predicted = (W*C)/(W+C)). This score was then compared to the participant’s actual CW score, and an interference score was calculated by subtracting the predicted CW score from the measured CW scores. A low or negative interference score is consistent with a poorer CW score than predicted. This reflects difficulty in attending to colors while suppressing attention to reading words, and may suggest poor selective attention ability (Chafetz & Matthews, 2004; Golden, 1978; Killian, 1985; Lansbergen et al., 2007).

IV. Listening Effort Measures

Listening effort was measured using a dual task paradigm. Participants were asked to recall seven-digit strings presented on a screen while listening to and repeating sentences in difficult listening conditions. The ability to correctly repeat the speech stimuli as well as recall the digits was measured in three different noise conditions: quiet, +16 dB SNR, and +12 dB SNR.

Stimuli

Vocoded speech materials were used as speech stimuli in order to mitigate ceiling effects in speech recognition performance in the normal-hearing participants. The speech stimuli used in the dual task paradigm were created at the New York University (NYU) Langone Medical Center’s Department of Otolaryngology by digitally processing AzBio sentences in quiet and in multi-talker babble noise (with +16 dB SNR and +12 dB SNR) (Spahr and Dorman, 2004, Spahr et al., 2012) as described by Kaiser and Svirsky (2000). Digital .wav files of the sentences were sent through a bank of eight analysis filters, and envelope detection by half wave rectification and second order low-pass filtering at 160 Hz was performed on their outputs (Kaiser & Svirsky, 2000). Each one of the eight envelopes was used to modulate a noise band, which was obtained by passing white noise through synthesis filters, each of which was identical to the corresponding analysis filter (Kaiser & Svirsky, 2000). Auditory stimuli for the listening effort tasks were presented to participants through a MATLAB program (MATLAB, 2013) on a laptop using an external sound driver and TDH-49 headphones. Stimuli were confirmed to achieve an average level of 70 dBA on the test equipment with a sound-level meter using an A-weighted scale.

Seven-digit strings were randomly generated using a MATLAB program provided by the NYU Langone Medical Center's Department of Otolaryngology. Digits were set to present on screen in bold, large typeface for 2 seconds in between sentence presentations. Each presentation of a digit string was preceded by the words "Watch for the numbers!" on the screen and verbal warning of "Ready?" The program would then prompt recall of the digits aloud with the words "Repeat the numbers" on screen and an audible tonal pulse.

Listening effort procedure

The sentences and digits were presented in each task and noise condition in the same format. A list of AzBio sentences was selected, and a seven-digit string was presented prior to the auditory presentation of the first sentence. The program then asked that the digit string be recalled after two sentences; each sentence was followed by a pause during which time participants could repeat the sentence if they were asked to do so. Another sentence was then presented followed by a pause, and then a new string of digits was presented. Ultimately, there were 20 sentences in each list, and the digit strings were presented seven times in each list. The presentation order of the stimuli and directions and can be represented as follows (D = digit presentation, S = sentence presentation, P = pause after sentence, R = program asks that digits are recalled):

D S P S P R S D S P S P R S D S P S P R S D S P S P R S D S P S P R S D S P S P R S D S P S P R

There were a total of three noise conditions (quiet, +16 dB SNR, and +12 dB SNR) and three task conditions within each noise condition (sentence recognition single task, digit recall single task, sentence recognition and digit recall dual task). Participants were given a full practice list prior to testing in each noise condition in order to familiarize them with the stimuli and the sentence recognition task. Participants were also given practice at the beginning of the testing period in performing the digit recall single task, and the sentence recognition and digit recall dual task. The order of presentation of noise and task for each participant was counterbalanced using a Latin square method within the older group and within the younger group. Participants were given instructions before the start of each condition’s test. For the sentence recognition single task, they were asked to ignore the digits on screen and focus on performing to their best ability on the sentence recognition task. For the digit recall task single task, they were asked to ignore the sentences they heard, and were asked to recall the string of digits in order to the best of their ability when prompted by the program to do so. For the sentence recognition and digit recall dual task, participants were asked to focus on repeating the sentences accurately to the best of their ability, and to try to recall as many digits as possible when prompted, but to consider that a secondary task.

Scoring was conducted by the experimenter, who was seated in the same room as the participant. Sentences were scored by words correct on a score sheet, and the score for each sentence was entered digitally into the MATLAB program during testing. Recalled digits were written down on a score sheet, and the recalled digits were also entered into the MATLAB program during testing. Recalled digits were scored as correct when the correct number was recalled in the same position in which it appeared in the original string. A percent correct was calculated in MATLAB for sentence recognition and for digit recall for test condition.

V. Procedures

Following cognitive and hearing screenings, participants continuing in the study were presented with the cognitive tests (the Stroop test and the digit span tests) and the listening effort test. The order of test presentation (listening effort versus cognitive tests) was counterbalanced between participants within the older group and the younger group. Test presentation of the cognitive tests was also counterbalanced within the groups. Breaks were provided throughout testing at the participant’s request; a break was also suggested to the participant by the experimenter at least once during the listening effort task. Participants were given the option to discontinue testing and return another day if preferred. Overall, the duration of testing was between one to two hours.

The experimenter was seated with the participant throughout testing (with the exception of the hearing screening). The listening effort speech recognition tasks were scored by the experimenter during the test, and while the participant was able to see the number of words correct entered into the MATLAB program for each sentence, they were not shown their final sentence recognition score or digit recall scores for any of the test lists. The Stroop test raw C, W, and CW scores were counted at the completion of testing. However, the predicted and interference scores were calculated at a later time. The digit span tests scores were also calculated at a later time.

VI. Data Analyses

Raw speech recognition score percentages and digit recall score percentages obtained from the listening effort task were transformed for each subject to rationalized arcsine units in order to normalize the data and facilitate comparison of scores (Sherbecoe & Studebacker, 2004; Studebaker, 1985). The raw scores were transformed to radians as described by Studebaker (1985) and Sherbecoe and Studebaker (2004) for small test sets (less than 150 items), and then transformed to rationalized arcsine units. These transformed data were then used for further data analyses.

In order to assess whether age, noise condition, and task difficulty had an effect on listening effort as measured by performance on the digit recall task in the single and dual task conditions, a repeated measures analysis of variance (RM ANOVA) was performed. The within-subject factors were noise conditions (quiet, +16 dB SNR, and +12 dB SNR), and task (digit recall single task score, and digit recall score in the speech recognition and digit recall dual task); the between-subjects factor was age group (younger and older). A RM ANOVA was also performed to assess the effects of age group, noise conditions, and task on speech recognition scores obtained during the listening effort task. Again, the within-subject factors were noise conditions (quiet, +16 dB SNR, and +12 dB SNR), and task (speech recognition single task score, and speech recognition score in the speech recognition and digit recall dual task); the between-subjects factor was age group (younger and older).

In order to assess whether cognitive ability, as measured by the Stroop test and digit span test, correlated with listening effort expended in demanding listening conditions, Pearson's product-moment correlations were calculated. A measure of listening effort, “mean digit recall difference score,” was calculated as described in the following results section, and was used as a co-variable in the Pearson product-moment correlation calculations with the Stroop test interference score, and the forwards and backwards digit span scores and longest digit spans.

Results

Note: an alpha of 0.05 was used for all inferential statistics unless specified otherwise.

I. Audiometric data

Pure-tone thresholds were obtained for all participants for octave frequencies 250-4000 Hz and at 3000 Hz for the left and right ears. Three-frequency (500, 1000, and 2000 Hz) and four-frequency (500, 1000, 2000, and 4000 Hz) pure-tone averages (PTAs) were calculated for all participants. Two-tailed t-tests using Bonferroni adjusted alpha levels of 0.003125 for each test revealed significant differences between the two age groups in pure-tone thresholds (in dB HL) for the right ear at 250 Hz (younger $M = 4.50$, $SD = 4.97$; older $M = 12.50$, $SD = 2.74$), $t(14) = -3.59$, $p = 0.003$; 500 Hz (younger $M = 3.50$, $SD = 4.12$; older $M = 12.50$, $SD = 4.18$), $t(14) = -4.21$, $p < 0.001$; 3000 Hz (younger $M = 0.50$, $SD = 5.99$; older $M = 11.67$, $SD = 6.06$), $t(14) = -3.60$, $p = 0.003$; and 4000 Hz (younger $M = 3.00$, $SD = 6.75$; older $M = 14.17$, $SD = 2.04$), $t(14) = -3.90$, $p = 0.002$. Significant differences were also found for four-frequency PTA in the right ear between the younger ($M = 2.88$, $SD = 3.78$) and older ($M = 10.42$, $SD = 2.58$) groups, $t(14) = -4.30$, $p < 0.001$. The data demonstrated that the above-mentioned thresholds and PTA for the older group were significantly higher (in dB HL) compared to the younger group. Nevertheless, pure-tone thresholds for all participants were 25 dB HL or better, consistent with hearing within normal limits.

II. Speech Recognition Scores

The mean speech recognition scores (SRS) and standard deviations in rationalized arcsine units (RAU) for the younger and older groups across each noise condition and each task condition are shown in Table 1.

Table 1. Speech recognition scores across groups, background noise conditions, and tasks

	Younger group						Older group					
Noise	Quiet		+16 dB SNR		+12 dB SNR		Quiet		+16 dB SNR		+12 dB SNR	
Task	S	D	S	D	S	D	S	D	S	D	S	D
Mean SRS	100.53	99.40	86.75	84.72	66.52	69.06	93.87	90.79	69.89	74.78	55.49	56.74
SD	5.80	5.97	10.37	7.60	8.32	8.75	4.13	5.65	3.51	3.16	6.20	5.11

Note: S = Single task, D = Dual task, SD = standard deviation; scores are in rationalized arcsine units (RAU)

A repeated measures analysis of variance (RM MANOVA) of the effect of noise condition (quiet, +16 dB SNR, +12 dB SNR), task condition (speech recognition single task, and speech recognition and digit recall dual task), and group (older and younger) yielded a main effect of group and noise on speech recognition scores, $F(1, 14) = 34.47$, $p < 0.001$ and $F(2, 28) = 365.94$, $p < 0.001$, respectively, demonstrating that the older group had poorer speech recognition scores than the younger group. A Tukey post-hoc analysis revealed significant differences in scores between the quiet and +16 dB SNR conditions, the quiet and +12 dB SNR conditions, and the +16 dB SNR and +12 dB SNR conditions ($p < 0.001$ for all comparisons). These results are consistent with the expectation that speech recognition performance would decline with more difficult listening conditions (i.e., lower SNRs). There were no significant interactions between task and group, $F(1, 14) = 0.18$, $p = 0.682$, task and noise, $F(2, 28) = 0.60$, $p = 0.556$, group and noise, $F(2, 28) = 2.74$, $p = 0.082$, or task, noise, and group, $F(2, 28) = 0.77$, $p = 0.475$. There was also no significant main effect of task on speech recognition scores, $F(1, 14) = 0.08$, $p = 0.783$. These results are consistent with maintenance of performance on the speech recognition task in both the single task and dual task conditions by the participants, on average.

Despite results showing that average performance on speech recognition tasks in the single and dual task conditions was not significantly different in any of the noise conditions, examination of individual data reveals that certain individuals did not maintain performance on the speech recognition task across task conditions. In particular, participants Y01, Y03, Y05, Y09, and Y10 in the younger group, and O03 in the older group were noted to have speech recognition scores in some noise conditions that varied significantly between single and dual task conditions, using criteria based on the 95% critical intervals obtained from the binomial distribution model put forth by Spahr et al. (2012). See Figures 1-6 below.

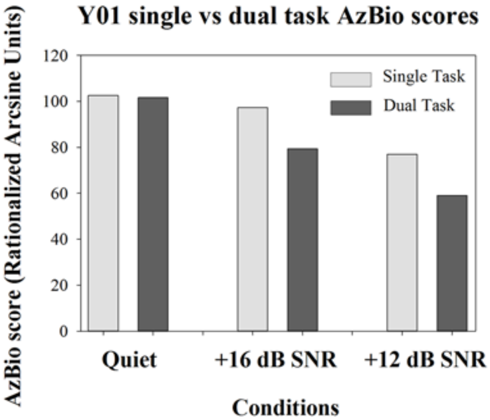

Figure 1. Y01 speech recognition scores

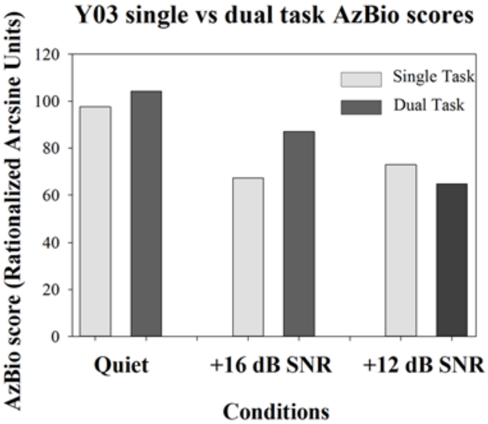

Figure 2. Y03 speech recognition scores

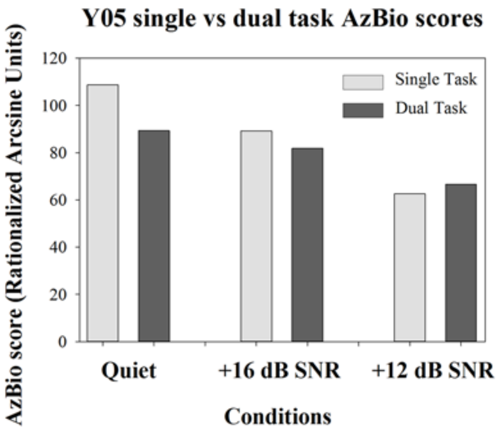

Figure 3. Y05 speech recognition scores

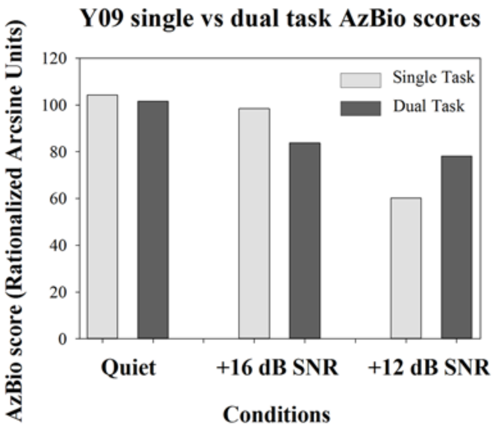

Figure 4. Y09 speech recognition scores

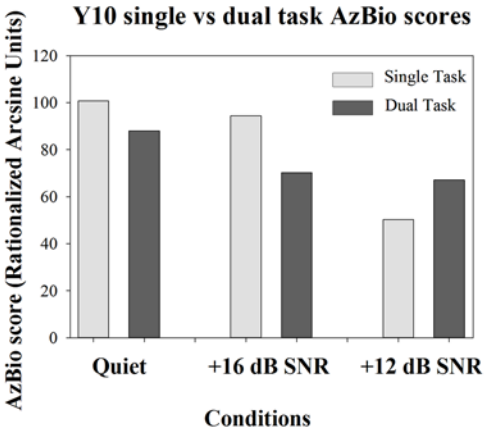

Figure 5. Y10 speech recognition scores

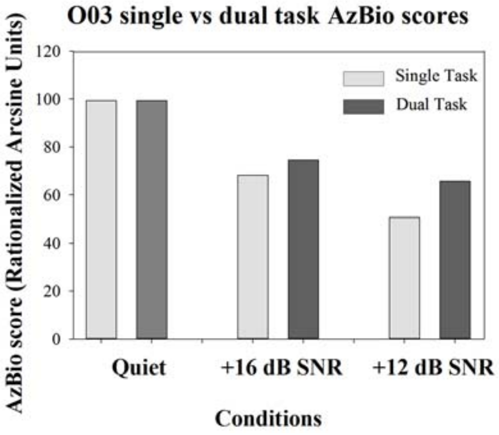

Figure 6. O03 speech recognition scores

III. Digit Recall Scores

The mean digit recall scores and standard deviations for the younger and older groups across each noise condition and each task are shown in Table 2.

Table 2. Digit recall scores across groups, noise conditions, and tasks

	Younger group						Older group					
Noise	Quiet		+16 dB SNR		+12 dB SNR		Quiet		+16 dB SNR		+12 dB SNR	
Task	S	D	S	D	S	D	S	D	S	D	S	D
Mean Score	91.30	71.92	94.28	63.12	99.47	59.70	92.79	50.00	94.09	53.65	96.83	47.43
SD	18.62	24.70	14.74	21.84	14.80	20.29	21.96	16.84	15.31	12.78	18.49	16.24

Note: S = Single task, D = Dual task, SD = standard deviation; scores are in rationalized arcsine units (RAU)

A RM MANOVA of the effects of noise condition (quiet, +16 dB SNR, +12 dB SNR), task condition (digit recall single task, and speech recognition and digit recall dual task), and group (older and younger) yielded a main effect of task, $F(1, 14) = 107.43$, $p < 0.001$, indicating that digit recall performance declined in the dual task condition compared to the single task condition. The main effects of group and noise condition were non-significant, $F(1, 14) = 0.98$, $p = 0.340$, and $F(2, 28) = 0.02$, $p = 0.977$, respectively. These results do not support the hypotheses that age or noise condition have an independent effect on listening effort, and therefore the null hypotheses cannot be rejected. There were also no significant interactions of noise and group, $F(2, 28) = 0.39$, $p = 0.681$, noise and task, $F(2, 28) = 3.01$, $p = 0.066$, task and group, $F(1, 14) = 3.87$, $p = 0.069$, or noise, task, and group, $F(2, 28) = 1.04$, $p = 0.367$, indicating that the hypothesis that age, noise condition, and task difficulty interact together to effect listening effort cannot be accepted, and the null hypothesis cannot be rejected. Post-hoc power analyses of age-related between-factor effects on digit recall scores for a sample size of 16 and medium and large effect sizes (*f* = 0.25 and *f* = 0.40, respectively, as described by Cohen, 1988) were conducted using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), and revealed low statistical power of 0.15 and 0.32 for detecting effects of medium and large sizes, respectively. Similarly, post-hoc power analysis for age and task interactions revealed low statistical power (0.08 and 0.13 for medium and large effect sizes, respectively), and analysis for age, task, and noise condition interactions revealed low statistical power (0.09 and 0.18 for medium and large effect sizes, respectively). Thus, the small sample size limited the statistical power of this study to detect medium or large effects of age, and interactions of age, noise condition, and task on the digit recall scores.

A one-way repeated-measures ANOVA of the effect of noise condition on digit recall scores in the single task condition did not yield a significant main effect, $F(2, 30) = 1.56$, $p = 0.226$, indicating that on average, individuals maintained baseline digit recall ability in each noise condition. Therefore, a mean digit recall score for the single task condition was calculated by averaging individual scores across all three noise conditions. This resultant mean digit recall score was used as the baseline score against which digit recall scores in each dual task condition were compared. A measure of listening effort was obtained by calculating the difference between an individual's digit recall score in a dual task condition and the baseline mean digit recall score for the single task conditions; this value will heretofore be referred to as "mean digit recall difference score." The mean digit recall difference score was used as a co-variable in correlation analyses with measured cognitive ability scores, described below.

IV. Cognitive Measures

The Stroop test's interference score was calculated as a measure of selective attention ability. The interference score was defined as follows (Golden, 1978; Killian, 1985): Interference score = CW-CW', where C = number of colors read in allotted time, W = number of words read in allotted time, CW = number of colors read in the allotted time in the incongruent condition, and CW' is the predicted number of colors read in the incongruent condition (defined as (W*C)/(W+C)). A lower interference score thus indicates that an individual read fewer colors in the incongruent condition than predicted by his or her performance on the simple word-reading and color-reading conditions. A two-tailed independent t-test revealed that average interference scores of the younger group (*M* = 8.12, *SD* = 6.64) were higher than those of the older group (*M* = -0.25, *SD* = 3.76); *t(*14) = 2.81, *p* = 0.014.

Results from a digit span test were obtained as a measure of working memory span in participants. Two-tailed independent t-tests showed no significant differences between groups for forward digit span (younger *M* = 7.50, *SD* = 0.97; older *M* = 8.33, *SD* = 1.51), *t*(14) = -1.36, *p* = 0.197, forward digit span score (younger *M* = 11.80, *SD* = 1.69; older *M* = 11.17, *SD* = 2.40), *t*(14) = 0.62, *p* = 0.544, backward digit span (younger *M* = 6.60, *SD* = 0.84; older *M* = 7.17, *SD* = 1.72), *t*(14) = -0.89, *p* = 0.388, backward digit span score (younger *M* = 10.20, *SD* = 1.55; older *M* = 9.67, *SD* = 2.16), *t*(14) = 0.58, *p* = 0.573, and total digit span test score (younger *M* = 22.00, *SD* = 2.62; older *M* = 22.00, *SD* = 2.83), *t*(14) = 0.00, *p* = 1.00.

Pearson product-moment correlations were calculated to assess the relationship between Stroop interference scores and mean digit recall difference scores, as well as between Stroop interference scores and digit recall difference scores in each noise condition (see Table 3). A moderate, negative correlation between Stroop interference scores and mean digit recall difference scores was found for the quiet condition, $r = -0.51$, n = 16, $p = 0.046$. However, with a Bonferroni corrected alpha value of 0.0125, this correlation was not significant. There were no other correlations between the variables. Pearson product-moment correlations were also calculated to assess relationships between mean digit recall difference scores and forward digit span, forward digit span score, backward digit span, backward digit span score, and total digit span test score. Correlations were also calculated for digit recall difference scores in each noise condition and the various digit span measures (see Table 4). No significant correlations were found between any of the digit span variables and digit recall difference scores.

Table 3. Correlations between Stroop interference scores and digit recall difference scores

Variable 1	Variable 2	*r*	*p*
Stroop interference score	Mean digit recall difference score	-0.19	0.488
Stroop interference score	Quiet digit recall difference score	-0.51	0.046
Stroop interference score	+16 dB SNR digit recall difference score	0.09	0.741
Stroop interference score	+12 dB SNR digit recall difference score	-0.01	0.972

Table 4. Correlations between digit span scores and digit recall difference scores

Variable 1	Variable 2	*r*	*p*
Forward digit span	Mean digit recall difference score	0.19	0.476
Forward digit span	Quiet digit recall difference score	0.17	0.532
Forward digit span	+16 dB SNR digit recall difference score	0.10	0.711
Forward digit span	+12 dB SNR digit recall difference score	0.23	0.394
Forward digit span score	Mean digit recall difference score	-0.17	0.535
Forward digit span score	Quiet digit recall difference score	-0.34	0.199
Forward digit span score	+16 dB SNR digit recall difference score	-0.002	0.993
Forward digit span score	+12 dB SNR digit recall difference score	-0.06	0.834
Backward digit span	Mean digit recall difference score	0.09	0.753
Backward digit span	Quiet digit recall difference score	-0.01	0.959
Backward digit span	+16 dB SNR digit recall difference score	0.08	0.768
Backward digit span	+12 dB SNR digit recall difference score	0.17	0.522
Backward digit span score	Mean digit recall difference score	0.04	0.891
Backward digit span score	Quiet digit recall difference score	-0.36	0.174
Backward digit span score	+16 dB SNR digit recall difference score	0.33	0.207
Backward digit span score	+12 dB SNR digit recall difference score	0.17	0.519
Total digit span test score	Mean digit recall difference score	-0.02	0.947
Total digit span test score	Quiet digit recall difference score	-0.31	0.236
Total digit span test score	+16 dB SNR digit recall difference score	0.17	0.540
Total digit span test score	+12 dB SNR digit recall difference score	0.15	0.571

Discussion

In the present study, the relationship between age, cognitive ability, and listening effort was examined in three different background noise conditions. The cognitive abilities of selective attention and working memory were assessed with the Stroop test and digit span test, respectively. Listening effort was measured through a dual task paradigm in which the primary task was a speech recognition task using spectrally degraded vocoded speech stimuli in quiet and in +16 dB SNR and +12 dB SNR multi-talker background noise conditions, and the secondary task was a digit recall task. Specifically, the change in digit recall score when going from a single task (i.e., performing only the digit recall task) condition to a dual task condition (i.e., simultaneously performing the speech recognition and digit recall tasks) was assessed in each noise condition among younger and older listeners as a measure of listening effort. Speech recognition performance alone was also measured in each noise condition. It was hypothesized that listening effort would increase (reflected by a greater change in digit recall scores when going from single to dual task conditions) as background noise conditions became more difficult, and that older performers would demonstrate greater listening effort compared to the younger listeners. Finally, it was hypothesized that there would be a correlation between the measures of selective attention and listening effort, and working memory and listening effort.

Unsurprisingly, results demonstrated that speech recognition scores were poorer in poorer signal-to-noise ratio conditions. In addition, this study found that overall, speech recognition scores were poorer for older listeners compared to younger listeners. Some studies have suggested that differences in speech recognition performance can largely be accounted for by differences in audiometric pure-tone thresholds (e.g., Abel et al., 2000; Souza & Turner, 1994; Takahashi & Bacon, 1992). However, similar to the results of this study, a number of studies in the literature have found age itself to be a factor in speech recognition scores. Dubno et al. (1994) looked at younger and older participants with and without hearing loss, and found that that while there was no effect of age on speech recognition performance in quiet, in background noise, older participants performed poorer compared to younger participants regardless of hearing loss. The authors (Dubno et al., 1994) also found that the predicted audibility of stimuli and the articulation index for each participant could not predict speech recognition performance in noise for older participants while it could for younger participants. Gordon-Salant and Fitzgibbons (1993) also found an effect of age on a speech recognition task of low-predictability sentences in background noise among younger and older participants with and without hearing loss. Tun and Wingfield (1999) found that while there was no age-related difference in speech recognition performance in quiet between younger and older participants with normal hearing up to 3000 Hz, older adults performed significantly worse in background noise. While the studies mentioned above used speech stimuli that were degraded only by the addition of background noise, more similar to the vocoded sentences used in this study, Schvartz, Chatterjee, and Gordon-Salant (2008) used vocoded stimuli that were spectrally shifted to simulate cochlear implant processing. The authors examined the phoneme recognition ability of younger, middle-age, and older normal-hearing participants, and found that the younger group performed significantly better than the two older groups, consistent with the findings of this study that the younger group performed better in the vocoded speech recognition tasks compared to the older group.

The measure of interest in this study, the performance change in the secondary digit recall task, did show that digit recall performance was significantly poorer in the dual task condition compared to the single task condition, suggesting that the word recognition task required sufficient extra cognitive resources, i.e., listening effort, causing a decline in performance on the secondary digit recall task. This interpretation of a decline in performance in a secondary task in a dual task paradigm is consistent with other studies that have used dual task paradigms to measure listening effort (Desjardins, 2011; Desjardins & Doherty, 2013; Downs, 1982; Gosselin & Gagne, 2011; Howard et al., 2010; McCreery & Stelmachowicz, 2013; McGarrigle et al., 2014; Ninio & Kahneman, 1974; Pichora-Fuller, 2003; Rabbitt, 1968; Rackerd et al., 1996; Tun et al., 2009; Wingfield et al., 2005). However, in this study, no significant relationship between age and listening effort, noise condition and listening effort, or age, noise condition, and listening effort was found. This finding is contrary to the evidence in the literature, which suggests that there is a relationship between age and listening effort, and between signal-to-noise ratio and listening effort. Findings from previous studies are discussed below.

Tun et al. (2009) found an overall effect of age, and an interaction of age and hearing loss on listening effort, as measured by decline in a secondary visual tracking task in a dual task paradigm, among younger and older listeners with normal hearing and hearing loss. In particular, the authors found that older adults' performance in the secondary task declined to a greater degree in the dual task condition compared to younger adults, and that hearing loss also had a greater effect on performance for the older adults than for the younger adults. A study examining listening effort using a dual task paradigm in younger normal-hearing listeners and older listeners with normal hearing and with hearing loss also found that older listeners demonstrated greater listening effort in speech in noise (two-talker babble and speech-shaped noise) recognition tasks compared to younger listeners, although no effect of age in the difficult six-talker babble background noise condition was found (Desjardins, 2011; Desjardins & Doherty, 2013). Gosselin and Gagne (2011) also found that older adults with normal hearing through 3000 Hz demonstrated greater listening effort, as measured by decline in accuracy and response in a secondary task in a dual task paradigm, compared to younger adults with normal hearing. They also found that better signal to noise ratios for the primary listening task resulted in less of a decline in performance for the secondary task for the older adults, reflecting less listening effort (Gosselin & Gagne, 2011). Tun, Wingfield, and Stine (1991) also found that listening effort, as measured by a dual task paradigm, was greater for older adults compared to younger adults who reported no hearing difficulties, although it is important to note that audiometric data was not collected for the participants in this study, and therefore an effect of hearing loss cannot be ruled out. Schurman, Brungart, and Gordon-Salant (2014) also found that listening effort, as measured by ability to rely on working memory to maintain speech recognition performance in a memory-dependent speech recall task, was greater for older adults with normal hearing compared to younger adults with normal hearing.

One explanation for the lack of significant findings in the relationships between age, background noise condition, and listening effort is that the data in this study had low statistical power due to the small sample size of 16 participants. As discussed in the preceding results section, the small sample size of this study did not provide sufficient power to detect large or medium effects of age on digit recall scores, or to detect large or medium interactions of age and task, or age, noise, and task on digit recall scores. Therefore, while results yielded non-significant effects and interactions, and the null hypotheses that there are no effects of the independent variables on the digit recall scores cannot be rejected, it should not be concluded that the null hypothesis must be accepted. While it cannot be said that the results of this study reflect the findings of other studies, it should not be concluded that this study provides contradictory findings; rather it should be stated that the data in this study are not sufficient to confidently evaluate the hypotheses put forth regarding the relationships between age, noise condition, and listening effort.

A possible explanation for the discrepancy in findings of an age effect and an interaction of age and signal-to-noise ratio on listening effort for this current study compared to findings in the literature is that speech recognition performance, i.e., the primary task in this study's experimental design, was not maintained across the single task and dual task conditions by all individual participants. While group analysis demonstrates that there was no significant difference between the single and dual task speech recognition scores, a closer look at individual data demonstrates that in some cases, individual participants did not maintain performance in the single and dual primary tasks, suggesting that a practice effect or participant attention or motivation may have confounded dual task speech recognition performance, and therefore secondary-task performance in the dual task condition for these participants may not accurately reflect effort exerted on the primary task. While the design of this study attempted to control for practice effects by counterbalancing the order of presentation of dual and single task conditions, as well as the order of presentation of background noise conditions, the small sample size, and the unequal group sizes (the older group consisted of 6 participants, and the younger group 10 participants) ultimately yielded a design in which order effects could not be expected to be adequately counterbalanced. In addition, while the experimenter in this study attempted to maintain participant performance on the speech recognition task by emphasizing that focus should remain on the primary speech recognition task and not the secondary digit recall task, and an attempt was made to avoid participant fatigue and boredom by enforcing at least one break during the experimental session, it is clear that such attempts may not have been sufficient.

Another possible reason for the difference in results seen in this study compared to the literature may stem from the type of task used for the secondary task. In this study, a string of seven digits was presented visually on a computer screen and participants were later asked to recall the string of digits. However, it cannot be ascertained whether participants used a strategy that would suggest reliance on a language-based skill (e.g., rehearsal) or if they utilized an image recall strategy, thereby tapping into a skill set that may not use the same resources as language-based tasks do. This in itself may not present as a problem, as it would be expected that regardless of the type of skill used to perform the secondary task, a decline in cognitive resources would cause a subsequent decline in secondary task performance. For example, Desjardins (2011) and Desjardins and Doherty (2013) used a visual tracking task, Gosselin and Gagne (2011) used a tactile pattern recognition task, and Sarampolis et al. (2009) used a word recall task as secondary tasks to successfully measure listening effort. However, while the modality the secondary task utilizes may not be important, it is important to consider whether a secondary task may allow for the use of multiple strategies or skills. It is possible that participants utilized and alternated between both, image recall and rehearsal, for the digit recall in the single and dual task conditions, thereby calling into question the validity of using a linear decline in digit recall performance from the single to dual task conditions as a measure purely of resources recruited to perform the primary speech recognition task. If one strategy, for example, image recall, utilizes less cognitive resources than rehearsal, then a participant who switches from using rehearsal during the single task to image recall during the dual task may not demonstrate a decline in secondary task performance, although he or she may still experience an increase in cognitive load, or listening effort. These potential difficulties in measuring listening effort in the current study may explain the lack of an effect of age and background noise seen on the decline in secondary task performance, in contrast to the effect seen in the literature, and may also impact the findings regarding the correlations of cognitive abilities and listening effort, described below.

This study compared the measure of listening effort to the interference scores calculated from the Stroop test, and to the digit span test scores. It was found that younger listeners had better scores on the Stroop test compared to the older listeners, consistent with other findings that older adults perform more poorly on selective attention tasks compared to younger adults (Cohn, Dustman, & Bradford, 1984; West & Alain, 2000). There was also a negative moderate correlation found between the Stroop test interference scores and the listening effort measure for the quiet noise condition, although this correlation did not rise to significance with a Bonferroni correction. There were no other significant correlations between the Stroop test and listening effort measures, or between the digit span scores or listening effort measures. The lack of significant correlations between Stroop interference scores and listening effort measures is consistent with a study reported by Desjardins (2011) and Desjardins and Doherty (2013), in which no correlations between Stroop test scores and listening effort measures were found. However, the finding of this study that there are differences in selective attention measures between age groups differs from the lack of age-related differences in Stroop interference scores reported by Desjardins (2011) and Desjardins and Doherty (2013). It is interesting to note, however, that the Stroop test administration used by the above-mentioned authors differed from the administration in this current study, as did the method used to calculate the Stroop interference score. In this study, the Stroop test was administered and interference scores were calculated by taking into account predicted performance on the selective attention task, as described by Golden (1978), while Desjardins (2011) and Desjardins and Doherty (2013) used an administration and calculation that did not account for predicted performance on the selective attention task, as described by Kemper, Schmalzried, Herman, Leedahl, and Mohankumar (2009). These differences in administration and calculations of the Stroop test may account for the discrepancy in findings about age-related differences in Stroop test scores between this study and that of Desjardins (2011) and Desjardins and Doherty (2013).

Similar to the results of this current study, Desjardins (2011) and Desjardins and Doherty (2013) found that there was no significant difference between the age groups in working memory scores, measured by the Reading Span Test (RST, as described by Daneman & Carpenter, 1980). However, the authors also found that while there was no correlation between RST scores and listening effort measures in the six-talker babble background noise condition, there were negative correlations between the RST scores and listening effort in the two-talker and speech-shaped noise background conditions (Desjardins, 2011; Desjardins & Doherty, 2013). A relationship between working memory and listening effort was also demonstrated by Schurman et al. (2014); the authors found a significant correlation between working memory span, as measured by the listening span test (LSPAN, as described by Daneman & Carpenter, 1980), and listening effort. Tun, et al. (1991) also found a relationship between working memory span and listening effort. The authors measured working memory using the LSPAN and RST (Daneman & Carpenter, 1980), and found that working memory ability in participants was correlated to listening effort; they also found that when participants in younger and older normal-hearing groups were matched by working memory ability, age-related differences in listening effort disappeared (Tun et al., 1991).

The discrepancy between results from this study, which did not show a relationship between working memory and listening effort, compared to evidence from the literature may be explained by the different measures of working memory used. In this study, working memory was measured with a digit span test adapted from the WAIS digit span test, using a simultaneous visual presentation of all digits in a string. Other studies mentioned above used measures of working memory that used auditory modalities or language-based skills; therefore it may be the case that the test used in this current study may not have been sensitive to working memory ability as it relates to listening tasks. This may be because when using a visual presentation of digits and asking participants to recall the stimuli, participants may not have solely relied on a language-based memory strategy (e.g., rehearsal), a similar issue described above as it pertains to the secondary task of the dual task paradigm in this study.

In conclusion, the results of this study are not consistent with the literature, which shows that older adults tend to exert greater listening effort than younger adults. As described above, methodological issues in this study may account for this discrepancy. While the use of the dual task paradigm to measure listening effort is supported by the evidence in the literature, it is critical that performance on the primary speech recognition task be maintained, and that a better method is implemented for monitoring individual participant's performance during data collection. One possible way to ensure maintenance of primary task performance may be to split data collection into two separate sessions for all participants, and provide a short break after each task, allowing time for the experimenter to examine and compare scores. Another possibility is to provide a reward for maintenance of performance on the primary task (Kerr, 1973), as well as a lesser reward for good performance on the secondary task. It is also important that the secondary task used in such a dual task paradigm require the use of a single processing strategy. Therefore, instead of a visual presentation of digit strings, an auditory presentation may be better suited for the dual task paradigm, or a different task, such as a visual tracking task similar to that reported by Desjardins (2011) and Desjardins and Doherty (2013). Finally, it is suggested that future studies utilizing a dual task paradigm to examine the effect of age, cognitive ability, and signal-to-noise ratio on listening effort obtain a larger sample size, to achieve acceptable statistical power, and to adequately counterbalance task order and background noise condition order. These efforts will help to ensure that participant motivation is controlled, that performance change in the secondary task reflects a valid measure of an increase in use of cognitive resources for the primary task, and that practice or order effects do not confound results.

The literature also suggests that there may be a correlation to listening effort and cognitive abilities, specifically working memory and processing speed. As mentioned above, the measure of working memory used in this current study, a digit recall task of visually presented stimuli, may not accurately and consistently measure working memory ability across individual participants. Future studies may consider using other working memory tests, such as the RST as described by Daneman and Carpenter (1980). The RST, which has shown more consistent correlation to speech recognition performance compared to the digit span test (Akeroyd, 2008), may also show more consistent correlation with listening effort, as it did in the study reported by Desjardins (2011) and Desjardins and Doherty (2013). Processing speed was not examined in this study, but it may be worth continuing to examine this cognitive ability in future studies, as the literature has shown a correlation to listening effort (Zekveld, Kramer, & Festen, 2011; Desjardins, 2011; Desjardins & Doherty, 2013). The potential for a correlation between selective attention and listening effort also seems worth pursuing in future studies. While the study reported by Desjardins (2011) and Desjardins and Doherty (2013) did not find a difference in selective attention ability between younger and older participants, and they did not find a correlation between selective attention ability and listening effort, this current study did find a significant difference in selective attention ability between age groups utilizing a different approach to the administration and interpretation of the Stroop test results (based on Golden, 1978). While no significant correlations were found between selective attention ability and listening effort, it is difficult to conclude whether the results were confounded by difficulties in measuring listening effort as described above. Therefore further investigation into the relationship between selective attention and listening effort for speech in noise recognition tasks may still yield interesting results.

Conclusion

In summary, the findings of this study are as follows:

1) Speech recognition ability was poorer for older adults compared to younger adults.
2) Speech recognition ability was poorer in poorer noise conditions, with performance being best in the quiet condition, and worst in the +12 dB SNR condition.
3) There was an overall decline in the digit recall scores in the dual task condition compared to the single task condition, indicating that overall, the dual task condition required greater listening effort than the single task condition.
4) Stroop test interference scores were higher for the younger group than the older group, consistent with better selective attention ability in the younger group.

References

Abel, S., Sass-Kortsak, A., & Naugler, J. (2000). The role of high-frequency hearing in age-related speech understanding deficits. *Scandinavian Audiol*ogy, *29*(3), 131-138.

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*(S2), S53-S71.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging affects neural precision of speech encoding. *The Journal of Neuroscience*, *32*(41), 14156-14164.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*(3), 189-208.

Bernarding, C., Strauss, D., Hannemann, R., Seidler, H., & Corona-Strauss, F. (2013). Neural correlates of listening effort related factors: Influence of age and hearing impairment. *Brain Research Bulletin*, *91*, 21-30.

Bertoli, S., & Bodmer, D. (2014). Novel sounds as a psychophysiological measure of listening effort in older listeners with and without hearing loss. *Clinical Neurophysiology*, *125*(5), 1030-1041.

Bess, F. H., Dodd-Murphy, J., & Parker, R. A. (1998). Children with minimal sensorineural hearing loss: Prevalence, educational performance, and functional status. *Ear and Hearing, 19*(5), 339-354.

Borg, G. (1990). Psychophysical scaling with applications in physical work and the perception of exertion. *Scandinavian Journal of Work, Environment & Health*, *16*(1), 55-58.

Chaftez., M., & Matthews, L. (2004). A new interference score for the Stroop test. *Archives of Clinical Neuropsychology*, *19,* 555-567.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohn, N. B., Dustman, R. E., & Bradford, D. C. (1984). Age-related decrements in Stroop color test performance. *Journal of Clinical Psychology*, *40*(5), 1244-1250.

Cox, R., & Alexander, G. (1995). The abbreviated profile of hearing aid benefit. *Ear and Hearing, 16*(2), 176–186.

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

Desjardins, J. (2011). *Age-related changes in listening effort for various types of masker noises*. (Unpublished doctoral dissertation). Syracuse University, Syracuse, NY.

Desjardins, J., & Doherty, K. (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.

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

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

Fitzgibbons, P., & Gordon-Salant, S. (2004). Age effects on discrimination of timing in auditory sequences. *The Journal of the Acoustical Society of America*, *116*(2), 1126-1134.

Folstein, M., Folstein, S., & McHugh, P. (1975). “Mini-mental state:” a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198.

Gatehouse S., Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *International Journal of Audiology, 43*(2):85–99.

Getzmann, S. (2012). Handicapped due to age? Behavioral and electrophysiological correlates of speech perception of dichotically presented narratives in young and middle-aged listeners. *Journal of Psychophysiology*, *26*(3), 132-144.

Golden, C. (1978). *Stroop color and word test*. Wood Dale, IL: Stoelting Co.

Gordon-Salant, S., & Fitzgibbons, P. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech, Language, and Hearing Research*, *36*(6), 1276-1285.

Gosselin, P., & Gagne, J. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech, Language, and Hearing Research*, *54*(3), 944-958.

Hällgren, M., Larsby, B., Lyxell, B., & Arlinger, S. (2005). Speech understanding in quiet and noise, with and without hearing aids. *International Journal of Audiology*, *44*(10), 574-583.

Harris, K. C., Eckert, M. A., Ahlstrom, J. B., & Dubno, J. R. (2010). Age-related differences in gap detection: Effects of task difficulty and cognitive ability. *Hearing Research*, *264*(1-2), 21-29.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139-183.

Hasher, L., Stoltzfus, E., Zacks, R., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 163-169.

Hicks, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, *45*(3), 573-584.

Hol. M. K., Kunst, S., Snik, A. F., Bosman, A. J., Mylanus, E., & Cremers, C. W. (2010a). Bone-anchored hearing aids in patients with acquired and congenital unilateral inner ear deafness (Baha CROS): Clinical evaluation of 56 cases. *Annals of Otology, Rhinology & Laryngology*, *119*(7), 447-454.

Hol, M. K., Kunst, S., Snik, A. F., & Cremers, C. W. (2010b). Pilot study on the effectiveness of the conventional CROS, the transcranial CROS and the BAHA transcranial CROS in adults with unilateral inner ear deafness. *European Archives of Otorhinolaryngology, 267*, 889-896.

Howard, C. S., Munro, K. J., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology*, *49*(12), 928-932.

Hughes, K. C., & Galvin, K. L. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear Implants International*, *14*(3), 121-129.

Humes, L. (1996). Speech understanding in the elderly. *Journal of the American Academy of Audiology*, *7*, 161-167.

Humes, L., Lee, J., & Coughlin, M. (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.

Kaiser A., Svirsky M. (2000, October 15-18). Using a personal computer to perform real-time signal processing in cochlear implant research. Paper presented at the Proceedings of the IXth IEEE-DSP Workshop, Hunt, TX.

Kemper, S., Schmalzried, R., Herman, R., Leedahl, S., & Mohankumar, D. (2009). The effects of aging and dual task demands on language production. *Aging, Neuropsychology, and Cognition*, *16*(3), 241-259.

Kerr, B. (1973). Processing demands during mental operations. *Memory & Cognition*, *1*(4), 401-412.

Killian, G. (1985). The Stroop color-word interference test. In Keyser, D. and Sweetland, R. (Eds.), *Test Critiques*, Vol. II, Kansas City, MO: Test Corporation of America, 751-758.

Lansbergen, M., Kenemans, J., & van Engeland, H. (2007). Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. *Neuropsychology*, *21*(2), 251-262.

Larsby, B., Hällgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology*, *44*(3), 131-143.

Mackersie, C., & Cones, H. (2011). Subjective and psychophysiological indices of listening effort in a competing-talker task. *Journal of the American Academy of Audiology*, *22*(2), 113-122.

MATLAB version 8.2.0 [Computer software]. (2013). Natick, MA: The MathWorks Inc.

McCreery, R. W., & Stelmachowicz, P. G. (2013). The effects of limited bandwidth and noise on verbal processing time and word recall in normal-hearing children. *Ear and Hearing*, *34*(5), 585-591.

McGarrigle, R., Munro, K., Dawes, P., Stewart, A., Moore, D., Barry, J., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper.' *International Journal of Audiology*, *53*(7), 433-445.

Ninio, A., & Kahneman, D. (1974). Reaction time in focused and in divided attention. *Journal of Experimental Psychology*, *103*(3), 394-399.

O’Bryant, S., Humphreys, J., Smith, G., Ivnik, R., Graff-Radford, N., Petersen, R., & Lucas, J. (2008). Detecting dementia with the mini-mental state examination in highly educated individuals. *Archives of Neurology*, *65*(7), 963-967.

O'Gorman, J., & Lloyd, J. (1988). Electrodermal lability and dichotic listening. *Psychophysiology*, *25*(5), 538-546.

Pichora-Fuller, M. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology*, *42*, 2S26-2S32.

Pichora-Fuller, M., & 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.

Pichora-Fuller, M., & Souza, P. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology*, *42*(S2), 11-16.

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

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

Rudner, M., Foo, C., Rönnberg, J., & Lunner, T. (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scandinavian Journal of Psychology, 50*(5), 405-418.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403-428.

Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research*, *52*(5), 1230-1240.

Schurman, J., Brungart, D., & Gordon-Salant, S. (2014). Effects of masker type, sentence context, and listener age on speech recognition performance in 1-back listening tasks. *The Journal of the Acoustical Society of America*, *136*(6), 3337-3349.

Schvartz, K. C., Chatterjee, M., & Gordon-Salant, S. (2008). Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listenersa). *The Journal of the Acoustical Society of America*, *124*(6), 3972-3988.

Sherbecoe, R., & Studebaker, G. (2004). Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units. *International Journal of Audiology*, *43*, 442-448.

Smith, S., Pichora-Fuller, M., Wilson, R., & MacDonald, E. (2012). Word recognition for temporally and spectrally distorted materials: The effects of age and hearing loss. *Ear and Hearing*, *33*(3), 349-366.

Souza, P., & Tuner, C. (1994). Masking of speech in young and elderly listeners with hearing loss. *Journal of Speech and Hearing Research*, *37*(3), 655-661.

Spahr, A., & Dorman, M. (2004). Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices. *Archives of Otolaryngology – Head and Neck Surgery*, *130*, 624-628.

Spahr, A., Dorman, M., Litvak, L., Van Wie, S., Gifford, R., Loizou, …, Cook, S. (2012). Development and validation of the AzBio sentence lists. *Ear and Hearing*, *33*(1), 112-117.

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

Takahashi, G., & Bacon, S. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech, Language, and Hearing Research*, *35*(6), 1410-1421.

Tremblay, K., Billings, C., & Rohila, N. (2004). Speech evoked cortical potentials: effects of age and stimulus presentation rate. *Journal of the American Academy of Audiology*, *15*(3), 226-237.

Tremblay, K., Piskosz, M., & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology*, *114*(7), 1332-1343.

Tun, P., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761-766.

Tun, P., & Wingfield, A. (1999). One voice too many: Adult age differences in language processing with different types of distracting sounds. *Journal of Gerontology: Psychological Sciences*, *54B*(5), P317-P327.

Tun, P., Wingfield, A., & Stine, E. (1991). Speech-processing capacity in young and older adults: A dual-task study. *Psychology and Aging*, *6*(1), 3-9.

van Rooij, J. C. G. M., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *The Journal of the Acoustical Society of America*, *88*(6), 2611-2624.

Wasson, J. H., Kairys, S. W., Nelson, E. C., Kalishman, N., & Baribeau, P. (1994). A short survey for assessing health and social problems of adolescents. *The Journal of Family Practice*, *38*(5), 489-494.

Wechsler, D. (2008). Wechsler’s Adult Intelligence Scale – Fourth Edition (WAIS-IV). San Antonio, TX: Pearson.

West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology*, *37*(2), 179-189.

Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood what it is and how it interacts with cognitive performance. *Current Directions in Psychological Science*, *14*(3), 144-148.

Wong, P. C., Ettlinger, M., Sheppard, J. P., Gunasekera, G. M., & Dhar, S. (2010). Neuroanatomical characteristics and speech perception in noise in older adults. *Ear and hearing*, *31*(4), 471-479.

Wong, P. C., Uppunda, A. K., Parrish, T. B., & Dhar, S. (2008). Cortical mechanisms of speech perception in noise. *Journal of Speech, Language, and Hearing Research*, *51*(4), 1026-1041.

Zekveld, A., Kramer, S., & Festen, J. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and Hearing*, *31*(4), 480-490.

Zekveld, A., Kramer, S., & Festen, J. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing*, *32*(4), 498-510.