Summer 2019

The Effect of Spectral Shaping on Perceptual, Acoustic, and Listening Effort Measurements in Young Normal Hearing Adults

Rachel Madorskiy

Recommended Citation

Madorskiy, R.(2019). *The Effect of Spectral Shaping on Perceptual, Acoustic, and Listening Effort Measurements in Young Normal Hearing Adults.* (Doctoral dissertation). Retrieved from https://scholarcommons.sc.edu/etd/5491

THE EFFECT OF SPECTRAL SHAPING ON PERCEPTUAL, ACOUSTIC, AND LISTENING EFFORT MEASUREMENTS IN YOUNG NORMAL HEARING ADULTS

by

Rachel Madorskiy

Bachelor of Arts
University of Louisville, 2014

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Communication Sciences and Disorders

The Norman J. Arnold School of Public Health

University of South Carolina

2019

Accepted by:

Daniel Fogerty, Major Professor

Judy Dubno, Committee Member

Al Montgomery, Committee Member

Jason Wigand, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ACKNOWLEDGEMENTS

Throughout the writing of this dissertation I have received a great deal of support and assistance. I am especially indebted to my advisor, Dr. Daniel Fogerty, whose expertise and advice was invaluable in the formulation of the research topic, methodology, and analysis of this project. Without your guidance and persistent help this dissertation would not have been possible. Furthermore, you provided me with the tools that I needed to successfully complete my dissertation. Finally, you have served as a role model for what a good scientist (and person) should be.

I would also like to thank the rest of my committee, Drs. Judy Dubno, Al Montgomery, and Jason Wigand for their insight and support, of which I am grateful.

In addition, I would like to thank my colleagues Dr. Bobby Gibbs, Ansley Potter, Blythe Vickery, Briemma Wilson, Caralee Lowry, and Lauren Sudduth with whom I had the pleasure of working with throughout the duration of this project, and who were essential in completing this research.

Finally, I would like to thank my husband, Arty Madorskiy, my parents, Dwight and Karen Miller, and my brother, Ian Miller, for their wise counsel and sympathetic ear. You have always been there for me, and that support was essential in the completion of this project.

ABSTRACT

Amplification is necessary to restore audibility to individuals with hearing impairment. However, frequency-specific amplification alters the spectral shape and level of speech. Previous research has demonstrated that amplification is associated with better speech recognition outcomes as compared to unaided listening (e.g. Humes, 2013). However, many hearing aid listeners report a lack of perceived benefit associated with hearing aids, and adherence is low (McCormack and Fortnum, 2013). This may, in part, be reflected by a common complaint of hearing aid users that speech is audible, but additional effort is required to understand speech. As amplification alters the spectral shape and overall level of speech, it is possible that more effort could be required to understand amplified speech, even in situations where accuracy remains equivalent. Furthermore, previous research has suggested that ratings of perceived sound quality differ with various hearing aid processing techniques (Neuman et al., 1998). Therefore, it is possible that varying spectral shape and speech level could have effects on perceived sound quality as well. This has clinical implications, as poor perceived sound quality is a top complaint of hearing aid users (McCormack and Fortnum, 2013). The current study aimed to systematically investigate the effect of spectral shape and speech level on measures of speech recognition, listening effort, and sound quality. Outcomes were investigated in three different background conditions. In addition, acoustic metrics were utilized to quantify the alterations made to speech stimuli amplified in a frequency-specific manner.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

GENERAL INTRODUCTION

One in eight U.S. adults has been diagnosed with hearing loss (Lin, Niparko, & Ferrucci, 2011). Despite this, less than 20 percent of these adults use hearing aids (NIDCD, 2010), and adherence is low in hearing aid (HA) users, in part, due to lack of perceived benefit (McCormack & Fortnum, 2013). Indeed, a top complaint of HA users is that speech is understandable, but listening is effortful and fatiguing. However, restoring audibility has long been associated with better speech recognition outcomes, especially in quiet (e.g. Humes, 2013). Assessing the subjective experiences associated with amplified speech is required to address understand HA user complaints, despite experimental evidence of gains related to HA use. Further research is needed to understand speech perception associated with amplification in order to better understand the amount of benefit experienced by HA listeners. This proposal seeks to systematically investigate the effects of amplification on recognition, as well as subjective, outcomes.

While amplification has been widely investigated in relation to speech recognition outcomes, it is unclear what perceptual processing is necessary to understand spectrally shaped speech. Two existing theories could explain how spectrally shaped speech is perceptually processed: the Ease of Language Understanding model (ELU; Ronnberg et al., 2017) and the peripheral hypothesis (Humes, 1996). A simple view of the peripheral hypothesis suggests that audibility, or access to the acoustic signal, is the most important factor in determining successful speech recognition. According to this theory, the more audible the signal is for the listener, the less perceptual processing demands are needed to understand the speech. This simple view suggests that spectrally shaped speech should be more easily understood than unamplified speech, due to increased audibility. In addition, when speech is spectrally shaped, there are changes in overall level, which could also affect audibility. In contrast, ELU states that the acoustic similarity between a stimulus and its stored mental representation affects lexical processing. Due to the acoustic changes that occur when speech in amplified, this mismatch between heard and stored representation could be larger for shaped than normal speech. Because of this mismatch, greater perceptual processing demands would be needed to understand spectrally shaped speech.

In order to better understand if perceptual processing demands increase, or decrease, as a result of spectral shaping, a behavioral outcome could be used to reflect these demands. This mechanism can be found in the Framework for Understanding Effortful Listening (FUEL, Pichora-Fuller et al., 2016). FUEL is a theoretical explanation for why listening can be effortful or fatiguing in certain listening conditions. FUEL suggests that there is a finite capacity of resources available when engaged in listening, a concept adapted from Kahneman’s model of attentional capacity (Kahneman, 1973). Generally, the more difficult the task, the larger the amount of resources are required to successfully listen to and understand speech. The more resources that a task requires, the more effortful it becomes. This concept of a capacity of resources may be a reflection of the amount of perceptual processing demands that are required to understand spectrally shaped speech. According to the simple peripheral hypothesis, amplification would increase audibility and reduce the resources required to understand speech. This would result in the perception of lower effort associated with the task. In contrast, according to ELU, a mismatch with stored mental representations as a result of amplification would require more resources to successfully understand speech. This, in turn, would result in the perception of increased effort. Listening effort outcomes could reveal the perceptual processing demands required in order to understand spectrally shaped speech.

Because of the acoustic changes that occur as a result of amplification, distortion may be introduced, which could increase or decrease perceptual processing demands. It is important to quantify this distortion in order to better understand how it might affect processing. Quantifying the amount of distortion introduced from spectral shaping may inform the processing resources required to understand speech, as well as how to interpret recognition and listening effort outcomes. One way to quantify distortion associated with amplification is by examining perceived sound quality. While speech recognition measures assess the intelligibility of amplified speech, sound quality measures may indicate how distorted, or unnatural, spectrally shaped speech sounds to the listener. In addition to this subjective indicator of distortion (or acceptability), acoustic analysis may objectively indicate the amount of distortion that is introduced from spectral shaping. One such acoustic metric is the Hearing Aid Speech Quality Index (HASQI; Kates & Arehart, 2014b). The combination of subjective and objective metrics could aid in the understanding of how factors of audibility and distortion affect perceptual processing demands that in turn affect recognition and effort.

While the peripheral hypothesis and ELU were developed within the context of a quiet listening environment, listeners are often tasked with understanding speech in noisy environments as well. In the presence of background noise, perceptual processing demands may increase within both theoretical constructs. In addition, the nature of the background noise, such as whether it is steady-state or fluctuating, could affect perceptual processing demands. Within ELU, the addition of noise could increase the ambiguity of the stimuli. With less information available, more perceptual processing may be required to match the stimulus to its mental representation. In fluctuating maskers, more information may be available due to glimpses in the masker, which could facilitate lexical retrieval in comparison to steady-state maskers. Within the simple peripheral hypothesis, the addition of noise would decrease audibility, which would increase perceptual processing demands. However, in fluctuating maskers, listeners may be able to take advantages of glimpses in the masker, where audibility is momentarily improved.

Examining speech recognition, listening effort, and sound quality outcomes associated with spectrally shaped speech will aid in better understanding of how amplified signals are perceptually processed. Measures of speech recognition inform how accurately the listener is able to perceptually process speech. However, speech recognition measures alone don’t predict the subjective experiences of hearing aid users, which often include a lack of perceived benefit and low adherence. The addition of listening effort outcomes could explain differences in perceptual processing demands for spectrally shaped speech in comparison to unaltered speech. Finally, sound quality measurements aid in understanding the level of distortion present in a stimulus, which could affect both perceptual processing demands, as well as speech recognition outcomes. In addition, collecting multiple outcomes measurements from the same listeners may more accurately explain HA user experience compared to one measure alone. The convergence of these multiple measures could help to explain the underlying perceptual model utilized when understanding amplified speech, which could inform hearing aid users' subjective experiences as well as audiological outcomes.

Current Study

When the spectrum of speech is altered to account for audibility, there are changes regarding both the speech level and spectral qualities. However, speech recognition outcomes with spectrally shaped speech have been mixed. In most cases, spectral shaping appears to aid in speech recognition (e.g. Horwitz, Ahlstrom & Dubno, 2008; Turner & Henry, 2002), however, there is also evidence of poorer speech recognition with spectral shaping, especially when high frequencies are amplified (e.g. Amos & Humes, 2007; Baer, Moore, & Kluk, 2002). It is unclear if this detriment is due to changes in spectral qualities or speech levels. In order to investigate this, the current study varied spectral shape and speech level in a speech recognition task for young normal hearing listeners. The young normal hearing population was tested to avoid factors related to altered suprathreshold processing due to cochlear pathology and focus on the consequences of acoustic HA processing relative to normal hearing without processing. Average thresholds from a group of listeners with hearing-impairment were used to simulate normal and hearing impaired thresholds. Speech recognition, listening effort, and sound quality ratings were evaluated in three listening conditions: quiet, steady-state noise, and speech modulated noise to assess the effect of spectral shape and speech level in common masking paradigms. This study represents a direct method of comparing the combined effects of level and shaping. The results of this study could be further extended to the hearing-impaired auditory system and could have implications for how we understand how speech is perceived via hearing aids.

The primary purpose of this project was to understand the relationship between spectral shape and speech level. Experiment 1 examined the effect of these factors on speech recognition. Acoustic measures were conducted in order to identify potential acoustic factors that may contribute to performance. Experiment 2 examined the effect of spectral shape and speech level on listening effort outcomes. Two measurements of listening effort were utilized to assess subjective and objective listening effort. Experiment 3 examined the effect of spectral shape and level on sound quality perception. An acoustic measure was also conducted to identify factors that may contribute to sound quality perception.

CHAPTER 2

EXPERIMENT 1: EFFECT OF SPECTRAL SHAPING ON SPEECH RECOGNITION IN QUIET AND IN NOISE.

In order to restore audibility, HAs shape the spectrum of speech and increase overall level. However, speech recognition outcomes with spectrally shaped speech have been mixed. In most cases, spectral shaping appears to aid in speech recognition (e.g. Turner & Henry, 2002; Horwitz et al. 2008), however, there is also evidence of poorer speech recognition with spectral shaping, especially when high frequencies are amplified (e.g. Amos & Humes, 2007; Baer et al. 2002). Similarly, there is evidence of poorer recognition for speech at high intensity levels (e.g. Dubno, Horwitz, & Ahlstrom, 2005; Summers & Cord, 2007). Therefore, while audibility is essential for successful speech recognition (e.g. Humes, 2007), there is evidence that suggests limitations regarding high frequency amplification. However, in much of this work, the benefit observed from amplification has been in quiet environments. The benefit obtained from amplification may be more limited in noise, but this effect has not been systematically addressed. Therefore, this study examined the effects of spectral shaping in both quiet and noise backgrounds. In addition, the interaction between spectral shaping, presentation level, and background noise was examined. The purpose of this investigation was to determine the relationship between spectral shaping and speech level on speech recognition in quiet and in noise. The results of this investigation have implications clinically for how amplification techniques affect speech perception differently depending on the speech level and background conditions.

Spectral Shaping

Spectral shaping is a technique commonly used to control for audibility across normal hearing listeners and listeners with hearing impairment. It involves amplifying a signal across frequencies according to an individual's audiometric thresholds. In most cases, high frequencies are more amplified in shaped speech due to high frequency hearing loss. Providing amplification of the high frequencies according to elevated thresholds accounts for primary declines in speech recognition due to audibility (Humes, 2013). However, previous research has shown that the perception of shaped speech is not necessarily equal to that of speech with a normal spectrum (Amos and Humes, 2007; Fogerty, Montgomery, & Crass, 2014; Fogerty, Ahlstrom, Bologna, & Dubno, 2015b; Fogerty, Bologna, Ahlstrom, & Dubno, 2017). One reason this might be is that unequal amplification of high versus low frequencies alters the consonant-vowel (CV) intensity ratio, or the relative amplitude of consonants in relation to the amplitude of vowels. Previous research has suggested that reduction of the CV intensity ratio, which occurs when high frequencies are amplified, changes amplitude cues that are used for speech recognition and consequently, can have negative effects on speech recognition (Freyman, Nerbonne, & Cote, 1991; Hickson & Byrne, 1997). Further research on reduced CV intensity ratios suggested that specifically, consonant amplification can reduce consonant recognition for familiar talkers, which could be due to an acoustic mismatch of the perceived consonant and its stored acoustic representation (Fogerty et al., 2014).

The unfolding of consonants and vowels over time creates amplitude variations of the temporal envelope, which has been shown to be important for speech recognition (Shannon, Zeng, & Wygonski, 1998). Spectral shaping can reduce time-varying amplitude differences, which in turn may impair perception of temporal envelope cues. Furthermore, research suggests that spectral shaping may affect how temporal cues are weighted across frequency regions. One study examined frequency-dependent temporal cues used for speech recognition by examining envelope and temporal fine structure cues independently in three spectral frequency bands. (Fogerty & Humes, 2012). The effect of hearing status was also examined and included a young normal hearing group that was tested with spectrally shaped stimuli. In comparison to young normal hearing listeners who heard speech with normal spectral qualities, the shaped group weighted envelope cues differently across different spectral regions. This indicates that spectral shaping may influence how temporal cues are compared and integrated across frequency regions, which could contribute to differences in perception.

Another common method of examining the effect of amplification in different frequency regions is by manipulating low- and high- pass filter cutoff frequencies for amplification. The majority of such studies have demonstrated the benefit of high frequency amplification for speech recognition (Hornsby & Ricketts, 2003; 2006; Skinner, 1980; Sullivan, Allsman, Nielsen, & Mobley, 1992; Turner & Henry, 2002). Utilizing this method, Hornsby and Ricketts (2003; 2006) found for both flat and sloping sensorineural hearing losses, high frequency amplification was associated with better speech recognition than without, although the magnitude of improvement was often small. In a similar study, Skinner (1980) found that hearing-impaired listeners had greatest speech recognition when high frequencies (2-4 kHz) were 0-15 dB above lower frequency regions (500-1000 Hz) in participants with hearing impairment.

However, improvement in speech recognition due to amplification is not always consistent or similar in magnitude at an individual level. A study by Simpson, McDermott, and Dowell (2005) examined consonant identification with increasing high frequency amplification and demonstrated a significant improvement at a group level. Individual results showed improvements with increasing high frequency audibility in nine of ten participants, but the remaining participant demonstrated little to no improvement. Similarly, Horwitz et al. (2008) examined individual differences in normal hearing- and hearing-impaired participants with amplification at increasing cutoff frequencies. For normal hearing- and hearing-impaired participants, a significant improvement in speech recognition was observed with increasing cutoff frequency. However, a number of participants in both normal hearing and hearing-impaired groups did not demonstrate an increase in performance for the maximum cutoff frequency. These studies suggest that while high frequency amplification is often helpful, the magnitude of benefit can be variable at an individual level. Few studies have explored the individual differences in improvement associated with spectral shaping, and none to date have examined them in the context of speech level, or varying background noise conditions.

Speech Level

Spectral shaping involves selective amplification of frequency regions. This results in a signal that has a higher overall level than without spectral shaping. This is, of course, necessary to restore audibility in individuals with hearing impairment. However, there is a long history of reduced speech recognition at high levels (e.g., Dubno et al., 2005; French & Steinberg, 1947; Hawkins & Stevens, 1950; Pollack & Pickett, 1958; Studebaker, Sherbecoe, McDaniel, & Gwaltney, 1999; Summers & Cord, 2007). More specifically, recognition of speech usually increases from low to middle levels and decreases at higher levels. Dubno et al. (2012), for example, examined recognition of noise-vocoded consonant-vowel syllables and sentences at 45, 60, and 85 dB SPL. Speech recognition performance increased from 45 to 60 dB SPL but declined from 60 to 85 dB SPL for both syllable and sentence materials. This detriment at high levels could be due to multiple factors, including lower frequency selectivity associated with high intensities, wider auditory filters, and an upward spread of masking caused by the overamplification of low frequency regions (as discussed in Dubno et al., 2012). Effects of high speech levels alone are important to consider when examining amplified speech, as well as the possible interaction between spectral shape and speech level.

Noise and amplification

Hearing-aid benefit is limited in the presence of noise (Humes, 1991; Magnusson, Claesson, Persson, & Tengstrand., 2013; Takahashi et al., 2007) and represents one of the top complaints of HA users (Kochkin, 2000). In addition, HA users often find background noise uncomfortably loud or annoying (Kochkin, 2000). Further research on the functionality of amplification in background noise is needed to understand why this is such a prominent complaint.

In general, speech recognition in background noise depends on the following factors: the spectral overlap between the speech and noise and the number and duration of glimpses, or periods of time where there is a favorable signal-to-noise ratio (SNR) (Cooke, 2003; 2006). Speech recognition for young normal hearing listeners is typically better in background noise that is temporally fluctuating, as they are able to take advantage of speech glimpses. (e.g., Festen & Plomp, 1990; Howard-Jones & Rosen, 1993; Miller, 1947; Miller & Licklider, 1950; Takahashi & Bacon, 1992; Wilson & Carhart, 1969). However, listeners with hearing impairment often do not demonstrate this release from masking, or better performance with fluctuating noise in comparison to steady state noise (Bronkhorst & Plomp, 1992; Eisenberg, Dirks, & Bell, 1995; Festen & Plomp, 1990; Gustafson & Arlinger, 1994; Hygge, Ronnberg, Larsby, & Arlinger, 1992; Phillips, Rappaport, & Gulliver, 1994). This is likely due to reduced audibility in listeners with hearing impairment (Bacon, Opie, & Montoya, 1998), as well as a reduced dynamic range (the difference between hearing thresholds and sensation levels) that prevents taking advantage of the dips in a masker. However, further research has pointed to a temporal-resolution deficit in hearing-impaired listeners as a cause (e.g. Dubno, Horowitz, & Ahlstrom, 2003; Nelson, Schroder, & Wojtczak, 2001; Phatak & Grant, 2012;). One such study found that the benefit gained from a modulated masker was similar between hearing-impaired and young normal hearing listeners when audibility and SNR were equalized across groups. (Jensen & Bernstein, 2019).

However, there is evidence that suggests that amplification may affect recognition depending on the nature of the background noise. For example, Fogerty et al. (2015a) examined speech recognition for sentences with consonants or vowels replaced with speech modulated noise. Participants who listened to spectrally shaped stimuli demonstrated an overall decrease in performance in comparison to participants who did not receive spectral shaping. This suggests that spectral shaping may hinder the perception of speech in complex maskers. Spectral shaping reduces the modulation depth of the temporal envelope, which could make “listening in the dips” of a modulated masker more challenging. Future research is needed to compare the effects of spectral shaping and level in multiple background noise conditions to better understand how amplification functions in different listening situations.

Current study

Amplification is necessary to improve audibility in hearing-impaired listeners, however, there are consequences regarding both the speech level and the spectrum of the speech. For the majority of investigations, amplification provides benefit to the hearing-impaired listener due to improvements in audibility, although the amount of improvement is mixed for individual listeners (Horwitz et al. 2008). In some cases, modification to the speech spectrum has been detrimental to normal hearing listeners on speech recognition tasks in comparison to unprocessed speech with a normal spectrum (e.g. Amos & Humes, 2007). However, it is unclear if this detriment is due to changes in spectral shape or speech levels, or if these factors may influence the degree of improvement that a listener gains from amplification. In order to investigate this, the current study systematically varied spectral shape and speech level in a speech recognition task for young normal hearing listeners in two different background noise conditions. The young normal hearing population was tested to avoid factors related to altered suprathreshold processing due to cochlear pathology and effects of varying sensation level due to elevated hearing thresholds. Furthermore, the consequences of acoustic HA processing can be interpreted relative to normal hearing without processing. Average thresholds from a group of listeners with hearing-impairment were used as a standard audiogram to define spectral shaping parameters. Speech recognition was evaluated in three listening conditions: quiet, steady-state noise, and speech-modulated noise to assess the effect of spectral shape and speech level in common masking paradigms.

Additionally, given that some perceptual learning might occur due to spectral shaping being a novel stimulus for normal hearing listeners, pre- and post-tests were administered to determine the pattern of learning after a brief exposure. In order to better understand behavioral results, acoustic analyses were conducted on the speech-in-noise stimuli to identify potential factors that may contribute to performance. This proposal represents a direct method of comparing the combined effects of level and shaping, and how these factors function in varying background conditions.

Methods

Participants

Forty-seven young normal hearing listeners were recruited from the undergraduate and graduate University of South Carolina population. Participants were between 18-29 years of age (Mean = 22 years, 45 female) and native speakers of American English. Criteria for normal hearing consisted of thresholds less than or equal to 20 dB HL at octave frequencies from 0.25-8 kHz (ANSI, 1997). Written consent was obtained from all participants. Listeners were compensated for participating via course credit or payment. A subset of 10 listeners also completed pre-/post-tests to assess perceptual learning over the course of the experiment.

Stimuli and design

Stimuli consisted of sentences from the IEEE corpus (IEEE, 1969). These sentences are phonetically balanced with five keywords and were recorded by a male talker (Loizou, 2007). Sentences are syntactically and grammatically correct but have low contextual predictability. No sentences were repeated in any condition.

Sentences were presented in normal and shaped spectral conditions and at three speech levels. All conditions were presented in quiet, steady-state noise, and speech-modulated noise. The experimental conditions entailed a 2 (shaped/normal) x 3 (speech level) x 3 (background type) design.

Spectral conditions.

Two spectral conditions, speech with a normal spectrum, and spectrally shaped speech, were tested to examine the effect of amplification. Spectrally shaped speech was amplified according to the mean thresholds of a group of older hearing-impaired participants, while no modifications were made to the spectrum of normal speech.

Normal spectrum. (Figure 2.1, Panel A)

Stimuli for the normal spectral condition were scaled to the same root mean square (RMS). All stimuli were passed through a lowpass, linear phase, finite-impulse-response, 128th order filter with a cutoff of 5623 Hz. No further changes were made to the speech spectrum.

Shaped spectrum. (Figure 2.1, Panel B)

Stimuli were first passed through the same lowpass filter as used for the Normal spectral condition. Target speech levels for the shaped spectral conditions were based on the average thresholds of fourteen older listeners with hearing-impairment recruited for a previous experiment (See Figure 2.2). The signal was first calibrated to a maximum level of 105 dB SPL. Attenuation was applied individually to 1/3 octave bands from 100 to 8000 Hz in order to increase the band level to a target level of 70 dB SPL, or to a target sensation level of more than 20 dB above the average thresholds of the older listeners with hearing-impairment, whichever level was higher. This procedure ensured that speech levels exceeded average audiometric thresholds from the hearing-impaired listeners by 15 dB up to 4000 Hz. To ensure overall presentation levels were equal, lower frequencies were reduced and higher frequencies were increased in shaped conditions relative to normal conditions.

Speech levels.

Normal and spectrally shaped conditions were then linearly scaled to one of three speech levels: 55, 70 and 82 dB SPL. These levels corresponded to the average sensation level of the group of older listeners with hearing impairment (i.e., difference between average hearing thresholds and the level of shaped speech), the level of unshaped speech, and the overall speech level of shaped speech following amplification for older listeners with hearing impairment, respectively. A group of nineteen participants were tested on stimuli 2 dB lower (55, 68, and 80 dB SPL), but results were not significantly different between groups and therefore, were combined for analysis.

Background conditions.

Three different background noise conditions, quiet, SSN, and SMN, were tested with both normal and spectrally shaped speech. Both SSN and SMN conditions were presented at -3 dB SNR.

Quiet.

In this condition, sentence stimuli were presented in quiet, without the presence of noise.

Steady-state noise. (SSN)

Sentence stimuli were concatenated and a SSN was created that matched the long-term average power spectrum. The target speech and a random sample of the competing noise was then combined during stimulus presentation. One-hundred milliseconds of noise was present before and after each sentence. Predicted performance for normal and shaped conditions in SSN were calculated at the three speech levels using the Speech Intelligibility Index (SII; ANSI, 1997). SII values and predicted performance were calculated using the third octave band frequency importance function for “Standard Speech” (ANSI, 1997), and the Hearing in Noise Test (HINT) transfer function (Eisenberg et al., 1998). SII values of 0.5 and predicted scores of 89% correct were observed for each condition. This indicates that both speech conditions at all three levels had similar predicted performance according to the SII.

Speech-modulated noise. (SMN)

The temporal envelope of the concatenated file was extracted by half-wave rectification and passed through a low-pass fourth-order Butterworth filter with a cutoff modulation frequency of 16 Hz. This temporal envelope was then used to amplitude modulate the SSN created above. The target speech and a random sample of the SMN was then be combined during stimulus presentation.

Procedure

Testing was completed in a sound attenuating booth. Participants listened to stimuli at a sampling rate of 48,828 Hz via Sennheiser HDA 200 headphones routed through a TDT System III digital-to-analog processor (RP2) and headphone buffer (HB7). The upper cutoff frequency for all stimuli was 5623 Hz. Stimuli were presented monaurally to the right ear. Participants were always presented with the quiet condition first, followed by the two background noise conditions. The order of presentation for the two noise conditions was counterbalanced across participants. Stimuli were presented in blocks of ten sentences according to the six spectral/level combinations within each condition (2 spectral x 3 speech levels). Sentences were presented in a fixed order; however, block order was randomized. A demonstration trial of ten sentences was presented at the beginning of each testing condition to familiarize the participant with the procedure and stimulus processing. Each background condition consisted of 90 sentences. Two additional conditions were assessed, but the current analysis focused on only six of the eight total conditions tested. Participants were required to repeat each sentence and encouraged to guess. Sentence presentation was self-paced on the computer. Listener responses were scored online as well as audio recorded. A response was scored as accurate if the participant repeated each word exactly (e.g. without missing or extra phonemes).

Perceptual Learning

A subset of 10 participants were administered a pre- and post-test in conjunction with speech recognition testing in order to assess perceptual learning of shaped and normal conditions over the duration of the testing session. The Ease of Language Understanding model (ELU; Ronnberg, 2003; Ronnberg et al., 2013) predicts that the acoustic similarity between a stimulus and its stored mental representation affects lexical processing. Due to the acoustic changes that occur when speech in amplified, this mismatch between heard and stored representation could be larger for shaped than normal speech. In this case, it is possible that participant’s lexical processing of shaped speech could improve as they are exposed to the novel speech throughout the testing session. Because there is not a mismatch between normal speech and its stored representation, a smaller effect of perceptual learning would be expected relative to that of shaped speech.

The perceptual learning test consisted of IEEE sentences presented in SSN and SMN backgrounds at -3 dB SNR. Only normal stimuli at 70 dB SPL and shaped stimuli at 82 dB SPL were assessed, which are the most relevant conditions ecologically (as amplified speech typical of a hearing aid results in the combined factors of elevated speech level and altered spectral shape). Four blocks of ten sentences were presented in each test, two blocks with SSN and two blocks with SMN. The order of noise condition as well as normal and shaped speech blocks within a noise condition were counterbalanced across participants. A demo block of four sentences was presented before each noise condition. Responses were scored by percentage of keywords correctly repeated.

Calibration

Two calibration noise files were created that matched the long-term average spectrum of the concatenated sentence materials for the normal and shaped stimuli. One-third octave band levels were measured for each of the three calibration noises at the output of the earphone using a Larson-Davis 800B sound level with linear weighting.

Results

Speech Recognition

Participant keyword recognition for shaped and normal stimuli in each of the three background conditions are displayed in Figure 2.3. All proportion keyword correct data was transformed using a rationalized arcsine procedure to stabilize the error variance

(Studebaker, 1985). In general, scores were highest within the quiet background condition, followed by SMN and SSN. Results were analyzed using a 2 (Spectral condition: normal or shaped) x 3 (Level: 55, 70, and 82 dB SPL) repeated measures analysis of variance for each background type. Pairwise comparisons were conducted for each background type to compare the effect of shaping at three speech levels and the effect of the two different spectral shapes (normal, shaped). All pairwise comparisons were conducted using a Bonferroni-Holms adjusted alpha level for multiple comparisons.

Quiet.

Keyword recognition in the quiet background condition is displayed in Figure 2.3, panel A for normal and shaped stimuli at three overall levels. A significant main effect of level was observed [F (1,46) = 10.4, p < .001] as well as a shape by level interaction [F (2,92) = 10.4, p < .001]. The main effect of shape was not significant (p > .05). Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels. For 55 dB SPL, performance was better with shaped speech than normal speech (p < .05). For 82 dB SPL, performance was better with normal speech than shaped speech (p < .05). No significant difference was observed between normal and shaped at 70 dB SPL (p>.05). In quiet, it appears that shaped speech yields better performance at lower overall presentation levels, while normal speech yields better performance at greater overall levels. T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. In general, performance was better as overall level increased, with the exception of shaped speech at 55 dB SPL. Performance with normal speech at 70 dB SPL and 82 dB SPL was significantly better than normal speech at 55 dB SPL (p < .05).

SMN.

Keyword recognition in the SMN background condition is displayed in Figure 2.3, panel B for normal and shaped stimuli at three overall levels. Significant main effects of shape [F (1,46) = 207.2, $p < .001$] and level [F (1,46) = 47.3, $p < .001$] were observed as well as a shape by level interaction [F (2,92) = 28.0, $p < .001$]. Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels. For 55, 70, and 82 dB SPL performance was better with shaped speech than normal speech ($p < .05$). In SMN, it appears that shaped speech yields better recognition regardless of level. T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. In general, recognition was best at the middle presentation level of 70 dB SPL. For both shaped and normal speech, performance at 70 dB SPL was significantly better than at 55 dB SPL ($p < .05$) or 82 dB SPL ($p < .05$).

SSN.

Keyword recognition in the SSN background condition is displayed in Figure 2.3, panel C for normal and shaped stimuli at three overall levels. Significant main effects of shape [F (1,46) = 1416.8, $p < .001$] and level [F (1,46) = 83.0, $p < .001$] were observed as well as a shape by level interaction [F (2,92) = 43.0, $p < .001$]. The main effect of level was not significant ($p > .05$). Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels. For 55, 70, and 82 dB SPL performance was better with shaped speech than normal speech ($p < .05$). Similar to the SMN background condition, it appears that shaped speech yields better recognition regardless of level. However, SII scores were the same across spectral conditions. T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. In general, recognition was better at lower levels with normal speech, and better at middle levels for shaped speech. For shaped speech, performance at 70 dB SPL was significantly better than at 55 dB SPL ($p < .05$) or 82 dB SPL ($p < .05$). For normal speech, performance at 55 dB SPL was significantly better than at 70 dB SPL ($p < .05$) or 82 dB SPL ($p < .05$).

Perceptual Learning

Keyword recognition for pre- and post- tests are displayed for both background conditions in Figure 2.4. In addition, difference scores were calculated between pre- and post- tests for both speech tests in the two background conditions (Figure 2.5). Results were analyzed using a 2 (Spectral Shape: normal or shaped) x 2 (Time: pre-/post-) repeated -measures analysis of variance for each background type. Test order was initially examined, but not found to be significant ($p > .05$), therefore groups were collapsed for this analysis. All pairwise comparisons were conducted using a Bonferroni-Holms adjusted alpha level for multiple comparisons.

SMN.

Keyword recognition for pre- and post- tests in the SMN background condition are displayed in Figure 2.4. A significant main effect of shape was observed [$F(1,9) = 10.1$, $p < .05$] as well as a shape by time interaction [$F(2,18) = 9.3$, $p < .05$]. Paired t-test comparisons were conducted to examine differences in performance between pre- and post- tests for the two speech conditions. Pre-test performance was significantly better than post-test performance for normal speech ($p < .05$). Pre- and post- test performance was not significantly different for shaped speech ($p > .05$). These results provide no evidence of perceptual learning in SMN for either normal or shaped speech.

SSN.

Keyword recognition for pre- and post- tests in the SSN background condition is displayed in Figure 2.4. Significant main effects of shape [F (1,9) = 18.1, $p < .01$] and time [F (1,9) = 27.2, $p < .01$] were observed as well as a shape by time interaction [F (2,18) = 5.9, $p < .05$]. Paired t-test comparisons were conducted to examine differences in performance between pre- and post- tests for the two speech conditions. Post- test performance was significantly better than pre-test performance for normal and shaped speech ($p <.05$). This suggests that perceptual learning may have occurred with both normal and shaped speech in the SSN background condition.

Difference scores.

Difference scores were calculated to quantify perceptual learning over the course testing for both normal and shaped speech in both background noise conditions. These difference scores are displayed in Figure 2.5. Paired t-test comparisons were conducted to compare the change in performance between pre- and post- tests for the two speech conditions. For both the SMN and SSN background conditions, differences between pre- and post- test performance were larger for shaped speech than normal speech. This suggests that greater perceptual learning may have occurred with shaped speech than normal speech throughout the course of the testing session.

Individual Differences in Speech Recognition

Previous research has suggested that speech recognition with amplification differs on an individual level (e.g. Horwitz et al. 2012). Scores for individual participants were examined to determine individual differences in response patterns with and without spectral shaping. Raw scores were compared between normal speech at 70 dB SPL and shaped speech at 82 dB SPL. These conditions were chosen for comparison because they represent the most ecologically valid listening conditions (unamplified speech and speech processed through a hearing aid, respectively) and are the most important clinically. In the quiet condition, 14 of the 47 participants (30%) had better scores with shaped speech, while four had better scores with normal speech (9%). The remaining 29 participants had the same raw score in both conditions, likely due to ceiling effects. Differences between conditions ranged from two to ten percentage-points, with a mean difference of 3.5 percentage-points. In SMN, 33 of the 47 participants tested had better scores with shaped speech (70%), while ten had better scores accuracy with normal speech (21%). The remaining four participants performed the same in both conditions. Differences between conditions ranged from 2-28 percentage-points, with a mean difference of 10 percentage-points. In SSN, all 47 participants had better scores with shaped speech than normal speech.

Acoustic Analysis

Temporal Envelope Analysis

The spectral shaping employed in this study primarily resulted in amplification of high frequencies (above 2 kHz). Because of this, portions of a given sentence that contain more high-frequency components received greater amplification than portions of the sentence that contained more low-frequency components. To explore this, an acoustic analysis was performed to demonstrate the difference in intensity contours between normal and shaped conditions. Intensity contours were calculated in Praat (Boersma & Weenink, 2017) for an example IEEE sentence processed via shaped and normal stimulus parameters. Results of this analysis are shown in Figure 2.6. As displayed in the figure, some portions of the sentence are at a higher intensity for the shaped condition. These portions contained a greater concentration of high frequencies, which were amplified in the shaped condition. These are segments of the sentence where, in noise, glimpsing opportunities may be greater for shaped conditions due to a more favorable SNR. This analysis demonstrates two consequences of spectral shaping: altered glimpsing opportunities and altered temporal envelope. In what follows, two acoustic metrics: HASPI (accounting for the altered temporal envelope) and a glimpsing model will be utilized to investigate the effect of these two altered aspects of the signal due to spectral shaping on keyword recognition.

Hearing Aid Speech Perception Index (HASPI)

The Hearing Aid Speech Perception Index (HASPI; Kates and Arehart, 2014a) was used to interpret behavioral recognition performance. This specific metric was chosen to explain the changes in the temporal envelope as a result of spectral shaping. This metric utilized a comparative acoustic analysis, where the output of an auditory model for an unprocessed signal (IEEE sentences) was compared to the output of an auditory model for a processed signal (amplified sentences). The metric combined a coherence measure, which examined temporal fine structure changes between a clean and noise signal, and envelope correlation, which examined changes to the signal envelope. In the following analysis the unprocessed IEEE sentence was compared to the same IEEE sentence processed in both normal and shaped conditions in SMN and SSN. An average was computed for each condition across all sentences tested. The HASPI output is constrained to be between 0 and 1, with 0 indicating poor predicted intelligibility and 1 indicating perfect predicted intelligibility. Normal and shaped conditions were analyzed in comparison to their relative clean signals.

HASPI Results.

Average HASPI scores for SSN and SMN background conditions for both speech types are displayed in Figure 2.7. Results were analyzed using a 2 (Spectral condition: normal or shaped) x 3 (Level: 55, 70, and 82 dB SPL) repeated-measures analysis of variance. In SMN, significant main effects of shape [$F(1,9) = 49.3$, $p < .001$] and level [$F(1,9) = 19.4$, $p < .001$] were observed, however the shape by level interaction was not significant ($p > .05$). In SSN, significant main effects of shape [$F(1,9) = 117.8$, $p < .001$] and level [$F(1,9) = 121.3$, $p < .001$] were observed, as well as a shape by level interaction [$F(1,9) = 41.0$, $p < .001$]. Paired t-test comparisons were conducted to examine the effect of the spectral condition on HASPI scores at three overall levels in both SMN and SSN. Tests were conducted using a Bonferroni-Holms adjusted alpha level for multiple comparisons. For 55, 70, and 82 dB SPL, HASPI scores were significantly better for shaped speech than normal speech ($p < .05$) in both SMN and SSN, indicating that shaped speech was predicted to be more intelligible than normal speech in both noise conditions. This follows the same pattern as speech recognition, where recognition of shaped speech was consistently higher than normal speech. Regarding level, 55 dB SPL had significantly better HASPI scores than at 70 dB SPL and 82 dB SPL in both shaped and normal conditions ($p < .05$), for both SMN and SSN. This contrasts with observed recognition results, where 70 dB SPL was associated with highest levels of performance for shaped and normal speech in both noise conditions. In summary, shaped speech had significantly better HASPI scores in SMN and SSN than in normal speech. This is perhaps due to better preservation of the amplified speech during glimpsed portions of the signal, which the next acoustic analysis assesses.

Correlation with Speech Recognition.

Average HASPI scores were correlated with average speech recognition scores across background conditions and level conditions to determine the relationship between the acoustic analysis and observed performance (See Figure 2.8). The two variables were positively correlated using Pearson's product moment correlations ($r = 0.88$, $p < 0.5$) indicating that HASPI scores captured 77% of the variance in speech recognition observed across conditions. However, scores predicted by HASPI were not consistent with observed speech recognition scores in every case.

Glimpse Proportion and Glimpse SNR

When speech is in the presence of a fluctuating masker, the amplitude modulation of the speech and masker provide the listener with intermittent "dips" where target speech can be perceived at a more favorable signal-to-noise ratio (SNR). Current methods used to model speech understanding in the presence of modulated maskers are based on predicting intelligibility on the basis of these "glimpsed" segments (Cooke, 2003, 2006; Li & Loizou, 2007). The purpose of the glimpse proportion and glimpse SNR analyses was to quantify the proportion of time-frequency units that were glimpsed in a sentence, or at a favorable SNR. The average SNR of glimpsed time-frequency units were also calculated. Stimuli were analyzed in both SMN and SSN conditions. Signal processing strategies for ideal-binary masked (IBM) speech were utilized. (e.g. Li & Loizou, 2007). The IBM was calculated from the combined speech and noise signal. Glimpse proportion was defined as the number of time-frequency units labeled by the ideal binary mask as 1, divided by the total number of time-frequency units across the sentence. A fast Fourier transform based short-time spectral analysis-modification-synthesis was utilized with a frame duration of 32 ms and a frame shift of 4 ms. Local SNR criterion was defined as -5 dB SNR. The proportion of glimpsed time-frequency units, as well as the SNR, was averaged for each sentence.

Proportion of Glimpses.

The proportion of the sentence that was glimpsed above a -5 dB SNR was calculated for each sentence across shaped and normal conditions. This analysis therefore identified speech conditions that preserved a higher proportion of glimpsed time-frequency units, and to determine the extent to which this metric might be related to better performance. The average glimpse proportions are displayed for natural and shaped speech in both background noise conditions in Figure 9, panel A. Paired t-test comparisons were conducted to compare the proportion of glimpsed segments for the two spectral conditions. Tests were conducted using a Bonferroni-Holms adjusted alpha level for multiple comparisons. Spectrally shaping speech significantly increases the proportion of speech glimpsed in SSN and SMN relative to natural speech ($p < .05$). This may explain why speech recognition performance was consistently higher for spectrally shaped speech in noise, due to more glimpsing opportunities relative to normal speech.

Glimpse SNR.

The SNR of each glimpsed time-frequency unit was calculated across sentence stimuli to examine the availability of target speech information in each noise condition with normal and shaped stimuli. The average SNRs are displayed for both speech and background conditions in Figure 2.9, panel B. Paired t-test comparisons were conducted to compare the average SNR of glimpsed segments for the two spectral conditions. Tests were conducted using a Bonferroni-Holms adjusted alpha level for multiple comparisons. For both the SMN and SSN background conditions, the shaped speech had significantly better SNRs in comparison to normal speech ($p < .05$). Therefore, spectrally shaping speech not only increased the proportion of the sentence glimpsed at a favorable SNR (above -5 dB), but the glimpses themselves were preserved at a better local SNR relative to those available for the unshaped, natural speech.

Discussion

The current study examined how spectral shaping and speech level interact as a function of background condition. While amplification is necessary to restore audibility for hearing-impaired listeners, factors related to speech level and background noise may interact to determine the benefit from amplification. These factors could underlie the lack of perceived benefit and low adherence reported by hearing aid users (NIDCD, 2010). The effects of spectral shaping, speech level, and noise background are discussed, as well as individual differences within these conditions.

The effect of spectral shaping

Spectrally shaped speech was associated with higher speech recognition scores than normal speech in all conditions except in the quiet background condition. However, participants had normal hearing, so audibility was sufficient. If audibility was limited (as with the hearing-impaired population) it is expected that spectrally shaped speech would also have higher speech recognition in the quiet background condition. The benefits of amplification on speech recognition have been well documented in previous literature, especially in background noise (e.g. Sullivan et al., 1992; Horwitz et al., 2008; Turner &

Henry, 2002). In this study, however, the effect of spectral shaping varied depending on the background condition. Spectrally shaped speech was especially beneficial in SMN and SSN conditions in comparison to normal speech. This result could be due, in part, to acoustic factors associated with spectrally shaped speech. As demonstrated by the acoustic analyses included in this study, high frequency amplification provides more frequent glimpse opportunities, and more favorable SNRs during these glimpses. This allows the listener to potentially access more acoustic cues of the target sentence, resulting in better speech recognition in background noise than when speech is not spectrally shaped. Similarly, acoustic analyses demonstrated that spectrally shaped speech had higher predicted intelligibility in noise in comparison to normal speech. A further explanation for the benefit of spectrally shaped speech could be related to perceptual learning. Because spectrally shaped speech is a novel stimulus to normal hearing listeners, it is feasible that learning could occur across the experimental session. Indeed, it appears that perceptual learning did occur in this study, especially with spectrally shaped speech in SSN as evidenced by pre-/post-test data. Perceptual learning may be another explanation for the benefit of shaping in noise background conditions. Participants improved in performance over time with shaped speech, while there was no evidence of perceptual learning with normal speech conditions.

While spectrally shaped conditions yielded better recognition than normal speech across noise backgrounds, in quiet, scores for normal speech were higher than shaped speech at 70 and 82 dB SPL. This finding could be due to the novel nature of the shaped speech: participants were always tested in the quiet condition first, and while perceptual learning may have occurred over time, there may have been initial detriments in performance with the unfamiliar stimulus. In addition, acoustic analyses demonstrated that benefits from spectrally shaped speech were related to glimpsing speech in noise, which isn’t a factor in quiet backgrounds. Indeed, the largest differences between normal and spectrally shaped speech were within noise, rather than quiet, conditions, though performance was compressed near ceiling in quiet conditions. However, the lack of benefit of shaped speech in quiet is important to note, as hearing aid fitting involves speech recognition in quiet, and amplification is utilized in quiet as well as noise conditions. Further research is needed to explore the effect of spectral shaping in quiet, as well as perceptual learning that may or may not be associated with this condition.

The effect of speech level

The effect of speech level on speech recognition varied depending on the background condition. This finding is similar to Studebaker et al. (1999), who suggested that increasing level had different effects on speech recognition depending on whether materials were presented in quiet or in noise. In the current study, an increase in performance was observed as speech level increased for quiet conditions, except for shaped speech at 55 dB SPL. Previous research has demonstrated increased or stationary performance with increasing overall levels in quiet, as audibility is greater as levels increase (e.g. Perez-Gonzalez & Lopez-Poveda, 2013; Studebaker et al., 1999). This pattern was not consistent across all background conditions, however. In SMN highest performance was observed at 70 dB SPL for both normal and spectrally shaped speech. This is a similar performance function as described in Dubno, et al. (2012), where performance was best at mid-levels, but decreased at high levels in broadband noise. Detriments at low speech levels are likely attributable to lower audibility. At higher speech levels, decreases in performance could be due to lower frequency selectivity at high intensities, wider auditory filters, as well as upward spread of masking from low frequency amplification (as discussed in Dubno et al., 2012). In the current study, the effect of performance at high levels begins at 70 dB SPL, where this decline begins at 60 dB in Dubno et al. (2012). This could be due to the difference in signal processing and SNR between the two studies. The effects of speech level were not consistent across the two background noise conditions, however. In SSN, two different patterns of performance were observed, depending on the spectral shape. For normal speech, performance decreased with increasing level. This is in contrast with SII results, which did not differ as level increased. While it is unclear why this pattern is different than that found for normal speech in SMN, we may speculate that it could be related to floor effects in performance. In other words, performance was so poor for mid-and high- levels with normal speech in SSN that listeners were unable to benefit from increasing level. In contrast, performance with spectrally shaped speech in SSN followed a similar pattern as in SMN, with highest performance observed at 70 dB SPL. As in SMN, it appears that better performance is associated with mid-levels for spectrally shaped speech.

Individual Differences

Individual differences were examined to determine if spectral shaping was associated with improvement for all individuals, and to explore the magnitude of improvement gained from spectrally shaped speech. The majority of participants benefited from spectral shaping on an individual level. However, not every individual experienced a benefit in all three background conditions, and the amount of benefit varied for each individual. In the quiet condition, for a few listeners the benefit of spectral shaping was observed, but the majority of listeners had equivalent percent correct for both speech conditions. Participants who did have better or worse performance with spectrally shaped speech in quiet experienced relatively small differences of 2-10 percentage points. This could be due to ceiling effects, which may have prevented variability in benefit between the two conditions. In SMN, spectrally shaped conditions were associated with better speech recognition at the group level. However, at an individual level, 30% of the participants performed the same or worse with spectral shaping. Unlike in the quiet background condition, the magnitude of difference varied widely, with one participant performing as much as 28 percentage points worse with spectrally shaped speech than normal speech. This is an important finding, as spectral shaping was beneficial at a group level, but for several individuals it was detrimental for speech recognition performance. In contrast to speech recognition in quiet and in SMN, all participants had better performance with shaped speech in SSN.

It appears that, depending on the background condition, there are differences on an individual level in terms of the benefit of spectral shaping. This finding is consistent with previous high frequency amplification studies that examined individual differences in recognition with hearing impaired participants (e.g. Horwitz et al., 2008, Hogan & Turner, 1998). Based on data from the current study and others, it appears that amplification may have varying amounts of benefit depending on the individual listener. It is possible that these individual differences could be driving the mixed findings of the benefit of amplification across the literature (Humes, 2013). More importantly, these findings could have clinical implications regarding variations in individual benefit associated with hearing aid use.

Summary and Conclusions

Amplification is provided in order to restore audibility to those with hearing loss. Typical sloping high frequency hearing loss associated with aging results in loss of audibility primarily in the high frequencies. High frequency amplification, such as the spectral shaping implemented here, intensifies high frequencies while preserving the natural intensity of low frequency sounds where audibility is already available. At a group level, results demonstrated that, for normal hearing listeners, amplification is beneficial for the majority of tested speech levels and background conditions. However, at high speech levels in quiet, spectral shaping was associated with lower speech recognition. At an individual level, speech recognition performance was not consistent, as some participants had similar or worse performance with spectrally shaped speech in comparison to normal speech, depending on the background condition. The magnitude of benefit associated with spectral shaping was contingent on both overall speech level and background noise condition. These findings have several important implications clinically. First, it is possible that amplification may not be as beneficial in quiet as in noise for high speech levels. However, participants in this study had normal hearing, and results may be different for participants with hearing impairment. In addition, overall speech levels should be considered when utilizing amplification techniques, as performance varied across background conditions as level increased. Finally, individual performance needs to be evaluated in varying background conditions, as the benefit of amplification was not always consistent across participants. The effects of spectral shape and speech level in different noise conditions needs to be further examined in the hearing-impaired population. The results from the current study indicate that spectral shaping is beneficial to speech recognition in normal hearing participants, but that individual benefit depends on background condition.

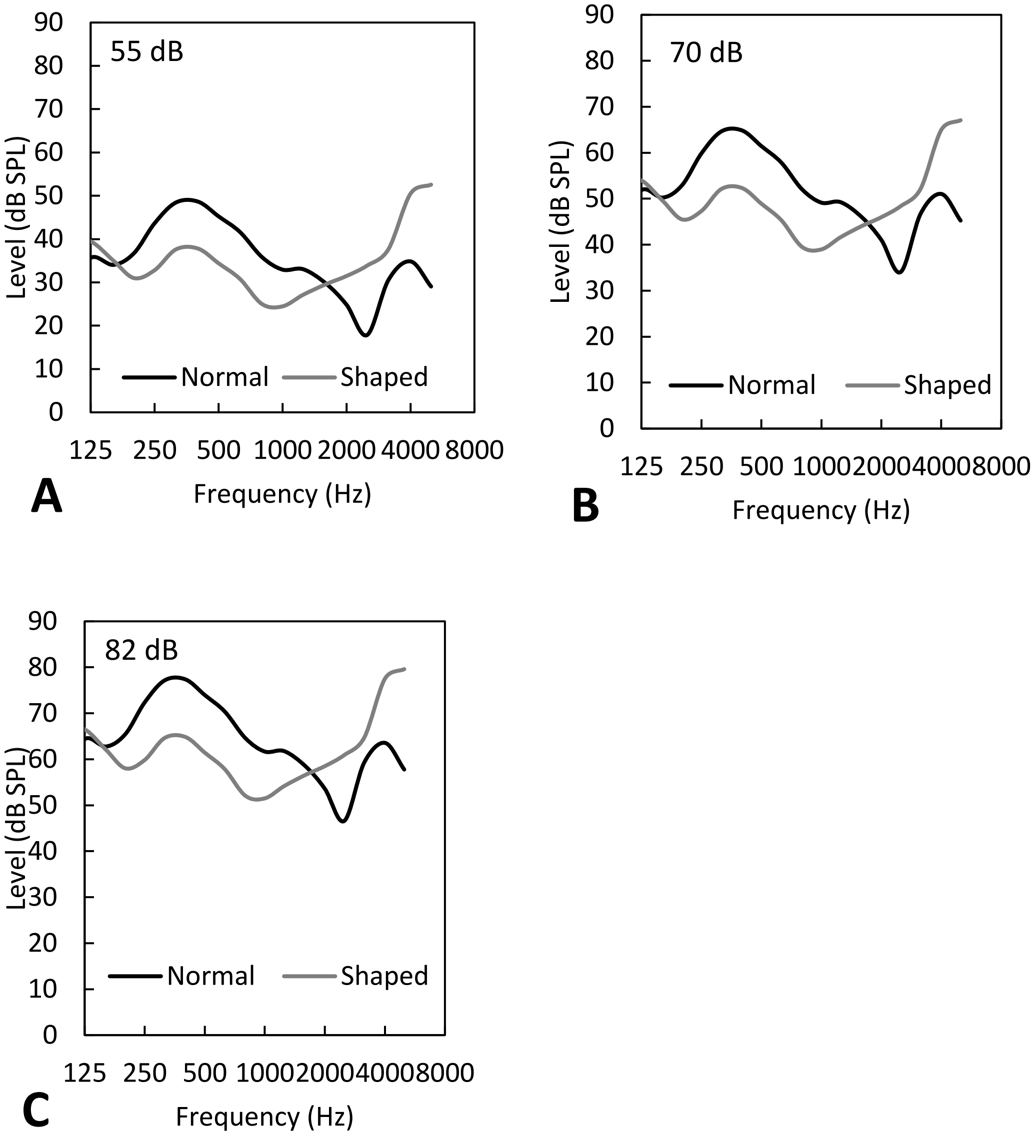

Figure 2.1. Spectral shaping conditions. Displays spectrally shaped and normal spectrums at 55 (Panel A), 70, (Panel B) and 82 (Panel C) dB SPL.

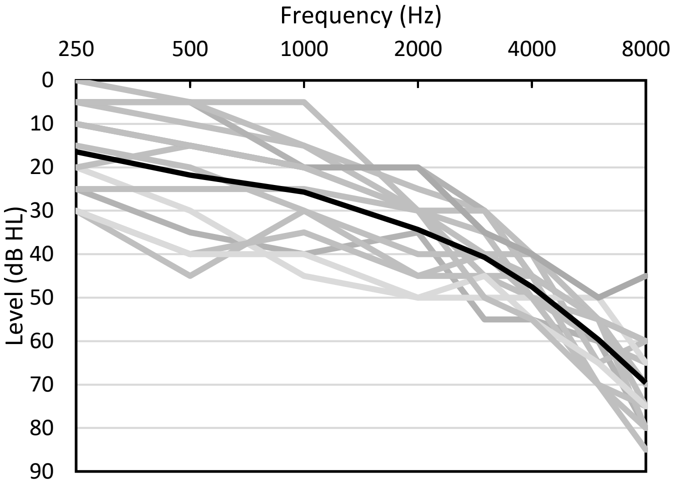

Figure 2.2. Audiometric thresholds of 14 older listeners with hearing impairment. Black line indicates the average thresholds that were used to define the spectral shaping implemented in this study for young normal hearing listeners.

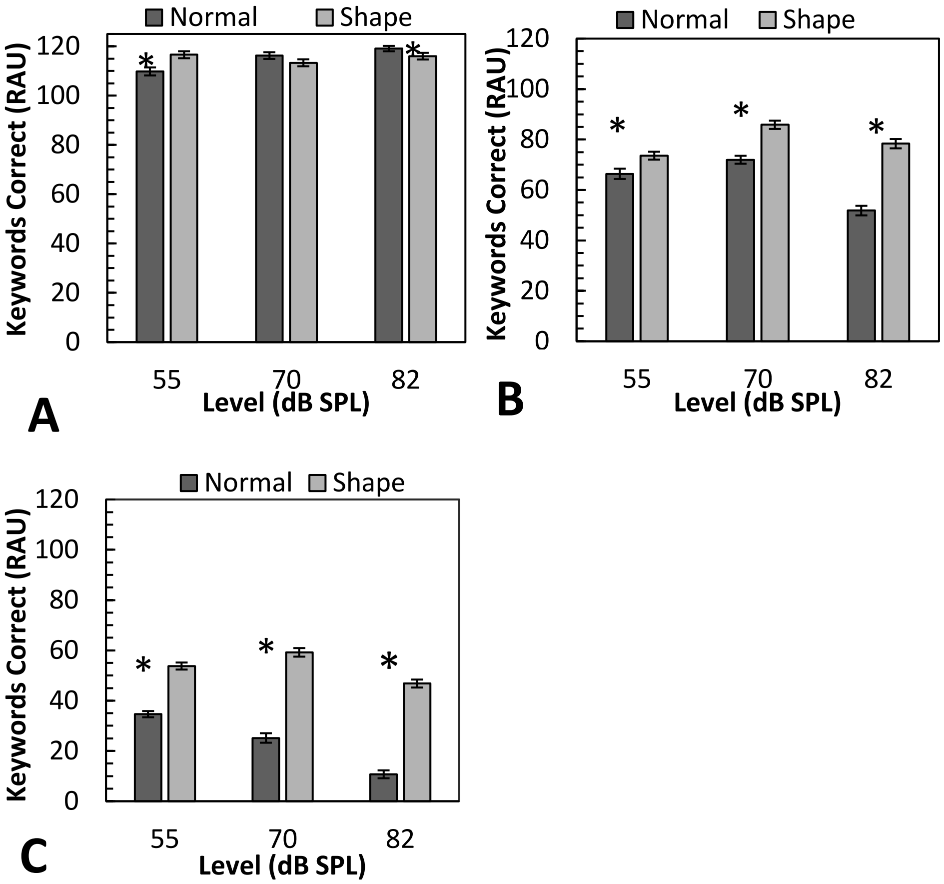

Figure 2.3. Keyword recognition accuracy for normal and shaped speech at three different levels. Shown in quiet (Panel A), SMN (Panel B), and SSN (Panel C). Error bars = standard error of the mean. Asterisks indicate significantly different performance between the two speech types (p<.05).

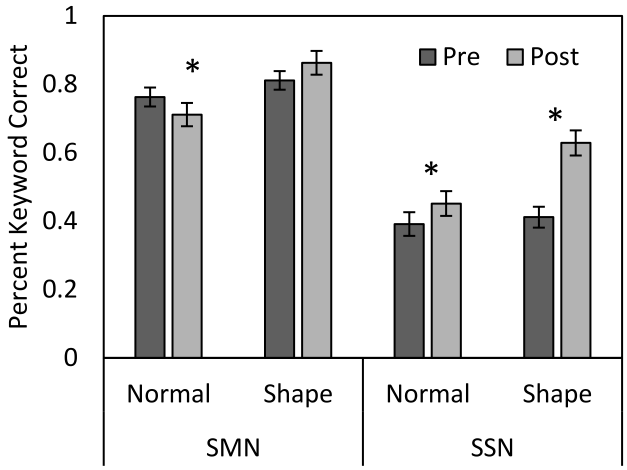

Figure 2.4. Keyword recognition accuracy for pre- and post- tests. Shown for SMN and SSN background conditions. Error bars = standard error of the mean. Asterisks indicate significantly different performance between the pre- and post- tests.

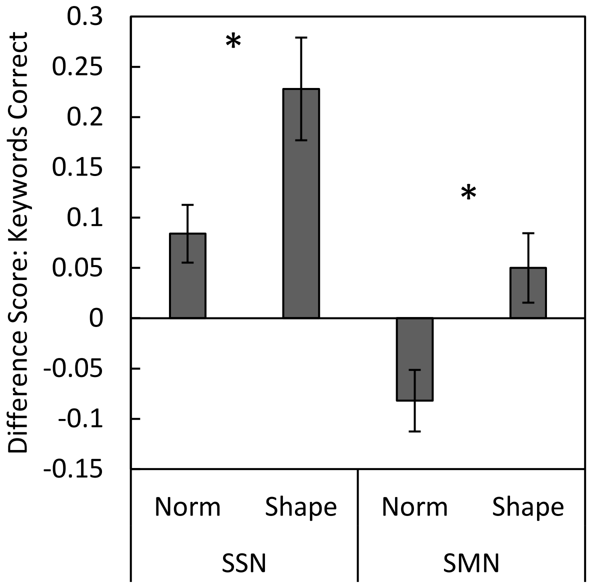

Figure 2.5. Difference scores for Pre-/Post-test. Shown in SMN and SSN background conditions. Error bars = standard error of the mean. Asterisks indicate significantly different performance between the two speech types.

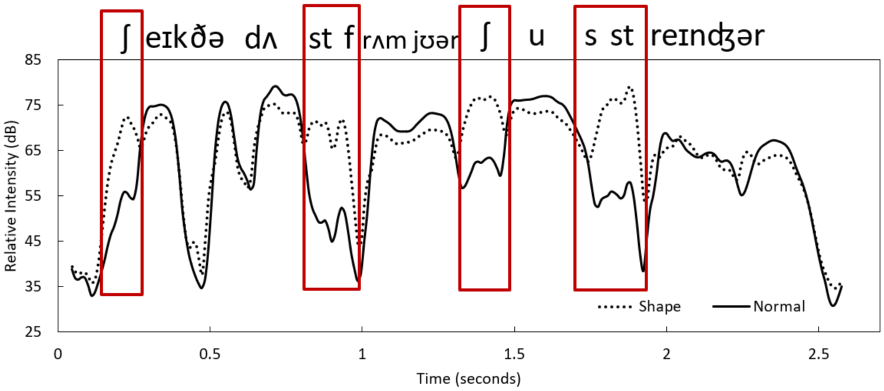

Figure 2.6. Intensity contour analysis for shaped and normal conditions for one IEEE sentence (“Shake the dust from your shoes, stranger.”). Each line displays the intensity contour for the same sentence processed according to the two experimental conditions..

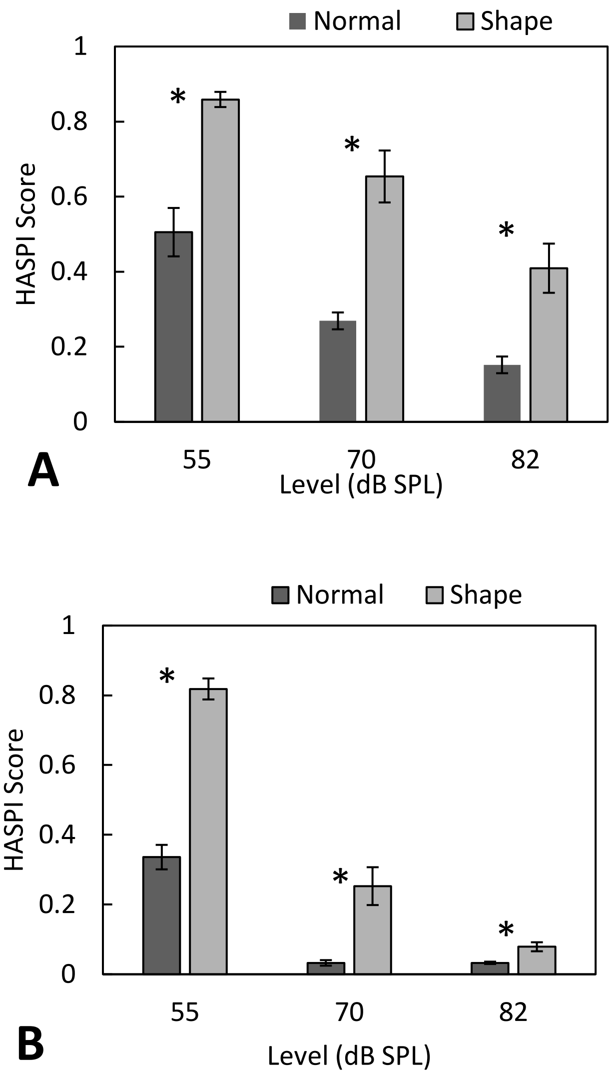

Figure 2.7. HASPI Scores for SMN (Panel A) and SSN (Panel B) background conditions. Average scores for shaped and normal speech conditions are displayed at 55, 70, and 82 dB SPL. Error bars = standard error of the mean. Asterisks indicate significantly different scores between the two speech types.

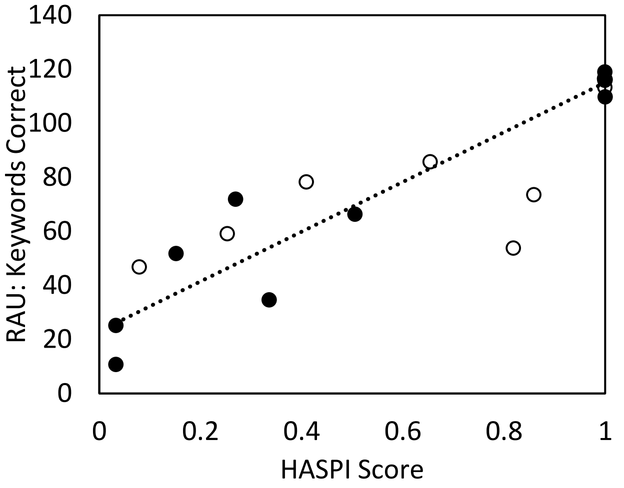

Figure 2.8. Scatterplot of the association between HASPI scores and speech recognition performance. Shown across normal (filled circles) and shaped (unfilled circles) conditions in quiet, SMN and SSN backgrounds.

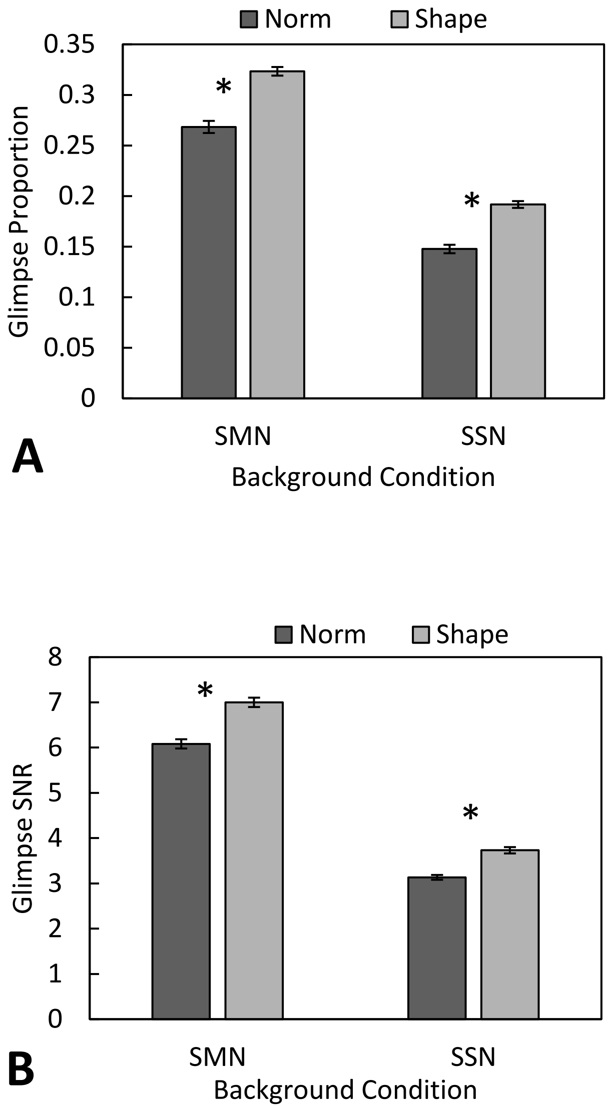

Figure 2.9. Glimpse Analysis. Panel A = Average proportion of glimpsed time-frequency units across a sentence for normal and shaped speech within SMN and SSN background conditions. Panel B = Average SNR for each glimpsed segment in SMN and SSN background conditions for normal and shaped speech. Error bars = standard error of the mean. Asterisks indicate significant difference between the two speech types.

CHAPTER 3

EXPERIMENT 2: EFFORT ASSOCIATED WITH LISTENING TO SPECTRALLY SHAPED SPEECH IN QUIET AND IN NOISE.

Introduction

A common complaint of hearing aid users is that speech is audible, but that additional effort is required to understand speech, particularly in noisy backgrounds. In order to restore audibility, hearing aids (HA) shape the spectrum of the speech and increase overall speech level. Spectral shaping alters the acoustics of the speech signal, which could create a mismatch between heard and stored mental representation (Ronnberg et al., 2017). Because of this mismatch, greater perceptual processing demands may be needed to understand spectrally shaped speech. Listening effort outcomes could reveal the perceptual processing demands required in order to understand spectrally shaped speech (Pichora-Fuller et al., 2016). Alternatively, it is possible that spectral shaping may decrease processing demands, and therefore lower listening effort, due to increased signal audibility. In general, HA amplification is associated with decreased listening effort in quiet and in noise (e.g. Ahlstrom, Horwitz, & Dubno, 2014; Bentler, Wu, Kettel, & Hurtig, 2008; Gatehouse & Gordon, 1990). However, perceived benefit associated with HA use, especially in noise, is low (McCormack & Fortnum, 2013). There are at least three potential explanations for lack of perceived benefit with HA use. First, when speech is spectrally shaped, high frequencies are often more amplified than low frequencies due to high frequency hearing loss. Listening to speech with an altered spectrum could be associated with an increase in listening effort, with may be related to perceived benefit associated with HA use. Second, when the spectrum of speech is shaped in order to restore audibility, the overall intensity level increases. There is evidence of poorer recognition for speech at high intensity levels (e.g. Dubno et al., 2005; Summers & Cord, 2007), but the effect of speech at high levels on listening effort has yet to be explored. Third, there has been limited research examining listening effort associated with spectral shaping in both quiet and noise environments. Listening effort associated with varying background noise could be related to perceived benefit experienced by HA users.

The purpose of the present study was to determine how manipulating spectral shape and speech level affects listening effort in quiet and noise. The results of this study have clinical implications for how HA amplification affects the amount of listening effort experienced by the user depending on speech level and background condition.

Listening Effort

Listening effort refers to ‘the amount of processing resources (perceptual, attentional, cognitive, etc.) allocated to a specific auditory task, when task demands are high (adverse listening conditions) and when the listener strives to reach a high-level of performance on the listening task (Gagne, Besser, & Lemke, 2017)’. Fatigue, or a feeling of depleted cognitive resources and decreased focus, is thought to be the result of expending high amounts of listening effort. Listening effort and fatigue do not have to be associated with accuracy for a task, and in fact, often are not. The Framework for Understanding Effortful Listening (FUEL, Pichora-Fuller et al., 2016) is a theoretical explanation for why listening can be effortful or fatiguing in certain conditions. FUEL suggests that there is a limited capacity of resources available when engaged in listening, a concept adapted from Kahneman's model of attentional capacity (Kahneman, 1973). Generally, the more difficult the task, the larger the amount of capacity is required to successfully listen to and understand speech. The more capacity that a task requires, the more effortful it becomes.

Listening effort has been measured in a variety of ways, including self-report and behavioral measures. Self-report measures usually consist of a rating scale or questionnaire (McGarrigle et al., 2014). These types of measures are useful in that they are quick and easy to administer and provide insight into the subjective experiences of the listener. These subjective experiences may be what guide choices regarding HA use, such as whether or not HAs are utilized in certain listening conditions. Behavioral measures, on the other hand, provide objective indicators of listening effort. These measures include a verbal or motor response (such as pressing a button; McGarrigle et al., 2014). One type of objective measure for listening effort is the participant's response time (RT) to a speech stimulus, which indicates the rate of speech processing (Gatehouse & Gordon, 1990; Houben, Doorn-Bierman, & Dreschler, 2013). Indeed, previous research has demonstrated that RT slows in more difficult listening situations, such as at poor signal-to-noise ratios (SNR; Houben et al., 2013). The current study includes an analysis of RTs and listening effort ratings in order to better understand listening effort associated with amplification. These estimates of effort were obtained during Experiment 1, which examined the accuracy of speech recognition.

Listening effort and HAs

Spectral shaping is a common technique to restore audibility in listeners with hearing impairment. In general, spectral shaping has been associated with less listening effort than in unaided conditions as indicated by RTs (Gatehouse & Gordon, 1990; Hornsby, 2013; Kulkarni, Pandey, & Jangamashetti, 2012; Picou, Ricketts, & Hornsby, 2013; Sarampalis, Kaluri, Edwards, & Hafter, 2009), subjective ratings (Ahlstrom et al., 2014), and dual task cost (Gustafson, McCreery, Hoover, Kopun, & Stelmachowicz, 2014; Pals, Sarampalis, & Baskent, 2013). In one such study, a dual task design was used to compare RTs to a visual stimulus during a speech recognition task for aided and unaided conditions in noise (Hornsby, 2013). Results indicated that RTs were faster in aided conditions, indicating that amplification reduced listening effort. Furthermore, RTs increased over time in unaided conditions, while they remained stable in aided conditions, suggesting that aided listening may also reduce fatigue. According to FUEL, it is possible that in these cases, providing amplification reduces the cognitive resources required for speech recognition. Consequently, compared to unaided conditions, restoring audibility reduces listening effort.

While hearing aids consistently improve speech recognition outcomes in quiet (Humes, 2013), and appear to mediate listening effort (Gatehouse & Gordon, 1990; Gustafson et al., 2014), only one in five individuals who could benefit from a HA actually uses one (NIDCD, 2010), and historically, regular HA use has been found to be low (Popelka et al., 1998; Upfold & Wilson, 1980; Weiss, 1973). Indeed, the one of the top reported reasons HA adherence is low is a lack of perceived benefit (Bertoli et al., 2009; Gianopoulos, Stephens, & Davis, 2002; Gopinath et al., 2011; Hartley, Rochtchina, Newall, Golding, & Mitchell, 2010; Kochkin, 2000; Tomita, Mann, & Welch, 2001; Vuorialho, Karinen, & Sorri, 2006). Adherence is further reduced due to complaints by HA users that performance is poor in background noise (e.g., McCormack & Forntum, 2013). This is curious, considering the evidence of better listening effort outcomes associated with amplification. Assessing the subjective experiences associated with amplified speech, especially in noise, is required to understand these complaints, despite experimental evidence of gains related to HA use. Further research is needed to understand the amount of listening effort associated with amplification in comparison to normal speech, as it may affect the amount of benefit experienced by HA listeners. In addition, previous research has not examined combined effects of amplification, speech level, and background condition on listening effort.

Listening effort and speech level

In order to restore audibility to individuals with hearing impairment, amplification results in higher overall intensity levels. However, there is substantial evidence of impaired speech recognition with high intensity levels (e.g., Dubno et al., 2005; French & Steinberg, 1947; Hawkins & Stevens, 1950; Pollack & Pickett, 1958; Studebaker et al., 1999; Summers & Cord, 2007). While high intensity effects have been widely investigated in relation to speech recognition, examinations of the relationship between speech level and listening effort have been limited. One study to date has systematically investigated speech level and listening effort in the hearing-impaired population (Gatehouse & Gordon, 1990), but no clear relationship was uncovered. The current study was designed to investigate the effect of speech level independent from the effect of altered spectral shape on listening effort outcomes.

Listening effort in noise

Previous research has suggested that the addition of noise increases the amount of perceived effort associated with speech recognition (e.g., Houben et al., 2013, Zekveld, Kramer, & Festen, 2010). However, there has been limited research investigating how different types of background noise affect listening effort, particularly in relation to the recognition of spectrally shaped speech. Initial investigations of listening effort in different types of maskers, e.g., steady-state noise (SSN) and modulated noise maskers, have been mixed. One study found higher listening effort associated with a modulated masker in comparison to SSN in younger adult listeners (Desjardins & Doherty, 2013). However, a similarly designed study suggested opposite results (i.e. higher effort associated with SSN in comparison to a modulated masker; Kruger, Schulte, Brand, & Holube, 2017a; Kruger et al., 2017b). Other studies have found no effect of noise type for younger normal hearing listeners (Pals et al., 2015) and older hearing-impaired listeners (Desjardins & Doherty, 2013). While normal hearing listeners obtain clear benefits in recognition for modulated versus steady-state maskers due to momentary improvements in SNR (e.g., Festen & Plomp, 1990; Howard-Jones & Rosen, 1993; Miller, 1947; Miller & Licklider, 1950; Takahashi & Bacon, 1992; Wilson & Carhart, 1969), the current results regarding effort associated with these conditions is mixed. This could be due to interactions of altered spectral shape, as well as overall level, with background noise type. The current study will investigate steady-state and modulated maskers in relation to listening effort outcomes in order to clarify this relationship.

Finally, for speech recognition tasks in noise, higher levels of listening effort are associated with listeners with hearing impairment in comparison to listeners with normal hearing (Desjardins & Doherty, 2013; Helfer, Chevalier, & Freyman, 2010; Picou et al., 2013). This evidence suggests that the presence of background noise may interact with factors related to amplification, such as altered spectral shape and overall presentation level. Currently, the relationship between background noise type, as well as interactions between spectral shape and level, are not well understood. The current study will address this gap by examining the effect of background noise type on listening effort outcomes in young normal hearing listeners, as well as possible interactions between background noise type and spectral shaping.

Current study

Spectral shaping is necessary to restore audibility to listeners with hearing impairment.   In general, amplification is associated with decreases in listening effort (Ahlstrom et al., 2014; Bentler et al., 2008; Gatehouse & Gordon, 1990; Gustafson et al., 2014; Hornsby, 2013; Korczak, Kurtzberg, & Stapells, 2005; Kulkarni et al, 2012; Pals et al, 2013; Picou et al., 2013; Sarampalis et al., 2009). However, spectral shaping also increases overall speech level. Although the association between poorer speech recognition and high presentation levels has been established (e.g. Dubno et al., 2005; Summers & Cord, 2007), the relationship between speech level and listening effort has not been systematically investigated. Furthermore, it has been highly noted that the perceived benefit of amplification in noisy backgrounds is poor (e.g. McCormack & Fortnum, 2013). Combined with the highly variable findings regarding listening effort in noise, these studies highlight a significant need to examine listening effort in different noise backgrounds, particularly as it relates to the subjective experience. In order to investigate these factors, the current study assessed two measurements of listening effort (subjective and objective) while independently varying spectral shape and speech level. A subjective rating scale was utilized to uncover listeners' perceptions of listening effort, while RT was used as an objective behavioral indicator. Young normal hearing adults were tested in order to avoid factors related to altered suprathreshold processing due to cochlear pathology and elevated hearing thresholds found in the hearing-impaired population. Average thresholds from a group of listeners with hearing impairment were used as a standard audiogram to define spectral shaping parameters. Listening effort was evaluated in three background conditions: quiet, steady-state noise, and speech-modulated noise. These results are an analysis of listening effort outcomes that were collected along with speech recognition data from Experiment 1.

Method

Participants

Forty-seven young normal-hearing listeners were recruited from the undergraduate and graduate University of South Carolina population. Participants were between 18-29 years of age (Mean = 22 years, 45 female) and native speakers of American English. Criteria for normal hearing consisted of thresholds less than or equal to 20 dB HL at octave frequencies from 0.25-8 kHz (ANSI, 1997). Written consent was obtained from all participants. Listeners were compensated for participating via course credit or payment. All participants completed the experiment, but response times were only analyzed for a subset of participants (the first nineteen to complete the experiment).

Stimuli and design

The same stimuli were used as in Experiment 1. This included IEEE sentences presented in normal and shaped spectral conditions at three speech levels, corresponding to the average sensation level of a group of older hearing impaired participants (55 dB SPL), conversational level (70 dB SPL), and amplified (82 dB SPL) conditions. Spectrally shaped speech was amplified according to the mean thresholds of a group of older hearing-impaired participants, while no modifications were made to the spectrum of normal speech. All conditions were presented in quiet, steady-state noise, and speech modulated noise, resulting in a 2 (shaped/normal) x 3 (speech level) x 3 (background type) design. Both steady-state and speech modulated noise matched the long-term average power spectrum of the normal (unshaped) stimuli for normal conditions, and spectrally shaped stimuli for shaped conditions. One-third octave band levels were calibrated to match the long-term average spectrum of sentence materials using a Larson-Davis 800B sound level meter with linear weighting.

Listening Effort Ratings.

NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988) is a subjective rating scale that assesses a listener’s perception of effort and demand associated with a listening task. Four categories were adapted from this rating scale for the current study. The scale contained the following categories and their accompanying questions: Mental Demand (“How mentally demanding was the task?”), Performance (“How successful were you in accomplishing what you were asked to do?”), Effort (“How hard did you have to work to accomplish your level of performance?”), and Frustration (“How insecure, discouraged, irritated, stressed and annoyed were you?”). The scales had endpoints of “very low” and “very high” for all ratings, and all ratings are made on a 21-point scale. Participants were instructed to consider each scale individually.

Response Time.

Stimuli and participant responses to sentences were audio recorded simultaneously in separate audio channels to allow for the measurement of RT. RT measurements were calculated in Audacity® by measuring the time delay between the speech offset in the stimulus to the start of the participant's oral response. The difference in these two time points was calculated for each sentence trial and averaged across a block of ten sentences for each condition. RTs were measured for 19 participants as the time between offset of the sentence stimuli to the onset of the participant's oral response. This was done using the "SoundFinder" function in Audacity® and manually adjusted by two trained scorers. This function divides an audio file by placing region labels for areas of sound that are separated by silence. Manual adjustment was required because of the noise offset in the stimuli, which occurred after the speech offset. The final RT was an average of the two scorers for each spectral condition within the three background conditions. In addition, RTs associated with trials that a participant either a) did not respond during the trial or b) repeated none of the sentence keywords correctly were excluded from the analysis in order to ensure that only trials that involved some perceptual processing were included. This resulted in 98 (9 percent) and 320 (28 percent) trials across all participants excluded from SMN and SSN conditions, respectively.

Procedure

Listening effort outcomes were collected during the speech recognition task in Experiment 1. This involved the participant listening to a sentence and repeating back what they heard. The three background conditions were presented in blocks, with the quiet block present first and the noise blocks counterbalanced. Each background noise block consisted of six experimental conditions (2 spectral shapes x 3 speech levels) with ten sentences per condition (60 sentences total). Two additional conditions were assessed, but the current analysis focuses on only six of the eight total conditions tested. At the end of each condition (i.e., set of 10 sentences), participants were asked to complete the NASA-TLX rating scale. A demonstration trial of ten sentences was presented at the beginning of each background block to familiarize the participant with the rating procedure. Stimuli and participant responses were audio recorded simultaneously in independent channels to allow for RT measurement.

Results

Reliability: Listening Effort Ratings

Cronbach's alpha was calculated for all four rating scales to determine if the four variables reliably measure the same latent variable (i.e., listening effort). This resulted in a Cronbach's alpha of 0.89, suggesting a high level of internal consistency for these variables. Cronbach's alphas were also calculated for each rating scale if removed from the set of variables. As a result of this analysis, it was determined that the Cronbach's alpha would be greater if one variable was removed from the group (the Performance scale: $\alpha = .92$ if removed). Therefore, for the remaining analyses, Mental Demand, Effort, and Frustration ratings were collapsed, by summation, into one "Listening Effort" rating, and Performance ratings were not included in the present analysis.

Background Type

Scores were collapsed across spectral and level conditions for each background noise type and paired t-test comparisons were conducted to examine the effect of background noise on listening effort ratings and RTs. Both listening effort and RTs demonstrated significantly reduced effort in quiet compared to the two noise background conditions (Listening effort: SMN [t (38) = 15.68, p < .05], SSN [t (38) = 11.22, p < .05]; RT: SMN [t (18) = 3.05, p < .05], SSN [t (18) = 2.78, p < .05]). Furthermore, listening effort ratings were also sensitive to differences between the two background noise types, with reduced subjective listening effort in SMN [t (38) = 3.46, p < .05], while RT was not sensitive to this difference [t (18) = 0.18, p > .05]. The effects of shaping and level are examined individually for each noise type and listening effort outcome in the subsequent analyses.

Listening effort ratings

The listening effort score (a composite of demand, effort, and frustration) was transformed using an arcsine transformation to meet assumptions of normality. These scores are displayed in Figure 3.1 according to background condition. In general, participants had lowest listening effort ratings in quiet, followed by SMN and SSN. Results for each background condition were analyzed using a 2 (Spectral condition: Normal or Shaped) x 3 (Level: 55, 70, and 82 dB SPL) repeated-measures analysis of variance for each background type. Pairwise comparisons were conducted for each background type to examine any interactions. Tests were conducted using a Bonferroni-Holms adjusted alpha level for multiple comparisons.

All three background conditions demonstrated a significant effect of level (quiet: [F (1,46) = 5.6, p < .01], SMN: [F (1,46) = 3.5, p < .05], SSN [F (1,46) = 12.2, p < .0001]). The main effect of shaping was only significant in noise (quiet: p > .05, SMN: [F (1,46) = 60.8, p < .0001], SSN: [F (1,46) = 26.3 p < .0001]). An interaction was observed in all three background conditions as well (quiet: [F (2,92) = 5.1, p < .01], SMN: [F

(2,92) = 4.7, $p < .05$], SSN: [F (2,92) = 7.0, $p < .01$]). Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels in the three background conditions.

Pairwise comparisons first investigated the effect of shaping at the three different speech levels. In quiet, shaping tended to increase effort ratings ($p < .05$), except at 55 dB SPL ($p > .05$). In contrast, in both SMN and SSN, shaping tended to decrease subjective listening effort (SMN: $p < .05$; SSN: $p < .05$) except in SSN at 82 dB SPL, where listening effort ratings were not significantly different between normal and shaped conditions ($p > .05$).

T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. In general, effects of level were minimal for spectrally shaped speech, only resulting in a significant difference in the SSN background condition. In this condition, a high speech level (82 dB SPL) was associated with a higher level of effort relative to 55 and 70 dB SPL. In contrast, greater effects of speech level were observed for normal speech. In this condition, higher speech levels were associated with a higher level of effort in both SMN ($p < .05$) and SSN ($p < .05$) background conditions.

Response Times

RTs were not normally distributed, so an inverse transformation was applied to the raw data to ensure normality. The raw scores are displayed in Figure 3.2. In general, participants had fastest RTs in the quiet background condition, followed by SMN and SSN. Results were analyzed as with listening effort ratings.

The main effect of level was only significant in quiet [F (1,18) = 3.4, $p < .05$] and the main effect of shaping was only significant in SMN [F (1,18) = 5.6, $p < .05$]. A shape by level interaction was observed in both the quiet condition and in SMN as well (quiet: [F (2,36) = 5.8, $p < .01$], SMN: [F (2,36) = 5.0, $p < .05$]). In SSN, no significant main effects or interactions were observed ($p > .05$). This was likely due to greater variability (Average SD: Quiet = 0.25; SMN = 0.32; SSN = 0.41). As a result of non-significance, SSN was not further analyzed.

Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels. In quiet, RTs were faster for shaped speech at the lowest level ($p < .05$) and were not significantly different at higher levels ($p > .05$). In contrast, in SMN, shaping tended to decrease RTs ($p < .05$) except at 82 dB SPL, where RTs were not significantly different between normal and shaped conditions ($p > .05$).

T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. In general, higher levels were associated with faster RTs in both quiet and in SMN for both normal and shaped speech (quiet: $p < .05$, SMN: $p <.05$). One exception to this trend were RTs for normal speech in SMN, which were significantly faster at 55 dB SPL than 70 dB SPL ($p < .05$).

Association between subjective ratings and response times

Pearson product moment correlations were used to investigate the relationship between listening effort ratings and RTs. Average scores for each condition pooled across background condition were correlated across listening effort ratings and RTs. The two variables were not significantly correlated ($r = 0.33$, $p > 0.5$). This is consistent with other research that has utilized subjective listening effort measurements in conjunction with a second listening effort outcome (Gosselin & Gagne, 2011; Fraser, Gagne, Alepins, & Dubois, 2011; Larsby, Hallgren, Lyxell, & Arlinger, 2005). This indicates that RT may capture processes independent of subjective experiences of effort. Although subjective measures provide evidence of perceived effort, RT appears to reflect the time required for perceptual processing in this study. That is, even if a high amount of processing resources are required for a task, this expenditure may not be perceived subjectively by the listener.

Discussion

The current study examined how spectral shaping and speech level affect listening effort, as defined by subjective ratings and RTs, as a function of background condition. These outcomes are important clinically as perceived and experienced effort may influence decisions made by HA users related to adherence. Indeed, two of the most commonly reported reasons for non-use of HAs are ineffectiveness in noisy situations and providing poor benefit (McCormack & Fortnum, 2013). These factors may be related to the amount of listening effort perceived by the HA user. The results of this study demonstrate that listening effort measures are sensitive to manipulations of spectral shaping, speech level, and noise background.

The effect of spectral shaping

Spectrally shaping speech resulted in lower effort as indexed by subjective ratings and RT measurements for both noise background conditions. The benefits of amplification in regards to listening effort have been documented, especially in background noise (e.g. Ahlstrom et al., 2014; Gatehouse & Gordon, 1990). In the current study, both RTs and listening effort ratings were consistent for noise conditions at higher presentation levels. Despite the consistency between these two measures, listening effort ratings and RTs were not correlated across conditions, indicating that the two measures may capture different aspects of listening effort.

In quiet, ratings of effort suggest that at higher levels, greater effort was perceived for shaped speech. However, at lower levels, shaped speech was associated with reduced processing time (i.e. RTs). According to the Ease of Language Understanding model (ELU; Ronnberg et al., 2017), this could be due to an acoustic mismatch between the stored representation (i.e., not spectrally altered) of speech and the acoustic signal. This mismatch could require more perceptual processing to resolve the signal, resulting in slower RTs, and higher listening effort.

In contrast to the quiet condition, in noise (SMN and SSN), listening effort ratings and RTs were significantly better for shaped speech. This result could be due to more favorable glimpsing opportunities for spectrally shaped speech, indicated by a greater number of glimpses at more favorable SNRs. This is due to the nature of spectral shaping: primarily high frequencies (above 2 kHz) were amplified, resulting in greater amplification of high frequency portions than low frequency portions of a sentence. This could create more glimpsing opportunities in the shaped speech condition, where certain temporal intervals of the sentence receive much greater amplification in comparison to the normal speech condition. This access, according to FUEL (Pichora-Fuller et al., 2016), may have resulted in reduced perceptual processing demands due to a better preserved signal, therefore mediating outcomes of listening effort.

The effect of background type

Previous research has suggested that the addition of noise increases the amount of perceived effort associated with speech recognition (e.g., Houben et al., 2013, Zekveld et al., 2010), which was also observed here. When averaged across spectral and level conditions, RTs and listening effort ratings were significantly better in quiet compared to

SMN and SSN backgrounds. Furthermore, listening effort ratings also indicated reduced effort in SMN compared to SSN backgrounds. Thus, glimpsing opportunities in modulated noise served to reduced perceived effort. This is in contrast to Desjardins and Doherty (2013) who found higher listening effort in modulated rather than steady-state maskers.

The effect of speech level

The effect of speech level on listening effort varied according to background condition. In general, better listening effort outcomes were associated with higher speech levels in quiet across spectral conditions. This was demonstrated by ratings of listening effort and RT decreasing as speech level increased. This supports previous research of speech recognition with varying speech level, which demonstrated increased or stationary performance with increasing overall levels in quiet, as audibility is greater as levels increase (e.g. Perez-Gonzalez & Lopez-Poveda, 2013; Studebaker et al., 1999).

A different pattern of results was observed in noise backgrounds. In noise, poorer listening effort outcomes were associated with higher speech levels across spectral conditions. This was demonstrated by ratings of effort and RT increasing as speech level increased. This result is also consistent with speech recognition research of speech level, as higher speech levels are associated with decreases in performance, especially in noise backgrounds (e.g. Dubno et al., 2012). This may be due to lower frequency selectivity at higher intensities, wider auditory filters, or an upward spread of masking from the amplification of low frequencies (as discussed in Dubno et al., 2012).

Summary and Conclusions

The purpose of amplification is to restore audibility to individuals with hearing loss. How this amplification affects listening effort is an important factor in clinical outcomes associated with hearing aids. For the group of normal hearing participants tested in this study, spectrally shaping speech increased listening effort in quiet, but reduced listening effort for the majority of conditions tested in noise. These results may be due to increased distortion in spectrally shaped speech, which may have increased perceptual processing resources required for lexical retrieval in quiet. However, in noise, improved glimpsing opportunities associated with spectrally shaped speech may have reduced perceptual processing demands. These findings have a few important implications clinically. First, it appears that spectral shaping is associated with better listening effort outcomes in young normal hearing listeners in noise. Second, overall speech levels should be considered in relation to listening effort and amplification techniques, as outcomes varied across background conditions as levels increased. While increases in level reduced listening effort in quiet, corresponding level increases in noise could be detrimental. This may potentially lead to the subjective complaints of HA users in noisy backgrounds. Third, listening effort outcomes were not correlated, indicating that RTs and perceived listening effort may provide related but qualitatively different indices of processing requirements. Future studies will need to examine how shaping and speech level affect listening effort outcomes in the hearing-impaired population where other factors, such as audibility and suprathreshold processing, may also influence results. In summary, for normal hearing listeners, spectral shaping is an effective means of reducing effort in noise backgrounds due to improved glimpsing, but increases in speech level should be minimized. The results from the current study suggest that spectral shaping may lower listening effort outcomes in normal hearing participants, but that benefit varies depending on background condition and speech level.

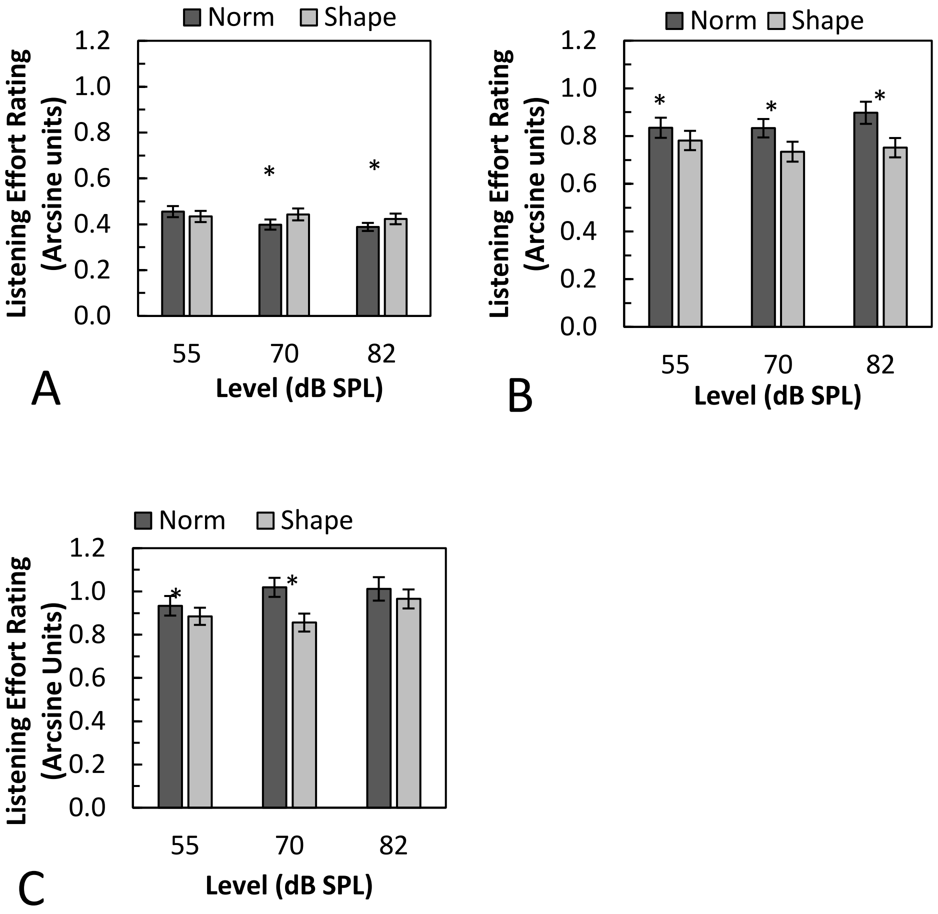

Figure 3.1. Listening effort ratings. Shown for the quiet (panel A), SMN (panel B), and SSN (panel C) background conditions. Ratings for shaped and normal speech conditions are displayed at 55, 70, and 82 dB SPL. Error bars = standard error of the mean. Asterisks indicate significantly different ratings between the two speech types.

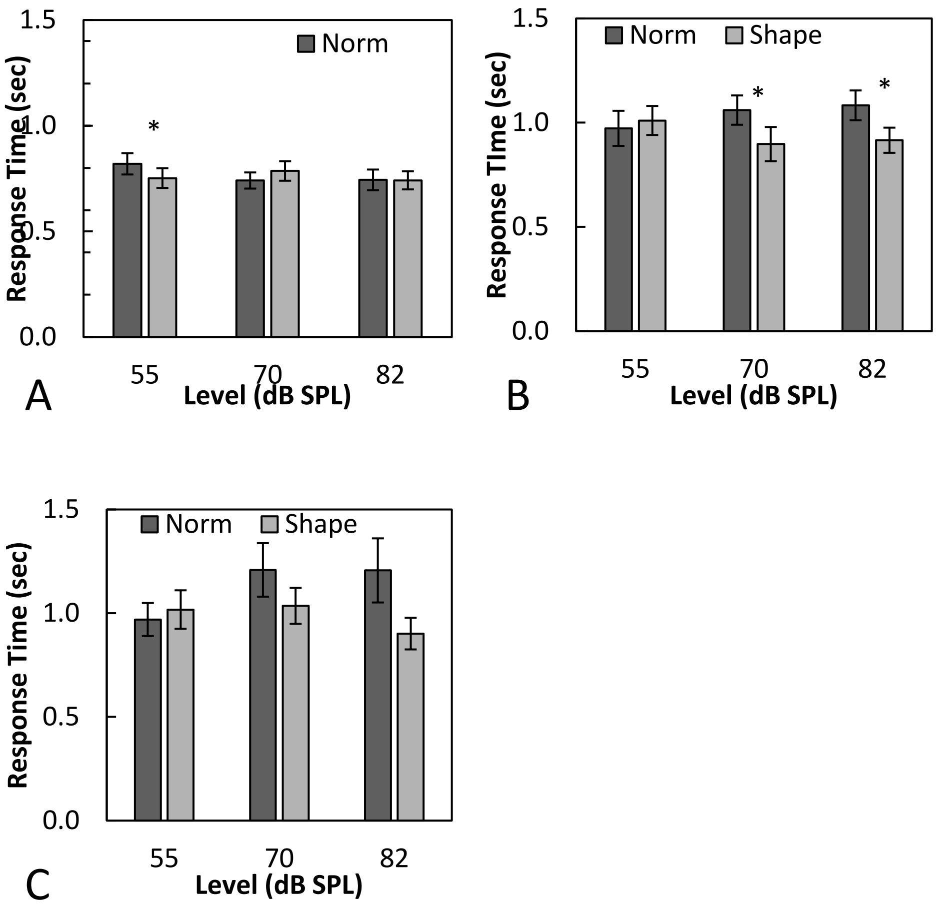

Figure 3.2. RTs. Shown for the quiet (panel A), SMN (panel B), and SSN (panel C) background conditions. RTs for shaped and normal speech conditions are displayed at 55, 70, and 82 dB SPL. Error bars = standard error of the mean. Asterisks indicate significantly different ratings between the two spectral types.

CHAPTER 4

EXPERIMENT 3: EFFECT OF SPECTRAL SHAPING ON SPEECH QUALITY RATINGS IN QUIET AND IN NOISE

Adherence is low in hearing aid (HA) users (McCormack & Fortnum, 2013) despite that one in eight U.S. adults are diagnosed with hearing loss (Lin et al., 2011). A possible explanation for poor adherence in HA users is poor perceived sound quality, which is a top complaint of hearing aid users (McCormack & Fortnum, 2013). Indeed, in a study that collected survey data, HA wearers ranked sound quality as the second most important area of improvement for HAs (Kochkin, 2002). In order to ensure audibility, HAs shape the spectrum of the speech and increase overall level. It is possible that these changes could influence the perception of sound quality. While spectral shaping and presentation level have been investigated in relation to speech recognition (e.g. Dubno et al., 2005; Horowitz et al. 2008), investigations utilizing sound quality ratings have primarily been focused on effects of compression and noise reduction algorithms (e.g. Moore, Fullgrabe, & Stone, 2011; Neumann, Bakke, Hellman, & Levitt, 1994; Neuman, Bakke, Mackersie, Hellman, & Levitt, 1998; Parsa, Scollie, Glista, & Seelisch, 2013). In addition, previous research has demonstrated that sound quality ratings are poorer in background noise (e.g. Anderson, 2010), indicating that perceived sound quality could vary depending on background conditions. Also, it is well established that higher presentation levels are associated with poorer speech recognition (e.g. Dubno et al., 2005; Summers & Cord, 2007). However, the combined effects of spectral shaping, presentation level, and background noise on perceived sound quality is not well understood. Given the reported complaints and low level of adherence in HA users, the measurement of sound quality has important clinical implications.

Sound quality

Sound quality ratings are often utilized when evaluating various hearing aid processing techniques, such as amplitude compression (e.g. Davies-Venn, Souza, & Fabry, 2007). These ratings are meant to provide information about dimensions of speech quality, such as pleasantness, intelligibility, loudness, and overall impression, which often differ from speech recognition outcomes in different background conditions (e.g., Lunner, Hellgren, Arlinger, & Elberling, 1997). "Sound quality" can be defined as the naturalness of a sound, as defined by its color, timbre, or character (Slawson, 1985). There are various ways to utilize sound quality ratings, including ranking procedures (whether a sound has higher quality than a competitor), semantic differential procedures (descriptors indicating why a sound is suitable to convey a message), and category scaling (how sound quality compares to a reference stimuli) (Fastl, 2005). The current study will utilize a noncomparative form of an interval scale in order to better understand sound quality associated with amplification.

Sound quality and spectral shaping

Spectral shaping is a technique used to control for audibility across listeners with and without hearing impairment. This involves amplifying a signal across frequencies according to the listener's audiometric thresholds. For most listeners, high frequencies are more amplified in spectrally shaped speech due to age-related hearing loss. High frequency amplification accounts for primary declines in speech recognition due to audibility (Humes, 2013). Indeed, the majority of studies examining the effect of amplification in different frequency regions have demonstrated the benefit of high frequency amplification for speech recognition for individuals with high frequency hearing loss (Hornsby & Ricketts, 2003; 2006; Skinner, 1980; Sullivan et al, 1992; Turner & Henry, 2002). Given this evidence of improved speech recognition with high frequency amplification, it is of interest to examine how this amplification affects sound quality.

Only a few studies to date have examined the effects of amplification on sound quality measurements. These studies involved manipulating low- and high-pass filter cutoff frequencies for amplification. Evidence from this research demonstrated an improvement in sound quality as bandwidth increased for normal hearing listeners (Franks, 1982; Ricketts, Dittberner, & Johnson, 2008), as well as hearing-impaired listeners (Ricketts et al., 2008). This evidence suggests that high-frequency amplification results in higher perception of sound quality. However, this does not support complaints of low sound quality in HA users (McCormack & Fortnum, 2013), many of whom utilize high frequency amplification.

Sound quality and speech level

In addition to spectral shaping, signals are amplified to higher overall intensity levels in order to restore audibility to HA users. There is substantial evidence of impaired speech recognition with high intensity levels (e.g., Dubno et al., 2005; French & Steinberg, 1947; Hawkins & Stevens, 1950; Pollack & Pickett, 1958; Studebaker et al., 1999; Summers & Cord, 2007). However, there is limited research on how high intensity levels affect sound quality measurements. One study examining amplitude compression obtained sound quality measurements at three intensity levels with hearing-impaired listeners, and found a preference for higher presentation levels (Davies-Venn et al., 2007). However, this effect has not been investigated in relation to spectral shaping. In addition, the relationship of speech level and background noise in regards to sound quality has not been established.

Sound quality and noise

The benefit of hearing aids is limited in background noise (e.g. Humes, 1991), and HA users often find background noise uncomfortably loud or annoying (Kochkin, 2000). In general, evidence suggests that perception of sound quality decreases in the presence of background noise (e.g. Anderson, 2010; Arehart, Kates, & Anderson, 2010; Neuman et al., 1994; Davies-Venn et al., 2007). However, the relationship between sound quality and type of background noise has not been investigated. It is widely established that speech recognition with normal hearing listeners is better in background noise that is temporally fluctuating, due to the ability to take advantage of speech glimpses in the masker (e.g., Festen & Plomp, 1990; Howard-Jones & Rosen, 1993; Miller, 1947; Miller & Licklider, 1950; Takahashi & Bacon, 1992; Wilson & Carhart, 1969). However, listeners with hearing impairment commonly do not benefit from this release from masking, or better performance in comparison to a steady state masker (Bronkhorst & Plomp, 1992; Eisenberg, Dirks, & Bell, 1995; Festen & Plomp, 1990; Gustafson & Arlinger, 1994; Hygge et al., 1992; Phillips, Rappaport, & Gulliver, 1994). Furthermore, there is evidence that amplification may affect speech recognition depending on the nature of the background noise, suggesting that the benefit from amplification may be reduced in complex maskers (Fogerty et al., 2015b). While recognition of amplified speech seems to differ depending on the background condition, it is unclear if perception of sound quality follows this trend. The current study compares the effects of spectral shaping and level in multiple background conditions to better understand how amplification affects sound quality perception in different listening situations.

Current study

Spectral shaping is a fundamental component of hearing aids that is used to restore audibility to individuals with hearing impairment. However, while speech recognition may improve, particularly in quiet, due to improved audibility (Humes, 2007), the perceived quality of speech following amplification continues to be a limitation to processing that underlies HA user complaints (McCormack & Fortnum, 2013). In general, high frequency amplification has been associated with better sound quality (Franks, 1982; Ricketts et al., 2008). However, this effect has not been investigated in relation to the combined effects of increased speech level, which can result in poorer speech recognition (e.g. Dubno et al., 2005; Summers and Cord, 2007). In addition, there is evidence that sound quality is poorer in background noise (Anderson, 2010; Arehart et al., 2010; Neuman et al., 1994), but this effect has not been compared between fluctuating and stationary types of background noise. In order to investigate these factors, the current study systematically varied spectral shape and speech level in quiet and two different background noise conditions. Sound quality ratings were obtained from young normal hearing listeners. This population was chosen to avoid factors related to altered suprathreshold processing due to cochlear pathology, as well as effects of elevated hearing thresholds. Average thresholds from a group of listeners with hearing impairment were used as a standard audiogram to define spectral shaping parameters. In order to model behavioral ratings, the Hearing Aid Speech Quality Index (HASQI, Kates & Arehart, 2014b) was utilized to quantify the amount of distortion introduced by spectral shaping. This study directly measured the combined effects of shaping, overall level, and background noise on sound quality perception.

Method

Participants

Twenty young normal hearing participants were recruited from the undergraduate and graduate University of South Carolina population. Participants were between 18 and 26 years of age (Mean = 22 years, two males) and native speakers of American English. Criteria for normal hearing consisted of thresholds less than or equal to 20 dB HL at octave frequencies from 0.25-8 kHz (ANSI, 1997). Written consent was obtained from all participants. Listeners were compensated for participating via course credit or payment.

Stimuli and design

The same stimuli were used as in Experiment 1. This included IEEE sentences presented in normal and shaped spectral conditions at three speech levels, corresponding to sensation level (55 dB SPL), conversational level (70 dB SPL), and amplified (82 dB SPL) conditions. Spectrally shaped speech was amplified according to the mean thresholds of a group of older hearing-impaired participants, while no modifications were made to the spectrum of normal speech. All conditions were presented in quiet, steady-state noise, and speech-modulated noise, resulting in a 2 (shaped/normal) x 3 (speech level) design in three background conditions. Both steady-state and speech modulated noise matched the long-term average power spectrum of the normal (unshaped) stimuli for normal conditions, and spectrally shaped stimuli for shaped conditions. One-third octave-band levels of steady-state and speech-modulated noise were calibrated to match the long-term average spectrum of sentence materials using a Larson-Davis 800B sound level meter with linear weighting.

Quality Ratings.

Following the procedures of Davies-Venn et al. (2007), the Speech Intelligibility Rating Test (SIR, Cox & McDaniel, 1984, 1989), was adapted to develop quality ratings for this experiment. This modification of the SIR contains four scales rating: overall impression, pleasantness, intelligibility, and loudness. The definitions for each aspect was defined as follows:

1. Overall Impression: Considering everything that you have heard, what do you think about the sound?
2. Loudness: How loud, strong, or forceful is the sound? The opposite of loud is soft, weak or timid/faint.
3. Intelligibility: How clear is the speech or what percent of the speech do you understand? The opposite of intelligible is extremely hard to understand and unclear.
4. Pleasantness: How pleasing is the tonal quality of the sound? The opposite of pleasant is unpleasant.

Each aspect was presented using a 1-10-point scale. In all aspects but loudness, 10 represents the optimum rating. Loudness ranges from 1, “not loud at all,” to 10, “very loud.” Ratings were made for each sentence in each processing condition.

Procedure

The three background conditions were presented in blocks, with the quiet block presented first and the noise blocks counterbalanced. Each background noise block consisted of 90 trials. Each block consisted of six sentence conditions (2 spectral shapes x 3 speech levels). Two additional conditions were assessed, but the current analysis focuses on only six of the eight total conditions tested. After listening to each sentence, participants were asked to complete the sound quality rating scale. A demonstration trial of ten sentences was presented at the beginning of each background condition to familiarize the participant with the rating procedure.

Results

Reliability of Quality Ratings

Cronbach's alpha was calculated for all four rating scales to determine if the four variables reliably measure the same latent variable (sound quality). This resulted in a Cronbach's alpha of 0.82, suggesting a high level on internal consistency for these variables. Cronbach's alphas were also calculated for each rating scale if removed from the set of variables. As a result of this analysis, it was determined that the Cronbach's alpha would be greater if one variable was removed from the group (the Loudness rating scale: $\alpha = .88$ if removed). Therefore, for the remaining analyses, Overall, Intelligibility, and Pleasantness ratings were collapsed, by summation, into one "Sound Quality" rating. Loudness ratings were not included in this analysis.

Sound quality ratings

As stated above, the composite sound quality score was calculated by summing the participant's Overall, Intelligibility, and Pleasantness ratings. The composite score met the assumption for a normal distribution. These scores are displayed in Figure 4.1 according to background condition. In general, participants had the highest quality ratings in quiet, followed by SMN and SSN. Results were analyzed using a 2 (Spectral condition: Normal or Shaped) x 3 (Level: 55, 70, and 82 dB SPL) repeated-measures analysis of variance for each background type. Pairwise comparisons were conducted for each background type to compare the effect of shaping at three speech levels and the effect of the two different spectral shapes (Normal, Shaped). All pairwise comparisons were conducted using a Bonferroni-Holms adjusted alpha level for multiple comparisons.

Background Condition

Scores were averaged across spectral and level conditions for each background noise type, and paired t-tests were conducted to examine the effect of background noise on sound quality ratings. The quiet background had significantly higher quality ratings than both SMN and SSN backgrounds for both shaped (SMN: [t (19) = 6.39, p < .001] SSN: [t (19) = 10.38 , p <.001]) and normal speech (SMN: [t (19) = 10.39, p < .001] SSN: [t (19) = 16.54 , p <.001]). In addition, the SMN background condition had significantly higher ratings than SSN for both shaped [t (19) = 3.13, p < .001] in normal speech [t (19) = 4.43, p < .001]. In this study, it appears that higher sound quality ratings are associated with quiet, than noise, backgrounds. In addition, this evidence suggests that greater glimpsing opportunities improved the perceived quality of the speech.

Quiet.

Compiled sound quality rating scores in the quiet background condition are displayed in Figure 4.1, panel A for normal and shaped stimuli at three overall levels. In general, in quiet, quality rating scores were better at higher presentation levels. In addition, at the lowest presentation level, ratings were higher for normal compared to shaped speech. Significant main effects of shape [$F(1,19) = 16.0$, $p < .01$] and level [$F(1,19) = 20.1$, $p < .001$] were observed as well as a shape by level interaction [$F(2,38) = 8.9$, $p < .01$]. Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels. At 55 dB SPL, quality ratings were significantly higher for normal speech than for shaped speech ($p > .05$). This effect was not significant at 70 and 82 dB SPL ($p > .05$). T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. Sound quality ratings for normal and shaped speech at 55 dB SPL were significantly lower than corresponding ratings at 70 or 82 dB SPL ($p < .05$). This suggests that, regardless of speech type, sound quality ratings are better with higher overall levels in quiet.

Acoustic analysis.

HASQI was used to investigate an acoustic model of sound quality that may explain behavioral results. This specific metric was chosen to explain the distortions introduced as a result of spectral shaping. Due to the SNR used in this study, HASQI could only be computed for stimuli in the quiet background condition (see Kates et al., 2018). This metric utilized a comparative acoustic analysis, where the output of an auditory model for an unprocessed signal (IEEE sentences) was compared to the output of an auditory model for a processed signal (spectrally shaped/normal sentences). The metric combined a nonlinear index, which captures differences in spectral shapes between stimuli over time, and a linear index, which captures the differences between long-term average spectra between stimuli. In the following analysis, the unprocessed IEEE sentence was compared to the same IEEE sentence processed in both normal and shaped conditions in quiet. An average was computed for each condition across all sentences tested in quiet (60 sentences total). The HASQI output is constrained between 0 and 1, with 0 indicating poor sound quality and 1 indicating perfect sound quality. Normal and shaped conditions were always analyzed in comparison to their relative clean signals.

Average HASQI scores for the Quiet background condition for both speech types are displayed in Figure 4.2. In general, higher scores were seen at the lowest presentation level and for normal speech. Results were analyzed using a 2 (Spectral condition: Normal or Shaped) x 3 (Level: 55, 70, and 82 dB SPL) repeated-measures analysis of variance. Significant main effects of shape [$F (1,9) = 1469.8$, $p < .001$] and level [$F (1,9) = 14.8$, $p < .001$] were observed, as well as a shape by level interaction [$F (2,18) = 185.1$, $p < .001$]. Paired t-test comparisons were conducted to examine the effect of spectral condition on HASQI scores at three speech levels. For 55, 70, and 82 dB SPL, HASQI scores were significantly better for normal speech than shaped speech ($p < .05$). Regarding level, 55 dB SPL had significantly better scores than 70 dB SPL and 82 dB SPL in both shaped and normal conditions ($p < .05$). This contrasts with sound quality ratings, where higher speech levels were associated with better sound quality ratings for shaped and normal speech.

HASQI scores and average sound quality ratings in the quiet background condition were examined to explore the relationship between the acoustic analysis and participant ratings (Figure 4.2). At 70 and 82 dB SPL, it appears that HASQI is more sensitive to differences in quality than participant ratings. However, at 55 dB SPL, a similar pattern emerged with both HASQI scores and participant ratings, with higher outcomes for normal speech. However, HASQI overpredicted sound quality measurements at 55 dB SPL in comparison to participant ratings. A correlational analysis was not examined between HASQI scores and participant ratings due to inconsistent effects across level. A lack of a meaningful correlation indicates that HASQI was not sensitive to average sound quality ratings. This could be due to multiple factors, including the sensitivity of the sound quality rating scale, as well as how reliable individual listener’s ratings were between conditions.

SMN.

Sound quality ratings in the SMN background condition are displayed in Figure 4.1, panel B for normal and shaped stimuli at three overall levels. A significant main effect of level [F (1,19) = 4.5, p < .05] was observed as well as a shape by level interaction [F (2,38) = 16.0, p < .001]. The main effect of shape was not significant (p > .05). Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels. For 82 dB SPL, sound quality ratings were significantly higher with shaped speech than normal speech (p < .05). In 55 and 70 dB SPL conditions, sound quality ratings were not significantly different between the two speech types (p > .05). T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. For shaped speech, 70 and 82 dB SPL conditions had significantly higher sound quality ratings than shaped speech at 55 dB SPL (p < .05). For normal speech, both 55 and 70 dB had significantly higher sound quality ratings than normal speech at 82 dB SPL. In general, it appears that shaped speech has better sound quality ratings at higher levels in SMN.

SSN.

Sound quality ratings in the SSN background condition are displayed in Figure 4.1, panel C for normal and shaped stimuli at three overall levels. Significant main effects of shape [F (1,17) = 26.3 p < .0001] and level [F (1,17) = 23.1, p < .0001] were observed as well as a shape by level interaction [F (2,34) = 48.6, p < .0001]. Paired t-test comparisons were conducted to examine the effect of spectral condition at three overall levels. For 55 and 70 dB SPL, sound quality ratings were higher with shaped speech than normal speech (p < .05). At 82 dB SPL, quality ratings were not significantly different between normal and shaped conditions (p > .05). T-tests were also conducted to examine the effect of level within shaped and normal speech conditions. For shaped speech, sound quality ratings at 55 and 70 dB SPL were significantly higher than at 82 dB SPL (p < .05). In addition, sound quality ratings at 70 dB SPL were significantly higher than at 55 dB SPL for shaped speech. For normal speech, sound quality ratings at 55 dB SPL were significantly higher than at 70 dB SPL (p < .05). Furthermore, sound quality ratings at 70 dB SPL were significantly higher than ratings at 82 dB SPL for normal speech (p < .05). In general, it appears that in SSN, sound quality ratings are poorer for higher presentation levels regardless of speech type.

Discussion

The current study examined how spectral shaping, speech level, and background condition affect the perception of speech quality by normal hearing listeners. Amplification is a necessary technique to restore audibility to hearing-impaired listeners, however, changes in spectral shape and speech level may interact to influence the perception of sound quality. These factors could underlie the perception of poor sound quality often experienced by hearing-impaired listeners (McCormack & Fortnum, 2013). The results of this study indicate that spectrally shaped speech is associated with better sound quality in noise, and poorer sound quality at low speech levels in quiet. In addition, sound quality varied significantly across background conditions.

Effect of spectral shaping

The relationship between sound quality and spectral shaping varied by background condition. In quiet, normal and shaped speech had similar sound quality ratings for higher presentation levels (70 and 82 dB). This is similar to results found with hearing-impaired listeners in previous research (Franks, 1982). However, a lack of difference between spectral conditions could be due to ceiling effects, as sound quality ratings were very high in quiet. At the lowest presentation level in quiet, however, normal speech was associated with higher sound quality than shaped speech This finding is supported by the acoustic analysis included in this study that demonstrated that normal speech had higher predicted sound quality in quiet in comparison to shaped speech.

In noise conditions, spectrally shaped speech was generally associated with higher sound quality than normal speech, especially at high levels in SMN, and mid- to low levels in SSN. This is consistent with previous literature that found better sound quality perception with increasing bandwidth in young normal hearing listeners (Franks, 1982; Ricketts et al., 2008). In addition, this finding is in line with speech recognition research, which has demonstrated the benefit of high frequency amplification extensively (Hornsby & Ricketts, 2003; 2006; Skinner, 1980; Sullivan et al, 1992; Turner & Henry, 2002). Given this combined evidence, it appears that spectral shaping has a positive effect on sound quality perception in noise, though the pattern of benefit varies by level and type of background noise.

In general, spectrally shaped speech was especially beneficial in SMN and SSN conditions in comparison to normal speech. This result could be due, in part, to acoustic factors associated with spectrally shaped speech. As demonstrated by ideal binary mask analyses in Experiment 1, high frequency amplification provides more frequent glimpse opportunities, and more favorable SNRs during these glimpses. This allows the listener to potentially access more acoustic cues of the target sentence in background noise. In quiet, however, glimpsing was not a factor, which may explain why normal speech had higher quality ratings than shaped speech in this condition.

A further explanation for the benefit of spectrally shaped speech could be related to perceptual learning. The Ease of Language Understanding model (ELU; Ronnberg, 2003, Ronnberg et al., 2013) predicts that the acoustic similarity between a stimulus and its stored mental representation affects lexical processing. Due to the acoustic changes that occur when speech in amplified, this mismatch between heard and stored representation could be larger for shaped than normal speech. In this case, it is possible that participant’s lexical processing of shaped speech could improve as they are exposed to the novel speech throughout the testing session. Because there is not a mismatch between normal speech and its stored representation, a smaller effect of perceptual learning would be expected relative to that of shaped speech. By the same turn, it is possible that in the quiet condition, which was always presented first, listeners did not have enough exposure to shaped speech to benefit from perceptual learning. However, the effect of perceptual learning on sound quality has not been investigated, and further research is needed to understand this relationship.

The effect of speech level

The effect of speech level on sound quality ratings varied depending on the background condition. Speech recognition research has suggested a similar finding, where increasing level has different effects depending on whether or not speech is presented in quiet or in noise (Studebaker et al., 1999). In the current study, quality ratings for both spectral conditions were better at higher presentation levels in quiet. Previous research has demonstrated this pattern in quiet as well with speech recognition outcomes, as audibility is greater as levels increase (e.g. Perez-Gonzalez & Lopez-Poveda, 2013; Studebaker et al., 1999). In SMN, however, level factors were dependent on the spectral condition. For shaped speech, quality ratings were better at higher speech levels, while ratings for normal speech were better at lower speech levels. In contrast, in SSN, sound quality ratings were poorer as overall level increased for both spectral conditions. This pattern has also been documented in speech recognition literature, where performance is best at mid-levels, but decreases at high levels in broadband noise (Dubno et al., 2012). This decrease in perceived sound quality at high speech levels could be due to several factors, including lower frequency selectivity at high intensities, wider auditory filters, as well as an upward spread of masking from amplification in lower frequencies. In sum, it appears that increasing speech level has a positive effect on sound quality ratings in quiet, and the opposite effect is observed in noise. This has important implications clinically, as speech levels may need to be adjusted depending on the background condition in order to ensure better sound quality. The exception to this pattern was shaped speech in SMN, where quality ratings were better at higher speech levels. This pattern is more similar to that of the quiet condition, where quality ratings also increased with overall level. We may speculate that this could be due to better glimpsing opportunities associated with shaped speech, such as more frequent glimpse opportunities, and more favorable SNRs during glimpses, as discussed above.

The effect of noise

Previous research has suggested that the addition of noise decreases perceived sound quality, which was also observed here (e.g. Anderson, 2010; Arehart et al., 2010; Neuman et al., 1994). When averaged across spectral and level conditions, sound quality ratings were significantly better in quiet compared to both SMN and SSN backgrounds. Furthermore, the results also indicated better sound quality ratings in SMN compared to SSN conditions. This evidence suggests that the glimpsing opportunities present in modulated noise served to increase the perception of sound quality. This is similar to evidence found in speech recognition literature, where young normal hearing listeners perform better in modulated noise in comparison to steady-state noise (e.g., Festen & Plomp, 1990; Howard-Jones & Rosen, 1993; Miller, 1947; Miller & Licklider, 1950; Takahashi & Bacon, 1992; Wilson & Carhart, 1969). However, it remains to be seen if this pattern would be similar with hearing-impaired adults, where the benefit of a modulated masker in comparison to steady-state noise is not often observed for speech recognition (Bronkhorst & Plomp, 1992; Eisenberg, et al, 1995; Festen & Plomp, 1990; Gustafson & Arlinger, 1994; Hygge, et al., 1992; Phillips, et al., 1994). Future research is needed to examine sound quality ratings in varying backgrounds in hearing-impaired listeners.

Summary and Conclusions

Amplification is necessary to restore audibility to individuals with hearing-impairment. How this amplification affects perceived sound quality is important clinically, as poor sound quality is a top complaint of hearing aid users (McCormack & Fortnum, 2013). In general, for the group of normal hearing listeners in this study, spectrally shaped speech was associated with better sound quality in noise. However, at low speech levels in quiet, spectral shaping was associated with lower sound quality. In addition, sound quality ratings varied across background conditions, with better outcomes in quiet than in noise.

These findings have a number of clinical implications. First, it is possible that hearing aid amplification may be related to the perception of poorer sound quality in quiet. This is important because hearing aid fitting is often conducted in quiet, and perception of sound quality may lead hearing aid users to form opinions about the perceived benefit of amplification, which could in turn affect adherence. In addition, overall speech levels should be considered when utilizing amplification techniques, as sound quality varied across background condition as levels increased. While sound quality was higher at increasing levels in quiet, corresponding level increases in noise were associated with poorer sound quality. This finding could be related to complaints of poor hearing aid performance in noise. Finally, sound quality outcomes in different background conditions need to be examined, as sound quality was not perceived equally across the conditions tested.

In summary, for normal hearing listeners, spectral shaping was associated with better sound quality outcomes in noise backgrounds, but not in quiet. In addition, increases in speech level in noisy backgrounds were associated with poorer sound quality. Finally, sound quality outcomes were better in fluctuating, rather than steady-state maskers.

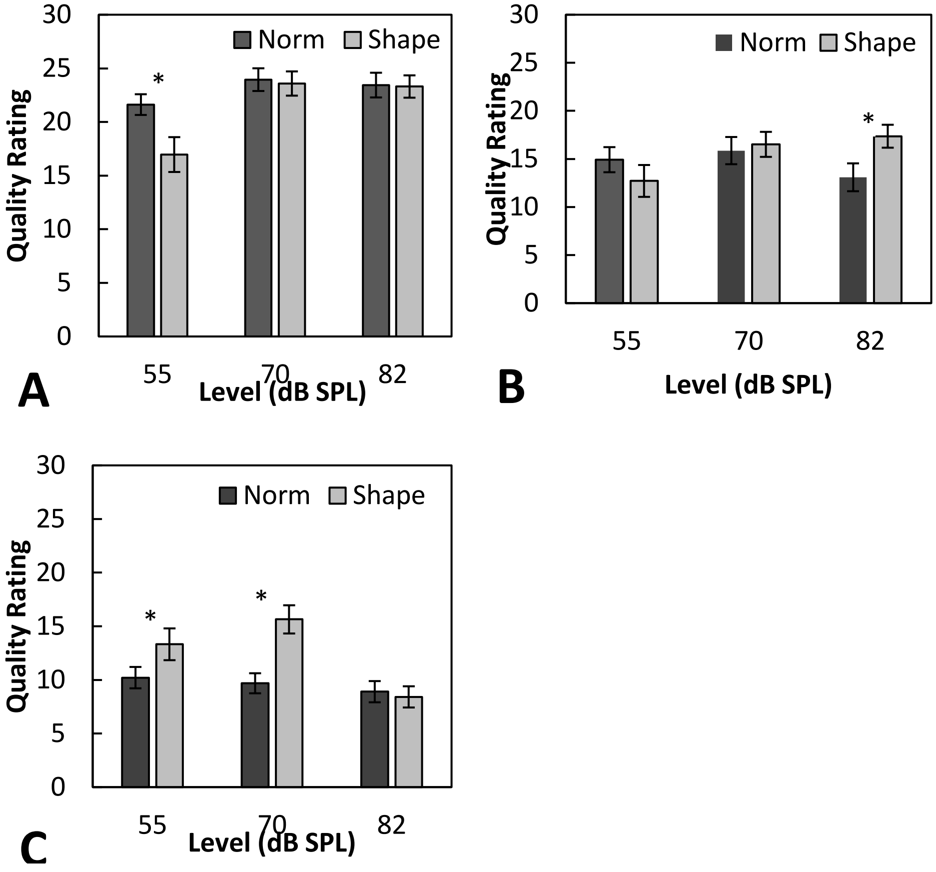

Figure 4.1. Sound quality ratings. Shown for the quiet (panel A), SMN (panel B) and SSN (panel C) background conditions. Ratings for shaped and normal speech conditions are displayed at 55, 70, and 82 dB SPL. Error bars = standard error of the mean. Asterisks indicate significantly different ratings between the two speech types.

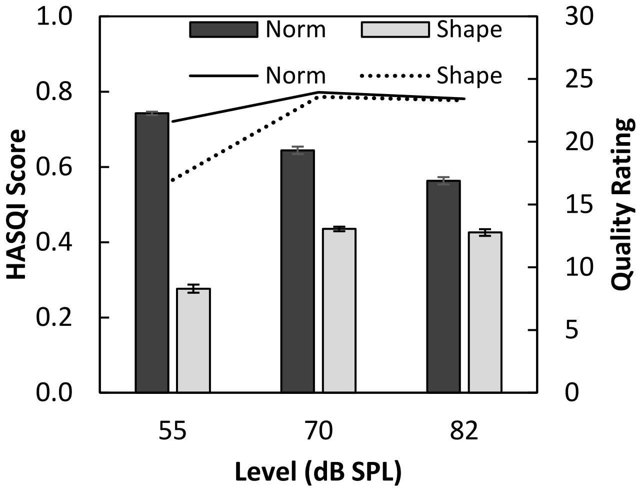

Figure 4.2. HASQI Scores (bars) and quality ratings (lines) for the quiet background condition. Scores for shaped and normal speech conditions are displayed at 55, 70, and 82 dB SPL. Error bars = standard error of the mean

CHAPTER 5

GENERAL DISCUSSION

The primary purpose of this project was to understand the relationship between spectral shape and speech level by assessing speech recognition, listening effort, and sound quality. In quiet, normal speech was generally associated with better recognition, listening effort, and sound quality outcomes with a few exceptions (as displayed in Figure 5.1, first column). In contrast, in both SMN and SSN backgrounds, spectrally shaped speech was associated with better speech recognition, listening effort, and sound quality outcomes (as displayed in Figure 5.1, second and third columns).

These results will be interpreted using two theories. First, the simple peripheral hypothesis (Humes, 1996) suggests that audibility dictates perceptual processing demands. This would indicate that shaped speech would have better outcomes due to increased audibility, particularly as level or glimpsing opportunities increase. This was evidenced by glimpse analyses for shaped speech. Second, ELU (Ronnberg et al., 2017) states that the amount of acoustic mismatch between a stimulus and its mental representation results in greater perceptual processing demands. This would indicate that greater distortion of a stimulus, such as through decreased audibility or signal processing of spectral shaping, could conceivably result in poorer outcomes due to increased perceptual processing demands. Both of these hypotheses have implications for the perceptual processing demands of the stimulus. In order to better understand if perceptual processing demands increase, or decrease, as a result of spectral shaping, a behavioral outcome could be used to reflect these demands. According to FUEL (Pichora-Fuller et al., 2016), when the cognitive resources required by a task increases, the more effortful it becomes. This mechanism would suggest that greater perceptual processing would result in a greater perception of effort. The results of this project will be discussed in terms of these theories below.

Perceptual processing requirements appear to have varied according to background condition. In quiet, shaped speech was associated with poorer outcomes than normal speech. This result is supported by ELU (Ronnberg et al., 2017), which suggests that a larger mismatch, or distortion, leads to increased processing resources. Because of this distortion, greater processing resources are needed to match the stimulus to its stored lexical representation. This increase in perceptual processing requirements can be demonstrated by poorer speech recognition outcomes in quiet, as well as greater subjective listening effort in quiet. This increase in perceptual processing requirements may be driven by the amount of distortion introduced when speech is spectrally shaped. The extent of this distortion was demonstrated through poorer perceived sound quality with spectrally shaped speech in quiet and reduced HASQI scores derived by acoustic analysis. This evidence suggests that spectral shaping increases perceptual processing requirements needed to understand speech when audibility of the signal is controlled. This is consistent with the simple peripheral hypothesis (Humes, 1996), as spectral shaping did not improve audibility according to SII measurements. However, because young normal hearing listeners were tested in this project, audibility may have been less of a factor in quiet environments than would be expected for listeners with hearing impairment.

In the two noise backgrounds, spectrally shaped speech was associated with better recognition, listening effort, and sound quality outcomes in comparison to normal speech. These results are consistent with both the simple peripheral hypothesis (Humes, 1996) and ELU (Ronnberg et al., 2017). According to the simple peripheral hypothesis, when speech is amplified, audibility at higher frequencies is increased although overall audibility was the same according to SII measurements in this study with young normal hearing listeners. Greater audibility reduces perceptual processing demands, which results in lower subjective listening effort (Pichora-Fuller et al., 2016). In addition, acoustic measurements indicated that spectrally shaped speech, when presented in noise, had an increased proportion of glimpses and better SNRs within these glimpses in comparison to normal speech. This was especially evident in fluctuating maskers, where listeners were better able to take advantage of glimpses in the masker, resulting in better speech recognition. According to ELU (Ronnberg et al., 2017), more favorable glimpsing opportunities could facilitate lexical retrieval as there is more acoustic information available to match the stimulus to a stored mental representation. These theories were supported experimentally by better speech recognition outcomes, as well as reduced listening effort, in noise. Finally, subjective sound quality was greater for spectrally shaped speech than normal speech in noise. This indicates that the perceived amount of distortion was lower in these conditions. A less distorted signal could facilitate lexical retrieval, as the acoustic mismatch between the stimulus and the mental representation was reduced as a result of spectral shaping (Ronnberg et al., 2017). This reduction in acoustic mismatch lowers processing demands required to resolve speech, resulting in better speech recognition and listening effort outcomes.

The effect of overall level also varied according to background condition. In quiet, increasing overall level resulted in better accuracy, higher quality ratings, and lower listening effort ratings. As level increases, audibility increases, which according to the simple peripheral hypothesis (Humes, 1996) results in better speech recognition outcomes. In addition, as overall level increases, less perceptual processing resources are required to match the acoustic signal to its lexical representation (Ronnberg et al., 2017), which according to FUEL (Pichora-Fuller et al., 2016), reduces perceived listening effort. In SMN, similarly to the quiet condition, mid- to high levels were often associated with better speech recognition and higher sound quality ratings. However, greater listening effort was also associated with higher levels for shaped speech in SMN. This could indicate that, while accuracy and quality remain better at higher levels in SMN, more perceptual processing demands may be needed to understand speech, resulting in a greater perception of listening effort. In contrast to quiet and SMN conditions, highest overall levels in SSN were associated with poorer speech recognition, quality ratings, and increased listening effort. At higher speech levels, decreases in performance could be due to lower frequency selectivity at high intensities, wider auditory filters, as well as upward spread of masking from low frequency amplification (as discussed in Dubno et al., 2012). These factors may have increased the amount of distortion perceived by the listener, as evidenced by poorer sound quality ratings. This distortion introduced at high intensity levels could increase the perceptual processing requirements needed to resolve high intensity speech, which could result in a perception of greater effort (Pichora-Fuller et al., 2016).

Correlations

In order to understand how speech recognition, listening effort, and sound quality were related, Pearson's correlations were conducted with data from ten participants who completed all three experiments. RT data was not included in this analysis. Data was collapsed across spectral condition, level, and background condition. All outcome measures were significantly correlated with each other (Figure 5.2, $p < .05$). When recognition was higher, listening effort tended to be lower. In addition, when recognition was higher, sound quality ratings also tended to be higher. This indicates that better sound quality and lower perceived listening effort may facilitate speech recognition outcomes. In addition, when sound quality ratings were higher, listening effort tended to be lower. This indicates that the perceived quality of speech may be related to how much listening effort the listener experiences.

While speech recognition, listening effort, and sound quality measurements were moderately correlated across participants, further correlations were conducted in order to understand how these measures were related on an individual level. Individuals with the highest and lowest mean scores within an outcome measure were examined. Results of this analysis are displayed in Figure 5.3. Row 1 displays two individuals who differed on listening effort, but happened to have similar outcomes in sound quality and speech recognition. Row 2 displays two individuals who differed on sound quality, but also happened to have similar outcomes in listening effort and speech recognition. Finally, row 3 displays two individuals whose means differed the most for speech recognition. This analysis illustrates that although outcome measurements were moderately correlated across participants, on an individual level, these measures can be used to dissociate individual listener characteristics. This is significant clinically, as even if two HA users perform similarly on a speech recognition task, it is possible that one listener is experiencing more effort, or lower perceived sound quality, than the other. Further research is needed to examine if these individual differences could impact the amount of perceived benefit experienced by the user, as well as likelihood of adherence.

Summary and Conclusions

Spectral shaping, determined based on mean audiograms from a group of older listeners with hearing impairment, facilitates speech perception in noise for young normal-hearing listeners. Furthermore, spectral shaping appears to mediate perceptual processing requirements needed to understand speech in noise, as evidenced by recognition, listening effort, and sound quality outcomes. However, in quiet, spectral shaping may limit the benefit obtained from amplification, as evidenced by poorer outcomes with spectrally shaped speech in young normal hearing listeners. But, the magnitude of these effects is difficult to assess due to ceiling effects. Thus, while older hearing-impaired individuals may find benefit in quiet, due to increased audibility (Humes, 2013), this benefit might be limited as indicated by these results. It is possible that amplification introduces acoustic distortion, which hinders lexical retrieval. This was evidenced by poorer sound quality and listening effort outcomes with shaped speech in quiet, where audibility was not a factor. In addition, combined outcomes of recognition, listening effort, and sound quality appear to differ across quiet, SMN, and SSN background types. This is clinically significant, as one of the most commonly reported reasons for non-use of HAs is ineffectiveness in noisy situations. The effects of spectral shape in different noise conditions needs to be further examined in the hearing-impaired population, where audibility and sensation level differ in comparison to young normal hearing listeners. The results of this project help to explain the underlying perceptual model utilized when understanding amplified speech, which could inform hearing aid user’s subjective experiences as well as audiological outcomes.

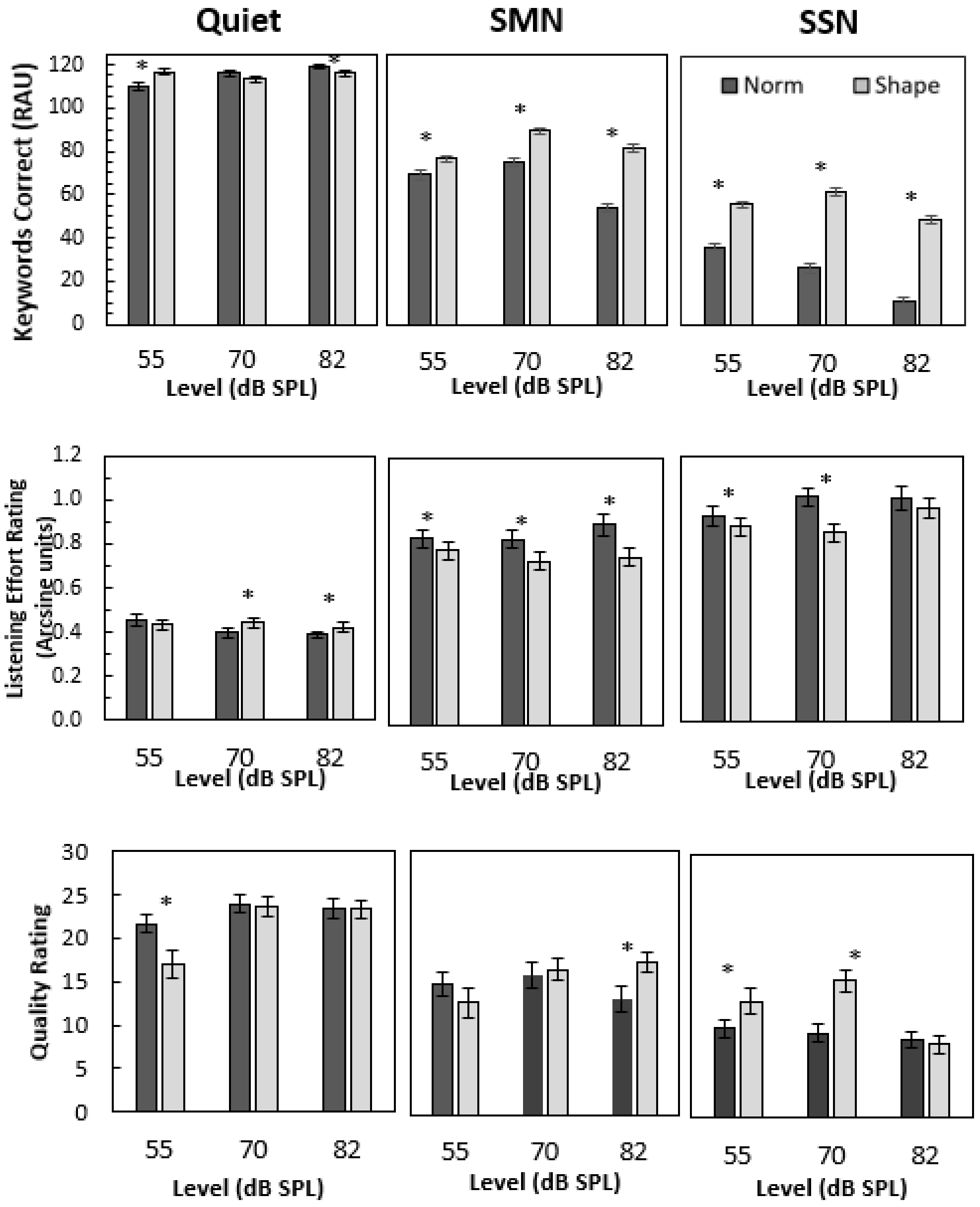

Figure 5.1. Summary outcomes. Outcomes are shown across the three experiments corresponding to each row. Keyword recognition accuracy (row 1), listening effort ratings (row 2) and sound quality ratings (row 3) for normal (dark grey) and shaped (light grey) speech at three different levels in quiet (column 1), SMN (column 2), and SSN (column 3). Error bars = standard error of the mean. Asterisks indicate significantly different performance between the two speech types (p<.05).

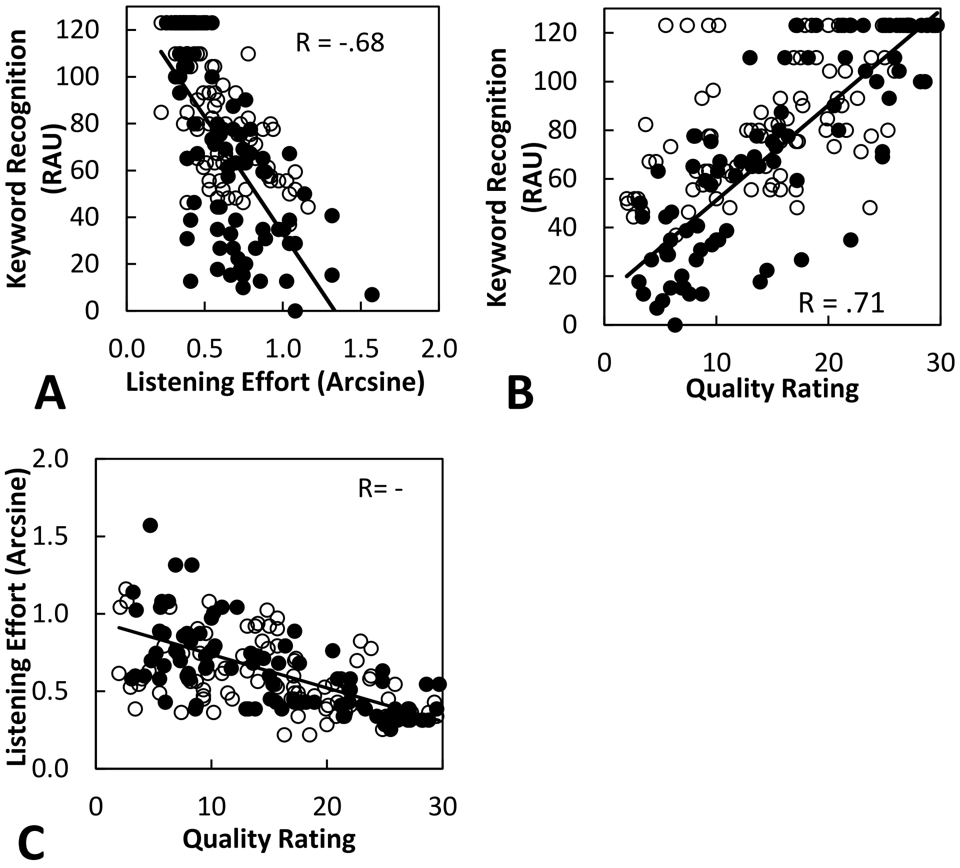

Figure 5.2. Scatterplots across measures. Scatterplots of the association between listening effort and speech recognition (panel A) speech recognition and sound quality (panel B), and sound quality and listening effort (panel C) are shown. Normal conditions are indicated by filled circles, and shaped conditions are indicated by unfilled circles.

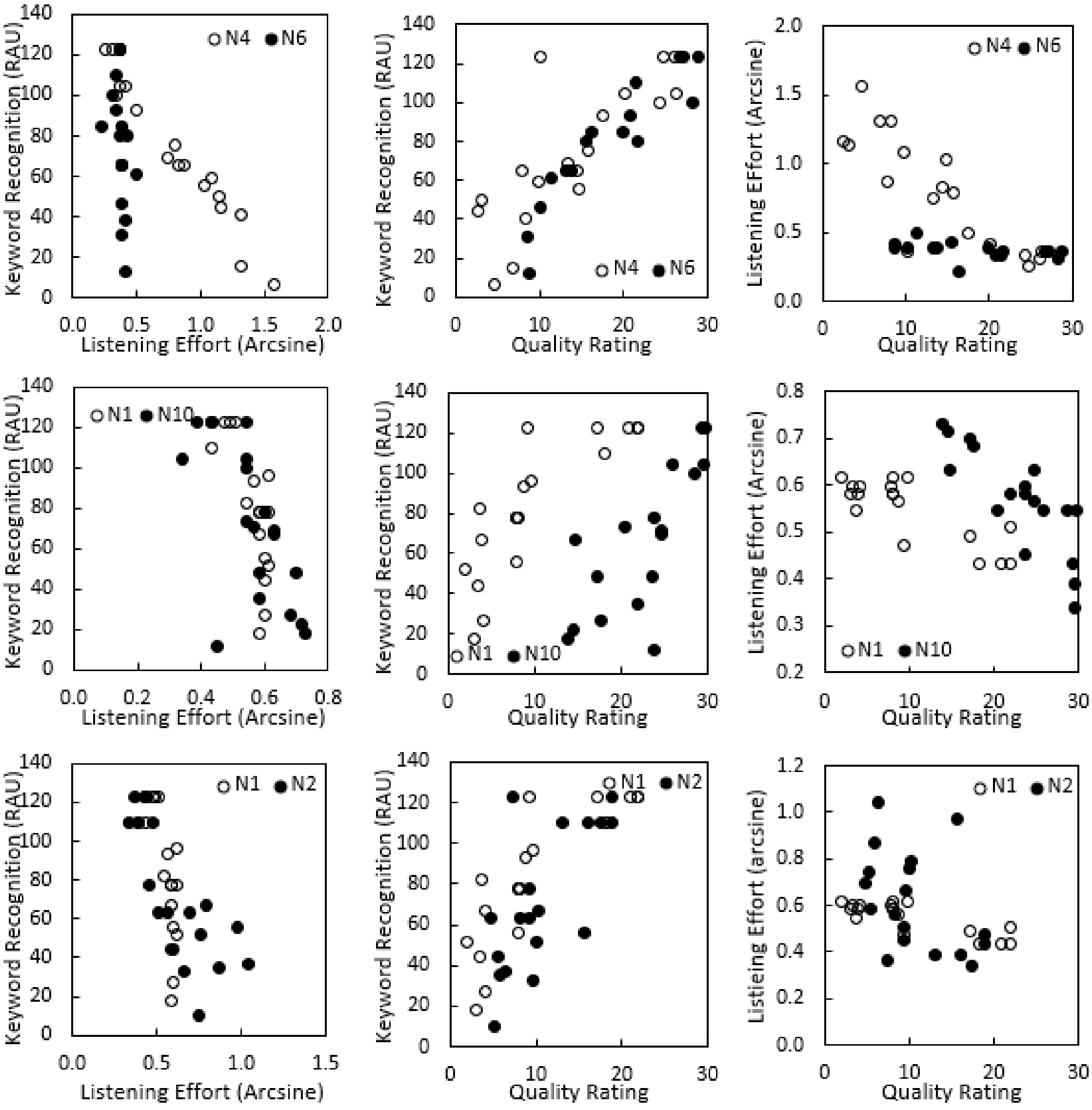

Figure 5.3. Individual correlations. Scatterplots of the association between recognition and listening effort (column 1) recognition and sound quality (column 2), and sound quality and listening effort (column 3) for pairs of individual listeners.

REFERENCES

Ahlstrom, J. B., Horwitz, A. R., & Dubno, J. R. (2014). Spatial separation benefit for unaided and aided listening. *Ear and Hearing*, *35*(1). https://doi.org/10.1097/AUD.0b013e3182a02274

American National Standards Institute. (1997). *American National Standard: Methods for Calculation of the Speech Intelligibility Index*. Acoustical Society of America.

Amos, N. E., & Humes, L. E. (2007). Contribution of High Frequencies to Speech Recognition in Quiet and Noise in Listeners With Varying Degrees of High-Frequency Sensorineural Hearing Loss. *Journal of Speech, Language, and Hearing Research*, *50*(4), 819–834. https://doi.org/10.1044/1092-4388(2007/057)

Anderson, M. (2010). The role of temporal fine structure in sound quality perception. *Speech, Language, and Hearing Sciences Graduate Theses & Dissertations*. Retrieved from https://scholar.colorado.edu/slhs_gradetds/3

Arehart, K. H., Kates, J. M., & Anderson, M. C. (2010). Effects of Noise, Nonlinear Processing, and Linear Filtering on Perceived Speech Quality. *Ear and Hearing*, *31*(3), 420. https://doi.org/10.1097/AUD.0b013e3181d3d4f3

Bacon, S. P., Opie, J. M., & Montoya, D. Y. (1998). The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *Journal of Speech, Language, and Hearing Research*, *41*(3), 549-563.

Baer, T., Moore, B. C. J., & Kluk, K. (2002). Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *The Journal of the Acoustical Society of America*, *112*(3), 1133–1144. https://doi.org/10.1121/1.1498853

Bentler, R., Wu, Y. H., Kettel, J., & Hurtig, R. (2008). Digital noise reduction: Outcomes from laboratory and field studies. *International Journal of Audiology*, *47*(8), 447-460.

Boersma, P., & Weenink, D. (2007). Praat (Version 4.5. 25)[Software]. *Latest version available for download from www. praat. org*.

Bertoli, S., Staehelin, K., Zemp, E., Schindler, C., Bodmer, D., & Probst, R. (2009). Survey on hearing aid use and satisfaction in Switzerland and their determinants. *International journal of audiology*, *48*(4), 183-195.

Bronkhorst, A., & Plomp, R. (1992). Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *The Journal of the Acoustical Society of America*, *92*(6), 3132-3139.

Cooke, M. (2003). Glimpsing speech. *Journal of Phonetics*, *31*(3-4), 579-584.

Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, *119*(3), 1562-1573.

Cox, R. M., & McDaniel, D. M. (1984). Intelligibility ratings of continuous discourse: Application to hearing aid selection. *The Journal of the Acoustical Society of America*, *76*(3), 758-766.

Cox, R. M., & McDaniel, D. M. (1989). Development of the Speech Intelligibility Rating (SIR) test for hearing aid comparisons. *Journal of Speech, Language, and Hearing Research*, *32*(2), 347-352.

Davies-Venn, E., Souza, P., & Fabry, D. (2007). Speech and Music Quality Ratings for Linear and Nonlinear Hearing Aid Circuitry. *Journal of the American Academy of Audiology*, *18*(8), 688–699. https://doi.org/10.3766/jaaa.18.8.6

Desjardins, J. L., & Doherty, K. A. (2013). Age-Related Changes in Listening Effort for Various Types of Masker Noises: *Ear and Hearing*, *34*(3), 261–272. https://doi.org/10.1097/AUD.0b013e31826d0ba4

Dubno, J. R., Ahlstrom, J. B., Wang, X., & Horwitz, A. R. (2012). Level-Dependent Changes in Perception of Speech Envelope Cues. *Journal of the Association for Research in Otolaryngology*, *13*(6), 835–852. https://doi.org/10.1007/s10162-012-0343-2

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2005). Recognition of filtered words in noise at higher-than-normal levels: Decreases in scores with and without increases in masking. *The Journal of the Acoustical Society of America*, *118*(2), 923–933. https://doi.org/10.1121/1.1953127

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2003). Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing. *The Journal of the Acoustical Society of America*, 113(4), 2084-2094.

Eisenberg, L. S., Dirks, D. D., & Bell, T. S. (1995). Speech Recognition in Amplitude-Modulated Noise of Listeners With Normal and Listeners With Impaired Hearing.

Journal of Speech, Language, and Hearing Research, *38*(1), 222–233. https://doi.org/10.1044/jshr.3801.222

Eisenberg, L. S., Dirks, D. D., Takayanagi, S., & Martinez, A. S. (1998). Subjective judgments of clarity and intelligibility for filtered stimuli with equivalent speech intelligibility index predictions. *Journal of Speech, Language, and Hearing Research*, *41*(2), 327-339.

Fastl, H. (2005). Psycho-Acoustics and Sound Quality. In J. Blauert (Ed.), *Communication Acoustics* (pp. 139–162). https://doi.org/10.1007/3-540-27437-5_6

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, *88*(4), 1725-1736.

Fogerty, D., Ahlstrom, J. B., Bologna, W. J., & Dubno, J. R. (2015a). Segmental interference by the masker modulation spectrum in sentences: Effects of age and hearing loss. *The Journal of the Acoustical Society of America*, *137*(4), 2435–2435. https://doi.org/10.1121/1.4920889

Fogerty, D., Ahlstrom, J. B., Bologna, W. J., & Dubno, J. R. (2015b). Sentence intelligibility during segmental interruption and masking by speech-modulated noise: Effects of age and hearing loss. *The Journal of the Acoustical Society of America*, *137*(6), 3487–3501. https://doi.org/10.1121/1.4921603

Fogerty, D., Bologna, W. J., Ahlstrom, J. B., & Dubno, J. R. (2017). Simultaneous and forward masking of vowels and stop consonants: Effects of age, hearing loss, and spectral shaping. *The Journal of the Acoustical Society of America*, *141*(2), 1133–1143. https://doi.org/10.1121/1.4976082

Fogerty, D., & Humes, L. E. (2012). A correlational method to concurrently measure envelope and temporal fine structure weights: Effects of age, cochlear pathology, and spectral shaping. *The Journal of the Acoustical Society of America*, *132*(3), 1679–1689. https://doi.org/10.1121/1.4742716

Fogerty, D., Montgomery, A. A., & Crass, K. A. (2014). Effect of initial-consonant intensity on the speed of lexical decisions. *Attention, Perception, & Psychophysics*, *76*(3), 852–863. https://doi.org/10.3758/s13414-014-0624-4

Franks, J. R. (1982). Judgments of hearing aid processed music. *Ear and Hearing*, *3*(1), 18-23.

Fraser, S., Gagné, J. P., Alepins, M., & Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. *Journal of Speech, Language, and Hearing Research*.

French, N. R., & Steinberg, J. C. (1945). Factors governing the intelligibility of speech sounds. *The Journal of the Acoustical Society of America*, *17*(1), 103-103.

French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *The journal of the Acoustical society of America*, *19*(1), 90-119.

Freyman, R. L., Nerbonne, G. P., & Cote, H. A. (1991). Effect of consonant-vowel ratio modification on amplitude envelope cues for consonant recognition. *Journal of Speech, Language, and Hearing Research*, *34*(2), 415-426.

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

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

Gianopoulos, I., Stephens, D., & Davis, A. (2002). Follow up of people fitted with hearing aids after adult hearing screening: the need for support after fitting. *Bmj*, *325*(7362), 471.

Gopinath, B., Schneider, J., Hartley, D., Teber, E., McMahon, C. M., Leeder, S. R., & Mitchell, P. (2011). Incidence and predictors of hearing aid use and ownership among older adults with hearing loss. *Annals of epidemiology*, *21*(7), 497-506.

Gosselin, P. A., & Gagné, J.-P. (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. https://doi.org/10.1044/1092-4388(2010/10-0069)

Gustafsson, H. Å., & Arlinger, S. D. (1994). Masking of speech by amplitude-modulated noise. *The Journal of the Acoustical Society of America*, *95*(1), 518-529.

Gustafson, S., McCreery, R., Hoover, B., Kopun, J. G., & Stelmachowicz, P. (2014). Listening effort and perceived clarity for normal hearing children with the use of digital noise reduction. *Ear and Hearing*, *35*(2), 183.

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

Hartley, D., Rochtchina, E., Newall, P., Golding, M., & Mitchell, P. (2010). Use of hearing aids and assistive listening devices in an older Australian population. *Journal of the American Academy of Audiology*, *21*(10), 642-653.

Hawkins, J. E., & Stevens, S. S. (1950). The Masking of Pure Tones and of Speech by White Noise. *The Journal of the Acoustical Society of America*, *22*(1), 6–13. https://doi.org/10.1121/1.1906581

Helfer, K. S., Chevalier, J., & Freyman, R. L. (2010). Aging, spatial cues, and single-versus dual-task performance in competing speech perception. *The Journal of the Acoustical Society of America*, *128*(6), 3625–3633. https://doi.org/10.1121/1.3502462

Hickson, L., & Byrne, D. (1997). Consonant perception in quiet: Effect of increasing the consonant-vowel ratio with compression amplification. *JOURNAL-AMERICAN ACADEMY OF AUDIOLOGY*, *8*, 322-332.

Hogan, C. A., & Turner, C. W. (1998). High-frequency audibility: Benefits for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *104*(1), 432-441.

Hornsby, B. W. Y. (2013). The Effects of Hearing Aid Use on Listening Effort and Mental Fatigue Associated With Sustained Speech Processing Demands. *Ear and Hearing*, *34*(5), 523. https://doi.org/10.1097/AUD.0b013e31828003d8

Hornsby, B. W. Y., & Ricketts, T. A. (2003). The effects of hearing loss on the contribution of high- and low-frequency speech information to speech understanding. *The Journal of the Acoustical Society of America*, *113*(3), 1706–1717. https://doi.org/10.1121/1.1553458

Hornsby, B. W., & Ricketts, T. A. (2006). The effects of hearing loss on the contribution of high-and low-frequency speech information to speech understanding. II. Sloping hearing loss. *The Journal of the Acoustical Society of America*, *119*(3), 1752-1763.

Horwitz, A. R., Ahlstrom, J. B., & Dubno, J. R. (2008). Factors Affecting the Benefits of High-Frequency Amplification. *Journal of Speech, Language & Hearing Research*, *51*(3), 798–813. https://doi.org/10.1044/1092-4388(2008/057)

Houben, R., Doorn-Bierman, M. van, & Dreschler, W. A. (2013). Using response time to speech as a measure for listening effort. *International Journal of Audiology*, *52*(11), 753–761. https://doi.org/10.3109/14992027.2013.832415

Howard-Jones, P. A., & Rosen, S. (1993). The perception of speech in fluctuating noise. *Acta Acustica united with Acustica*, *78*(5), 258-272.

Humes, L. E. (1991). Understanding the speech-understanding problems of the hearing impaired. *Journal of the American Academy of Audiology*, *2*(2), 59-69.

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

Humes, L. E. (2013). Understanding the Speech-Understanding Problems of Older Adults. *American Journal of Audiology*, *22*(2), 303. https://doi.org/10.1044/1059-0889(2013/12-0066)

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

Jensen, K. K., & Bernstein, J. G. (2019). The fluctuating masker benefit for normal-hearing and hearing-impaired listeners with equal audibility at a fixed signal-to-noise ratio. *The Journal of the Acoustical Society of America*, *145*(4), 2113-2125.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kates, J. M., & Arehart, K. H. (2014a). The Hearing-Aid Speech Perception Index (HASPI). *Speech Communication*, *65*, 75–93. https://doi.org/10.1016/j.specom.2014.06.002

Kates, J. M., & Arehart, K. H. (2014b). The Hearing-Aid Speech Quality Index (HASQI) Version 2. *Journal of the Audio Engineering Society*, *62*(3), 99–117.

Kates, J. M., Arehart, K. H., Anderson, M. C., Kumar Muralimanohar, R., & Harvey, L. O. (2018). Using Objective Metrics to Measure Hearing Aid Performance: *Ear and Hearing*, 1. https://doi.org/10.1097/AUD.0000000000000574

Kochkin, S. (2000). MarkeTrak V:"Why my hearing aids are in the drawer" The consumers' perspective. *The Hearing Journal*, *53*(2), 34-36.

Kochkin, S. (2002). 10-year customer satisfaction trends in the US hearing instrument market. *Hearing Review*, *9*(10), 14-25.

Korczak, P. A., Kurtzberg, D., & Stapells, D. R. (2005). Effects of sensorineural hearing loss and personal hearing aids on cortical event-related potential and behavioral measures of speech-sound processing. *Ear and hearing*, *26*(2), 165-185.

Krueger, M., Schulte, M., Brand, T., & Holube, I. (2017). Development of an adaptive scaling method for subjective listening effort. *The Journal of the Acoustical Society of America*, *141*(6), 4680–4693. https://doi.org/10.1121/1.4986938

Krueger, M., Schulte, M., Zokoll, M. A., Wagener, K. C., Meis, M., Brand, T., & Holube, I. (2017). Relation Between Listening Effort and Speech Intelligibility in Noise.

American Journal of Audiology, *26*, 378–392. https://doi.org/10.1044/2017_AJA-16-0136

Kulkarni, P. N., Pandey, P. C., & Jangamashetti, D. S. (2012). Multi-band frequency compression for improving speech perception by listeners with moderate sensorineural hearing loss. *Speech Communication*, *54*(3), 341-350.

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

Li, N., & Loizou, P. C. (2007). Factors influencing glimpsing of speech in noise. *The Journal of the Acoustical Society of America*, *122*(2), 1165-1172.

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

Lunner, T., Hellgren, J., Arlinger, S., & Elberling, C. (1997). A digital filterbank hearing aid: Three digital signal processing algorithms-User preference and performance. *Ear and hearing*, *18*(5), 373-387.

Magnusson, L., Claesson, A., Persson, M., & Tengstrand, T. (2013). Speech recognition in noise using bilateral open-fit hearing aids: The limited benefit of directional microphones and noise reduction. *International Journal of Audiology*, *52*(1), 29-36.

McCormack, A., & Fortnum, H. (2013). Why do people fitted with hearing aids not wear them?. *International journal of audiology*, *52*(5), 360-368.

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & 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. https://doi.org/10.3109/14992027.2014.890296

Miller, G. A. (1947). The masking of speech. *Psychological bulletin*, *44*(2), 105.

Miller, G. A., & Licklider, J. C. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, *22*(2), 167-173.

Moore, B. C. J., Füllgrabe, C., & Stone, M. A. (2011). Determination of Preferred Parameters for Multichannel Compression Using Individually Fitted Simulated

Hearing Aids and Paired Comparisons: *Ear and Hearing*, *32*(5), 556–568. https://doi.org/10.1097/AUD.0b013e31820b5f4c

National Institute on Deafness and Other Communication Disorders (NIDCD). (2010). Quick statistics. Retrieved from http://www.nidcd.nih.gov/health/ statistics/Pages/quick.aspx

D. A., Schroder, A. C., & Wojtczak, M. (2001). A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *110*(4), 2045-2064.

Neuman, A. C., Bakke, M. H., Hellman, S., & Levitt, H. (1994). Effect of compression ratio in a slow-acting compression hearing aid: Paired-comparison judgments of quality. *The Journal of the Acoustical Society of America*, *96*(3), 1471–1478. https://doi.org/10.1121/1.410289

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1998). The effect of compression ratio and release time on the categorical rating of sound quality. *The Journal of the Acoustical Society of America*, *103*(5), 2273–2281. https://doi.org/10.1121/1.422745

Pals, C., Sarampalis, A., & Başkent, D. (2013). Listening Effort With Cochlear Implant Simulations. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1075–1084. https://doi.org/10.1044/1092-4388(2012/12-0074)

Pals, C., Sarampalis, A., van Rijn, H., & Başkent, D. (2015). Validation of a simple response-time measure of listening effort. *The Journal of the Acoustical Society of America*, *138*(3), EL187–EL192. https://doi.org/10.1121/1.4929614

Parsa, V., Scollie, S., Glista, D., & Seelisch, A. (2013). Nonlinear Frequency Compression: Effects on Sound Quality Ratings of Speech and Music. *Trends in Amplification*, *17*(1), 54–68. https://doi.org/10.1177/1084713813480856

Perez-Gonzalez, P., & Lopez-Poveda, E. A. (2013, June). Level-dependent effects on speech intelligibility. In *Proceedings of Meetings on Acoustics ICA2013* (Vol. 19, No. 1, p. 050065). ASA.

Phatak, S. A., & Grant, K. W. (2012). Phoneme recognition in modulated maskers by normal-hearing and aided hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 132(3), 1646-1654.

Phillips, D. P., Rappaport, J. M., & Gulliver, J. M. (1994). Impaired word recognition in noise by patients with noise-induced cochlear hearing loss: contribution of temporal resolution defect. *The American journal of otology*, *15*(5), 679-686.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., … Wingfield, A. (2016). *Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL)*. *37*, 23.

Picou, E. M., Ricketts, T. A., & Hornsby, B. W. Y. (2013). How Hearing Aids, Background Noise, and Visual Cues Influence Objective Listening Effort: *Ear and Hearing*, *34*(5), e52–e64. https://doi.org/10.1097/AUD.0b013e31827f0431

Popelka, M. M., Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E., & Klein, R. (1998). Low prevalence of hearing aid use among older adults with hearing loss: the Epidemiology of Hearing Loss Study. *Journal of the American Geriatrics Society*, *46*(9), 1075-1078.

Pollack, I., & Pickett, J. M. (1958). Masking of Speech by Noise at High Sound Levels. *The Journal of the Acoustical Society of America*, *30*(2), 127–130. https://doi.org/10.1121/1.1909503

Ricketts, T. A., Dittberner, A. B., & Johnson, E. E. (2008). High-Frequency Amplification and Sound Quality in Listeners With Normal Through Moderate Hearing Loss. *Journal of Speech, Language, and Hearing Research*, *51*(1), 160–172. https://doi.org/10.1044/1092-4388(2008/012)

Ronnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology*, *42*, S68-S76.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., … Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*. https://doi.org/10.3389/fnsys.2013.00031

Rothauser, E. H. (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans. on Audio and Electroacoustics*, *17*, 225-246.

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. https://doi.org/10.1044/1092-4388(2009/08-0111)

Shannon, R. V., Zeng, F. G., & Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. *The Journal of the Acoustical Society of America*, *104*(4), 2467-2476.

Simpson, A., McDermott, H. J., & Dowell, R. C. (2005). Benefits of audibility for listeners with severe high-frequency hearing loss. *Hearing Research*, *210*(1), 42–52. https://doi.org/10.1016/j.heares.2005.07.001

Skinner, M. W. (1980). Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation. *The Journal of the Acoustical Society of America*, *67*(1), 306-317.

Studebaker, G. A. (1985). A" rationalized" arcsine transform. *Journal of Speech, Language, and Hearing Research*, *28*(3), 455-462.

Slawson, W. (1985). *Sound color*. Yank Gulch Music.

Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., & Gwaltney, C. A. (1999). Monosyllabic word recognition at higher-than-normal speech and noise levels. *The Journal of the Acoustical Society of America*, *105*(4), 2431–2444. https://doi.org/10.1121/1.426848

Sullivan, J. A., Allsman, C. S., Nielsen, L. B. M., & Mobley, J. P. M. (1992). Amplification for Listeners with Steeply Sloping, High-Frequency Hearing Loss. *Ear & Hearing*, *13*(1), 35–45.

Summers, V., & Cord, M. T. (2007). Intelligibility of speech in noise at high presentation levels: Effects of hearing loss and frequency region. *The Journal of the Acoustical Society of America*, *122*(2), 1130–1137. https://doi.org/10.1121/1.2751251

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

Takahashi, G., Martinez, C. D., Beamer, S., Bridges, J., Noffsinger, D., Sugiura, K., ... & Williams, D. W. (2007). Subjective measures of hearing aid benefit and satisfaction in the NIDCD/VA follow-up study. *Journal of the American Academy of Audiology*, *18*(4), 323-349.

Tomita, M., Mann, W. C., & Welch, T. R. (2001). Use of assistive devices to address hearing impairment by older persons with disabilities. *International journal of rehabilitation research*, *24*(4), 279-290.v

Turner, C. W., & Henry, B. A. (2002). Benefits of amplification for speech recognition in background noise. *The Journal of the Acoustical Society of America*, *112*(4), 1675–1680. https://doi.org/10.1121/1.1506158

Upfold, L. J., & Wilson, D. A. (1980). *Hearing aid distribution and use in Australia: The Australian Bureau of Statistics 1978 survey*. Audiological Society of Australia.

Vuorialho, A., Karinen, P., & Sorri, M. (2006). Counselling of hearing aid users is highly cost-effective. *European Archives of Oto-Rhino-Laryngology and Head & Neck*, *263*(11), 988-995.

Weiss, C. E. (1973). Why more of the aged with auditory deficits do not wear hearing aids. *Journal of the American Geriatrics Society*, *21*(3), 139-141.

Wilson, R. H., & Carhart, R. (1969). Influence of pulsed masking on the threshold for spondees. *The Journal of the Acoustical Society of America*, *46*(4B), 998-1010.

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil Response as an Indication of Effortful Listening: The Influence of Sentence Intelligibility. *Ear and Hearing*, *31*(4), 480. https://doi.org/10.1097/AUD.0b013e3181d4f251

.