EVALUATION OF THE TELEPHONE SPEECH ENHANCEMENT ALGORITHM IN OLDER ADULTS USING INDIVIDUAL AUDIOGRAMS

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation *with research distinction* in Speech and Hearing Science in the undergraduate college of the Ohio State University

by

Jessica Stacy Freesen

****

The Ohio State University
June 2009

Project Advisors: Dr. Christina Roup & Dr. Lawrence Feth, Department of Speech and Hearing Sciences

Table of Contents

Abstract

After experiencing an influx of challenging telephone conversations with elderly individuals with hearing loss, counselors at the Franklin County Ohio Office on Aging called upon researchers in the Departments of Speech and Hearing Science and Electrical and Computer Engineering at The Ohio State University. They developed a Telephone Speech Enhancement Algorithm (TSEA) to help hearing impaired, older listeners with telephone communication problems. The TSEA is based on a hearing aid compression strategy. Specifically, the algorithm processes speech signals by applying amplification to soft sounds (e.g. consonants) without amplifying louder sounds (e.g. vowels) (Roup et al., In Review). This type of processing is known to improve speech understanding for those with hearing loss. The algorithm has been tested and revised over the past several years and is now implemented in a “real-time” format using technology developed in collaboration with FutureComm located in Gahanna, Ohio (Harhager, 2007). Past studies focused on an average audiogram, or default TSEA, derived from a set population of patients from the Columbus Speech and Hearing Center between 1996 and 2000. For this particular study, the goal was to test the significance of applying ten subjects’ individual audiograms to the TSEA as opposed to the default TSEA alone. The preliminary results of applying the individual audiogram to the TSEA varied significantly depending on the listener’s hearing loss. Depending on the loss, applying the individual’s personal audiogram to the original TSEA did not always allow for maximum performance on the telephone. If the individual’s hearing loss was closer to normal at certain frequencies than those of the average audiogram, the sound pressure level delivered over the telephone was lower. In general, the closer to normal the subject’s hearing was, the worse they performed with their individual

audiogram applied to the TSEA. Also, the preliminary results showed that the default TSEA had better percent improvement than when the individual audiogram was applied to the TSEA. Other factors that were not measured but potentially affected the results included: 1) the participant's motivation and interest in the test; 2) word recognition ability; and 3) overall communication experience and familiarity with the telephone.

Dedicated to My Grandfather, Oscar Robert “Bob” Freesen

I would like to thank Dr. Christina Roup and Dr. Lawrence Feth for the continued support and encouragement throughout this entire project.

CHAPTER 1

Introduction and Literature Review

For many older adults with hearing loss, telephone communication is extremely difficult. Reasons for its difficulty depends on the characteristics of the telephone and the individuals personal impairments. Such factors that contribute to the difficulty of telephone use includes the type, degree, and configuration of the individual's hearing loss, frequency range of the telephone, absence of visual cues, line noise that is present at the receiver, background noise, and monaural listening (Harhager, 2007). After experiencing an influx of challenging telephone conversations with elderly individuals with hearing loss, counselors at the Franklin County Ohio Office on Aging collaborated with researchers in the Departments of Speech and Hearing Science and Electrical and Computer Engineering at The Ohio State University to develop a Telephone Speech Enhancement Algorithm (TSEA) to help this communication problem. The TSEA is based on a hearing aid compression strategy. Specifically, the algorithm processes speech signals by applying amplification to soft sounds (e.g. consonants) without amplifying louder sounds (e.g. vowels) (Roup et al., In Review). This type of processing is known to improve speech understanding for those with hearing loss. The algorithm has been tested and revised over the past several years and is now implemented in a "real-time" format using hardware and software developed in collaboration with a local small business, FutureComm located in Gahanna, Ohio. (Harhager, 2007)

The current version of TSEA uses an average audiogram as a representative sample of hearing loss for this population. This average, however, may not be appropriate for all listeners. Specifically, an individual may have more or less hearing loss than the average loss used by the algorithm, potentially affecting the effectiveness of the algorithm. The TSEA has the capability of changing the reference audiogram. The purpose of the present study, therefore, is to use each individual participant's audiogram as the reference hearing loss for the TSEA. The amount of improvement in speech recognition from this "individualized" algorithm will be measured for each participant. The present study tested the hypothesis that the individual audiograms will be more successful at improving speech recognition, because of the attention to each participant's unique hearing loss.

1.1 Hearing Loss Overview

Hearing loss can be caused by several factors such as aging, noise exposure, ototoxicity, heredity, pregnancy complications, and various other conditions. The disorder can be described through four different types of hearing losses: conductive, sensorineural, mixed, and central. Conductive losses involve the abnormality or malfunction of the outer and middle ear structures. In a conductive loss, the ear has problems directing sounds to the inner ear and sounds are reduced. Types of conductive losses include cerumen build up, otitis media (fluid build up behind the ear drum), otitis externa (inflammation of the canal), otosclerosis (hardening of the ossicular chain), and many more. Sensorineural losses, however, involve the deterioration or damage of the outer and inner hair cells causing sounds to become distorted and specific frequency regions on the basilar membrane to be permanently lost. Types of sensorineural losses include trauma, viruses, presbycusis or aging, ototoxic medications, and genetics. Central hearing losses involved the impairment of the auditory nerve and/or processing of the sounds in the brain. These hearing disorders do not necessarily involve the loss of hearing but rather the difficulty in listening and/or processing. Therefore, the most common term used for these disorders is (central) auditory processing disorders. (Humes, 2003)

1.2 Hearing Loss and the Older Adult

Hearing loss has become an extremely prevalent disorder among the elderly (Heine & Browning, 2001). According to the United States Department of Health, Education, and Welfare, presbycusis (age-related hearing loss) is the fourth most prevalent major chronic disability among people 65 and older. With the life expectancy on the rise and the 65 and older age group growing more rapidly than any other age group, the number of older adults reporting a hearing loss is likely to increase (Heine & Browning, 2001; Harhager, 2007). It is estimated that 21 million older adults will have an age related hearing loss by the year 2030 (Garstecki,1996; Weinstein, 2000).

1.2A *Types and Characteristics of Presbycusis*

Presbycusis, literally meaning elderly hearing, is the universal term applied to age-related hearing loss due to the contributions of a lifetime of abuse to the auditory system (Gates & Mills, 2005). The term encompasses all conditions associated with hearing loss in elderly individuals. In general, its symptoms include a reduction in hearing sensitivity and speech understanding in noisy environments with low signal to noise ratios, slowed central processing of acoustic information, and impaired localization of sound sources (Gates & Mills, 2005). Consequently, the impairment causes difficulty with conversations and participation in social activities. The degree of presbycusis can range from mild to profound, causing the effects of the disorder to vary between each individual.

There are four classic types of presbycusis – sensory, strial, cochlear conductive, and neural – that can be found either independently or as a combination of the types (Gates & Mills, 2005). The most common first sign of presbycusis is a loss of threshold sensitivity in the high-frequency region of the hearing spectrum, typically 2-4 kHz region. Figure 1.2 illustrates an audiologic example of presbycusis and the common phonemic information lost due to the impairment. Strial (or metabolic) presbycusis is the eventual decrease of sensitivity in the lower frequency regions as well as the high frequency regions, characteristic of the declining metabolic functions of the strial vascularis, responsible for maintaining the cochlea's bioelectric and biochemical properties (Connelly, 2003). Sensory presbycusis is defined by the loss of outer hair cells on the basal end of the cochlea typically caused by constant or sudden exposure to noise. This type of loss is commonly illustrated on an audiogram with a notch near the 4 kHz region (Gates & Mills, 2005). Cochlear conductive presbycusis is defined as the unfavorable stiffness of the basilar membrane micromechanics that affects the membranes motion (Connelly, 2003). Neural presbycusis is described as a diminished quantity of auditory neurons in the cochlea and auditory pathways. This type of presbycusis is not evident until the number of normally functioning neurons falls below the minimum amount required to adequately transmit a signal (Connelly, 2003).

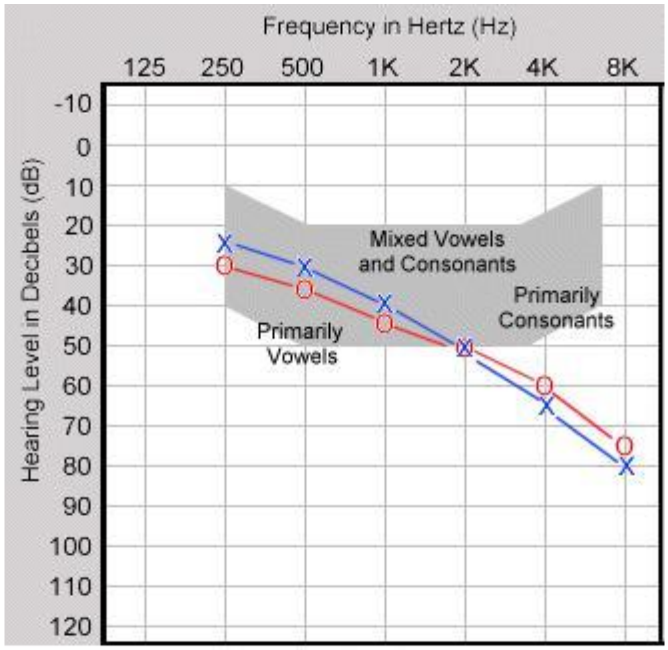


Figure 1.2: The “Speech Banana” – Typical Presbycusis Audiogram (Ross, 2004)

1.2B *Social Consequences of Hearing Loss*

Hearing loss causes many psychosocial, cognitive, emotional, and financial problems. According to Gates and Mills (2005), “many people regard presbycusis as an inevitable rite of passage into their senior years and are reluctant to seek help because of cost, vanity, and inconvenience” (pg. 1111). Depending on the degree and type, hearing loss can inhibit the individual from enjoying the things they often do, influencing their quality of life. It can lead to an increased sense of vulnerability and insecurity, a lack of self-esteem, depression, frequent exhaustion, and an inability to adjust to new circumstances and environments (Heine & Browning, 2002). Such symptoms often lead to the individual isolating themselves from social situations.

Hearing loss does not solely affect the individual; it can also affect the relationships in that person’s life. In a study conducted by Chmiel and Jerger (1993), it was concluded that an older adult with a hearing loss judges his/her impairment less harshly than his/her spouse or partner depending on the degree of loss and presence of a central lesion. Problems within relationships due to the existence of a hearing loss require significant aural rehabilitation, resulting in the purchase of a hearing aid or other assistive listening device (ALD) (Chmiel & Jerger, 1993).

1.3 Telephone Communication and Hearing Loss

Telephone use has become a major part of almost every household in the United States (Mormer & Mack, 2003). According to a census report, recent data analysis shows that the older householders have an increased chance of owning a telephone

(U.S. Census, 2000). More importantly, due to limited social opportunities that accompany aging, many elderly people rely heavily on telephone communication for maintaining social contact. Social contact for older adults is connected to the telephone, often becoming the long-distance tie to family and friends, allowing the individual to have a sense of connectedness and preservation of family relationships (Mormer & Mack, 2003). Communication over the telephone, however, is often difficult for older adults due to the presence of hearing loss. With the inability to depend entirely on auditory skills, visual cues, such as lipreading, becomes increasingly important to speech understanding for the elderly. Because a telephone is entirely auditory with no visual cues, many may experience difficulty communication over the telephone and therefore find themselves isolated from the social world. This is especially the case for the majority of elderly people with hearing loss that go through daily life without assistive listening devices or hearing aids (Lesner, 2003). Though it was meant to bring people closer together and simplify life’s basic communication restrictions, the invention of the telephone eventually lead to an emphasis on independence.

1.3A *The Telephone and the Individual*

In 1876 when Alexander Graham Bell invented the telephone, it was meant to be a tool that provided extra security when faced with unexpected emergencies. To those who could afford it, the role of the telephone was simple – a manmade appliance meant to help make life easier. As the pace of life began to speed up, the demand for telephone began to increase dramatically. According to Mack and Mormer (2003) the percent of phoneless households in the U.S. in 1960 was just over 20%. By 1990, however, the percent of phoneless households dropped to approximately 5%. Though the telephone still acted as a home security device, the motivation to own a telephone was more personal. Mack and Mormer (2003) discussed that the individual's motivations for owning a telephone are broken down into three primary categories: (1) sociability (purpose of social interaction), (2) instrumentality (tool into daily lives), and (3) reassurance (psychological aspects that provides emotional relief). Together these categories construct the well-being and quality of life for the modern American (Mack & Mormer, 2003).

1.3B *The Telephone and the Older Adult*

Before the telephone became a major component in American households, older family members lived with their immediate family. Since then, 10.5 million older Americans were living alone in 2003. The proportion varies greatly by age, with 29.6 percent ages 65 to 74, 47.6 percent ages 75 to 84, and 57 percent age 85 and older living alone (U.S. Census, 2006). When living alone, the telephone quickly becomes a major part of the individual's life. Not only does it keep the older person safe if they fall down while alone but it also keeps them in touch with family and friends. For the many older individuals with hearing loss, the telephone is an isolating factor in their lives rather than a social and safety outlet. An estimated 30-35 percent of adults between the ages of 65-75 and 40-45 percent of adults above the age of 75 years experience some degree of presbycusis (Smith, n.d.). Fueled by the aging of the pre-Baby Boomer generation, the total population of individuals 65 years of age and older has increased five percent from 2000 to 2005. Different segments within this particular age group, however, have changed more rapidly than others. For example, as shown in Figure 1.3, the 85 and older age group has increased in population by approximately 20 percent in the U.S. (U.S. Census, 2005). Therefore, the population increase leads to an increase in hearing loss prevalence among the older population.

The Franklin County Office on Aging (FCOA) created a newsletter to help keep older adults who live alone safe at home. Most of their services that aid in ensuring senior safety required the use of a telephone such as making sure a phone is located in bathrooms, kitchens, and by stairways, enabling the person to call in case he/she falls or injures him/herself. With a greater emphasis placed on telephone use and the increase in the older population who experience presbycusis, it can be concluded that there may be more people experiencing difficulty when using the telephone.

1.3C *The Telephone and Its Limitations*

The invention of the telephone was not beneficial to all Americans. For the hearing impaired, even those with mild losses, telephone use can be extremely frustrating. One of the biggest complaints among older adults with hearing loss is that when talking on the telephone speech understanding is extremely difficult, if not impossible. The difficulty center around the limitations of the telephone: (1) a reduced bandwidth, (2) a lack of visual cues, and (3) a reduced dynamic range (Roup et al, In Review).

According to Terry et al (1992), due to efficiency and economic reasons, the telephone mechanics exploits on the redundancy of the speech signal and transmits a limited bandwidth of the signal ranging from 300-3000 Hz. The speech spectrum,

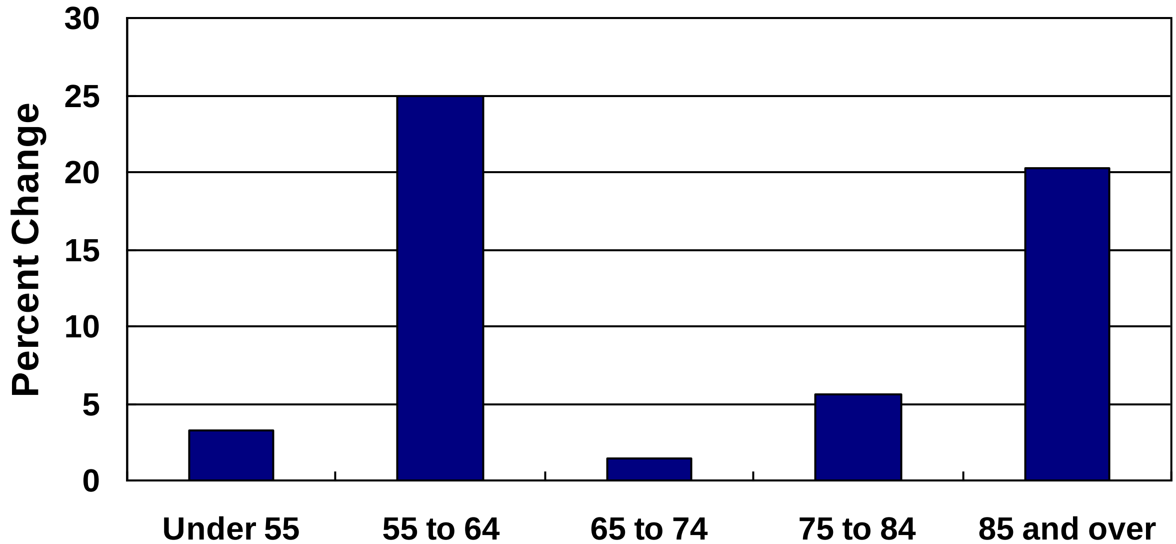


Figure 1.3: Percent Increase in Resident Population by Age from April 1, 2000, and July 1, 2005 (U.S. Census, 2005)

however, includes frequency information up to 8000 Hz. In fact, the high frequency range, 1000 to 8000 Hz, holds 95 percent of speech intelligibility (Gerber, 1974). This means that speech energy above 3000 Hz, which is required to correctly recognize most consonant sounds, such as the fricatives, is lost (Harhager, 2007). On the other hand, environmental sounds that transmit the presence and realism of the world are present in the low-frequency sounds, below 300 Hz (Harhager, 2007). For example, the fundamental frequency of a male voice is 125 Hz while for a female voice, the fundamental frequency is 250 Hz. The lack of this specific information caused distortion in the talker's voice, making them sound unrealistic. Along with a lack of phonemic information above 3000 Hz and below 300 Hz, visual cues are lost as well. This does not allow hearing loss individuals to use lip-reading and facial cues which are helpful in determining specific details such as lip-rounding and placement information of phonemes (Terry et al, 1992). The reduced dynamic range of the telephone is caused by peak clipping. The telephone signal is linear in nature up to the point of maximum amplitude. If a signal exceeds the available amplitude range of the telephone, however, the signal is clipped at that point causing nonlinearity and/or distortion. The signal may also include interference and/or distortion generated during the electrical transmission through the telephone line (Roup et al, In Review). By excluding such vital acoustical information, the telephone enhances the present limitations of speech understanding for older people with hearing loss.

1.4 Background and Motivation of TSEA

The Franklin County Ohio Office on Aging (FCOA) realized that their counselors were having problems communicating with the older adults with hearing loss who called in for assistance. In the late 1990s, FCOA contacted The Ohio State University Department of Speech and Hearing Sciences to help find a solution to this problem. The FCOA has offered the Senior Options Program since 1993 providing Franklin County residents 60 years of age and older with services such as medical transportation, home delivered meals, and minor home repair (FCOA, 2004).

1.4A *Telephone Speech Enhancement Algorithm*

In response to the FCOA's request, a research group from The Ohio State University Department of Speech and Hearing Sciences and the Department of Electrical and Computer Engineering created a speech enhancement algorithm to alleviate the complications with telephone communication for older adults with hearing loss. The speech enhancement system processes speech before being sent over the telephone on an outside line (Harhager, 2007). The algorithm was designed to compensate for the limited bandwidth of the telephone (300-3000 Hz range) and the listener’s hearing loss. Currently, the algorithm uses an average audiogram of 100 older adults within the age of 74 to 93 years of age. Each subject was referred to and tested at the Columbus Speech and Hearing Center between 1996 and 2000 (Harhager, 2007). The objective of the algorithm is to use compression amplification to improve intelligibility within the limited bandwidth of the telephone without amplifying the signal above the listener’s threshold of discomfort (Roup et al, In Review).

The average hearing loss of the 100 older adults used to produce the telephone speech enhancement algorithm (TSEA) is pictured in Figure 1.4A illustrated by the dotted line. The average thresholds were increased at the 2 kHz, 3 kHz, and 4 kHz frequencies due to the roll-off of the frequency response of the telephone that is displayed about 2 kHz. The thresholds, however, were adjusted to compensate for the effect of the lost frequencies over the phone line. Such roll-off effects cause an additional “hearing loss” at those frequencies. The modified audibility curve is shown by the solid line in Figure 1.4A (Harhager, 2007).

In addition to the compensated bandwidth, the algorithm adjusts the amount of gain according to the intensity of the speech signal. For example, using a multi-channel dynamic range compression algorithm, more gain was provided to the less intense consonants rather than the more intense components of speech such as the vowels. While applying the appropriate gain to the speech signal, it keeps the amplitude within the dynamic range of the listener and the telephone (Roup et al, In Review).

While the gain is applied to the specific speech signals, the speech enhancement algorithm compensates for the steep roll-off of the frequency response of the telephone above 2 kHz (Harhager, 2007). The roll-off arises because the telephone introduces nonlinearity by clipping the signal. This is because the speech signal is beyond the available amplitude range of the telephone. The roll-off on the frequencies about 2 kHz can simulate additional hearing loss that can occur in the 2 kHz to 4 kHz frequency range. Consequently, additional gain is provided between this particular frequency range to counteract for the additional loss from the limitations of the telephone system (Harhager, 2007).

Sensorineural hearing loss has been characterized by the decreased dynamic range of hearing and reduced spectral resolution (Tejero-Calado, 2001). It has been

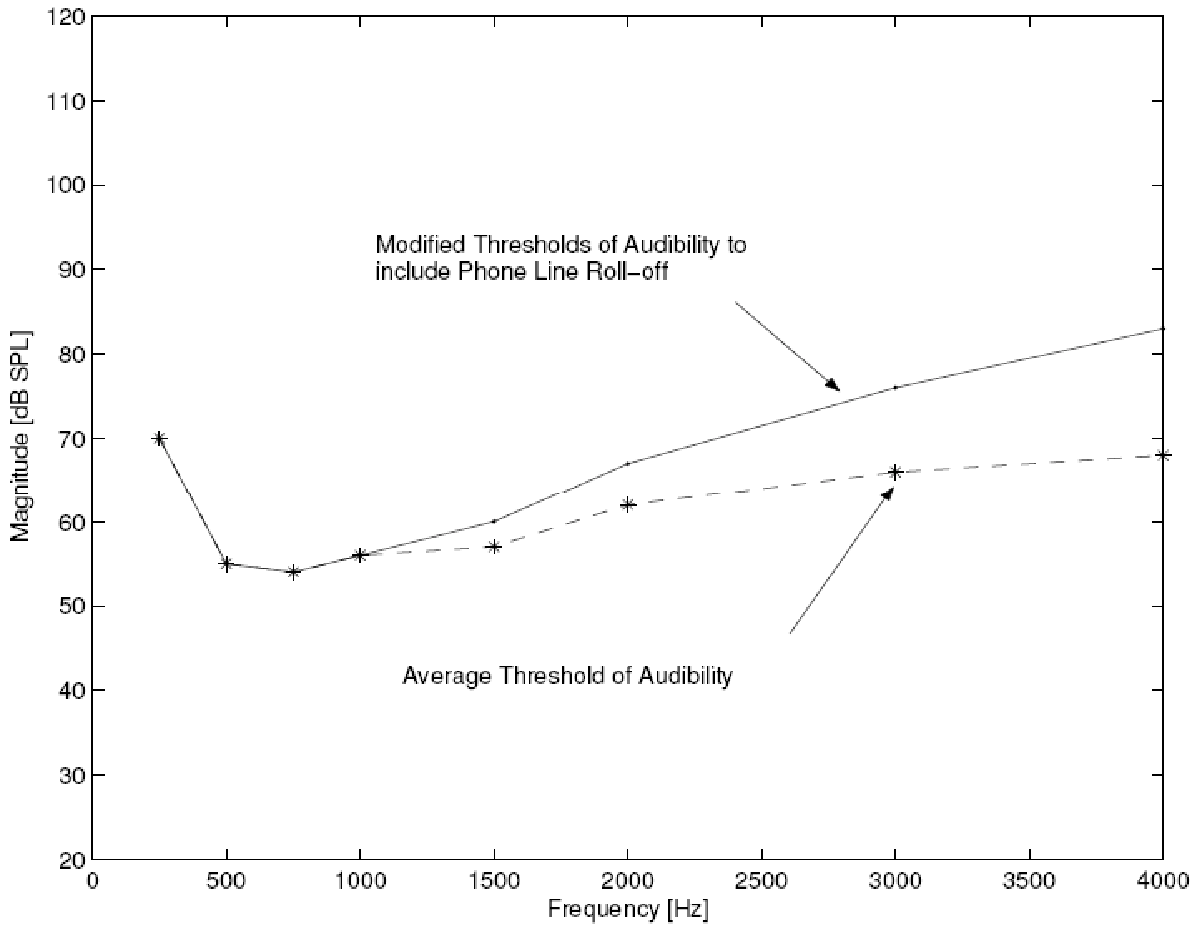


Figure 1.4A: Average Audibility Thresholds and Modified Average Thresholds. The average threshold of audibility curve (dotted line) is the average of audiograms of 100 older adults. The modified thresholds of audibility (solid line) is the average audiogram compensated for the phone line frequency roll-off. Figure from Harhager (2004).

shown that by preserving the spectral contrast during processing, speech intelligibility can be improved for such individuals. It is impossible to preserve spectral peak-to-valley ratios throughout the frequency range in interest in a multi-band compression, however, the given algorithm attempts to preserve the essential formant frequencies in the speech signal (Tejero-Calado, 2001). The algorithm selects the most dominant peaks and applies appropriate gain to those frequencies while still remaining within the individual's dynamic range. First, the peaks are divided into 32 milliseconds frames with 50% overlapping between the following frames (Harhager, 2007). Determined by a Fast Fourier Transform, the spectral information provided in each frame is determined and processed. If the spectral signal is below the noise spectrum, it is classified as "non-speech" and no gain is applied. After the information is processed, the average spectral levels are obtained and applied to the critical bands in each frame. Then, the integrated channels are passed through a peak detection module and found by using the three most important peaks available (Harhager, 2007). The three major peaks relate to the formant frequencies most essential in the given speech signal.

After all the peaks are divided and the information is processed, "the gains are determined and the compression ratio is applied to each channel calculated based on the average threshold of the model audiogram in that channel along with the spectrum level" (Harhager, 2007). Next, the gain is smoothed across the frames to avoid applying gain too quickly across each frame. If the gain is applied to fast, the signal may be uncomfortable for the listener. Lastly, the frames are transformed into a time domain and processed into a speech signal. An "over-lap add technique" is used to combine the frames and complete the process. The processed speech signal is close to real-time, however, processed speech spoken over the transducer has a delay of 16 ms

(Harhager, 2007). The general steps of the overall process are mapped out in Figure 1.4B.

1.4B *TSEA and the Individual Audiograms*

For this particular project, the objective was to apply the current algorithm to each participant's individual audiogram. The idea behind the use of individual audiograms was to examine the differences between the characteristics of the average algorithm (or default audiogram) and the subject's particular hearing loss. Examining each aspect of the two processes allows the individual to take advantage of the maximum performance level the algorithm provides while communicating on the telephone. Depending on the hearing loss, applying the individual's personal audiogram to the original TSEA may not allow for maximum performance on the telephone. If the hearing loss is better (closer to normal) at certain frequencies than those of the average audiogram, the final dB SPL signal processed over the telephone will be lower, allowing for less gain to be applied. Table 1.4 illustrates the average frequencies for the 100 older adults and the amount of gain applied at each frequency to convert dB HL to dB SPL. By using the individual audiograms as the reference hearing loss for the TSEA, it was predicted that the overall improvement in performance on the three speech understanding tests under the three conditions would be greater than the improvement in performance on the default audiogram.

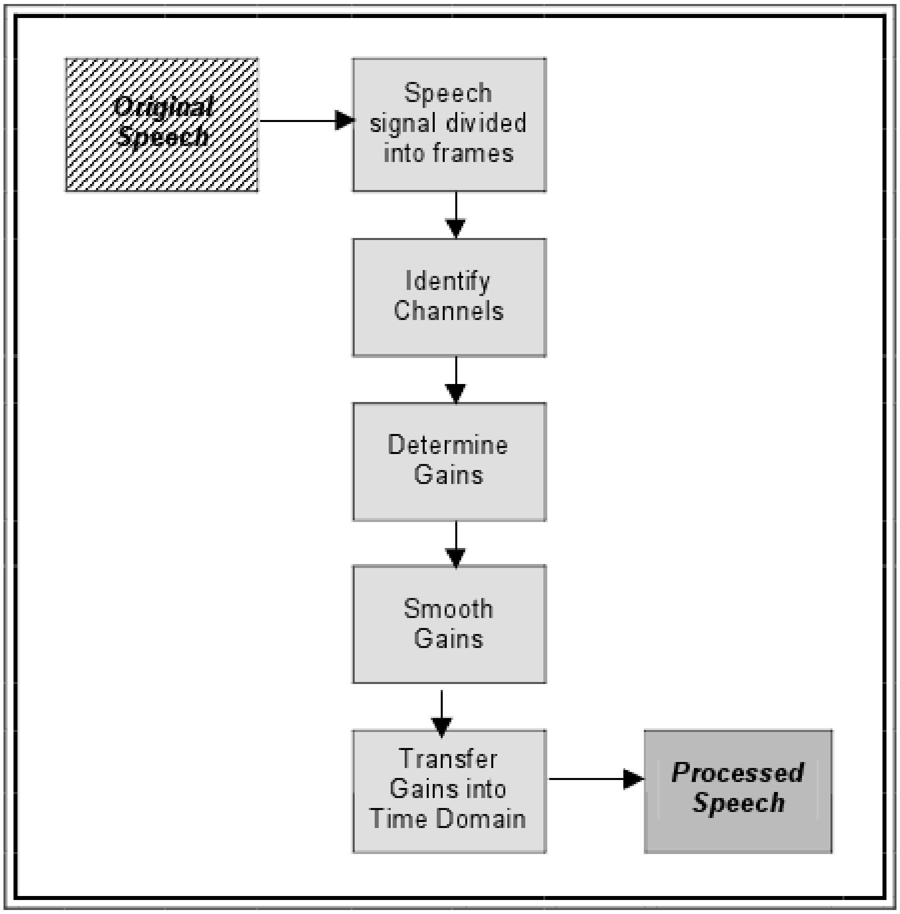


Figure 1.4B: Main Modules of the Speech Processing Algorithm (Harhager, 2007)

	250Hz	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz
dB HL AVERAGE	42	42	49	45	53	49	59	61
Gain Added	**25.5**	**11.5**		**7**	**6.5**	**9**	**10**	**9.5**
dB SPL FINAL	68	53		52	60	58	69	70

Table 1.4: Conversion from dB HL to dB SPL for TSEA

CHAPTER 2

Methods

2.1 Subjects

Ten older adults (4 females, 6 males) with sensorineural hearing loss (SNHL) were recruited for the present study. The subject's ages ranged from 66 to 84 years of age with a mean of 76.9 years of age. All subjects were patients at the OSU Speech, Language, and Hearing Clinic and had participated in previous research studies at OSU. The degree of sensorineural hearing loss varied from mild to severe, including both unilateral and bilateral impairments. The configurations of the hearing losses ranged from flat to precipitously sloping. In order to determine eligibility for the study, each participant was given a complete audiologic evaluation. The evaluation included: 1) air and bone conduction audiometry; 2) tympanometry and acoustic reflexes; and 3) otoscopy. To participate in the study, the subjects had to meet the following criteria: 1) mild to severe SNHL in at least one ear; 2) normal otoscopic findings; 3) tympanometry within normal limits; 4) English as a first language; and 5) 55 years or above in age. Figure 2.1 illustrates the mean pure-tone thresholds of the participants used for this study compared to the default audiogram used in the current version of TSEA. All subjects were recruited from The Ohio State University Speech Recognition and Aging Laboratory research database.

2.2 Stimuli

Three parameters of speech were evaluated in this study: 1) phoneme

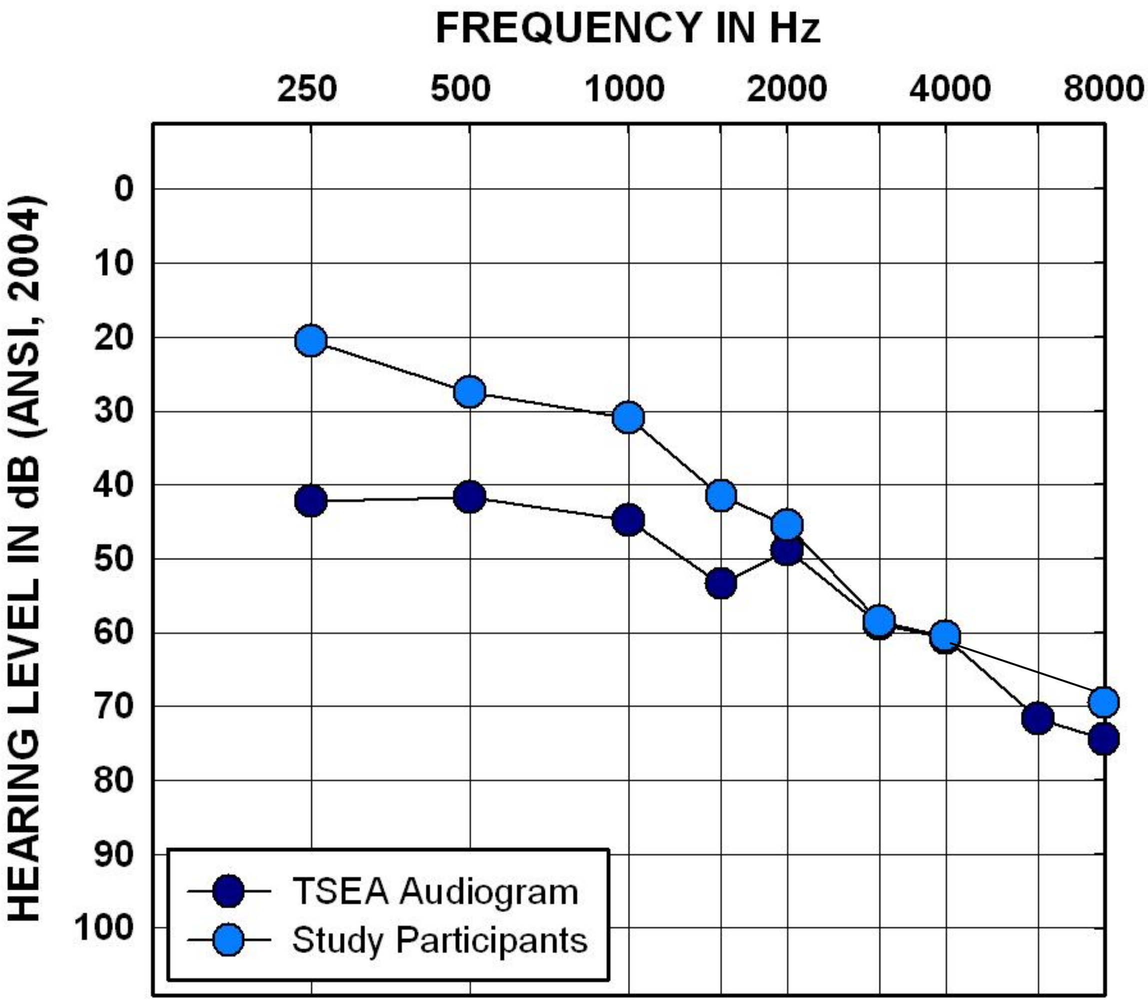


Figure 2.1: Average audiogram for the 10 subjects used in the present study compared to the default audiogram used in the current version of TSEA.

discrimination; 2) word recognition; and 3) sentence recognition. The effectiveness of the TSEA was measured using three tests of speech understanding: the Modified Rhyme Test (MRT), the Speech Perception in Noise test (SPIN), and the Quick Speech in Noise test (QSIN). The MRT examines an individual's ability to discriminate specific phonemes with the aid of a word bank. It is a closed-set test with 50 items, each set consisting of six monosyllabic words that vary in the initial or final consonant. The listener had a word list in front of them for the entire MRT portion of the study, allowing the participant to use visual reinforcement to aid in phonemic discrimination. The listener was told to repeat the word they heard following the carrier phrase "mark the ______ please" (Kreul et al., 1968). The SPIN test requires an individual to repeat the last word they hear in sentences that may or may not have context. The SPIN includes eight lists of 50 sentences, 25 with low predictability and 25 with high predictability. High predictability (HP) sentences allow the participant to predict the final word based on context cues while low predictability (LP) sentences do not provide any context cues to aid in predicting the final word in the sentence. The SPIN test requires the participant to repeat the last word they hear in each of the sentences provided over the telephone (Bilger et al., 1984). Lastly, the QSIN test requires an individual to repeat the entire sentence he/she hears while five target words are scored. The entire test consists of 12 sentences with 60 target words total (Killion et al., 2003). The order of the tests was counter-balanced across participants. Each of the speech intelligibility tests (MRT, SPIN, and QSIN) were digitized and stored on a desktop PC hard drive.

2.3 Procedures

The current study required two separate sessions: 1) Session 1 measured performance at 3 S/N's in order to determine the 50%-correct point on the psychometric function; and 2) Session 2 measured performance under 3 conditions at the S/N determined in Session 1: unprocessed, processed with the default audiogram (Proc-D), and processed with the individual audiograms (Proc-IA). Two sessions were required in order to limit fatigue and minimize the learning effect. Prior to beginning Session 1, each subject was provided with forms to read and sign. These forms include the Consent for Participation in Social and Behavioral Research and the Authorization to Use Personal Health Information in Research. Subjects were provided with a parking pass and compensated for their time in the study.

2.3A *Session 1*

Session 1 measured speech understanding in a multi-talker babble for each test unprocessed, or without the TSEA. The multi-talker babble was presented out of two separate speakers: one directly behind the participant and one directly in front of the participant. In order to maintain consistency, the subject is told to use the same ear on the telephone (whichever they prefer) for the entire study. Each test was given a minimum of three times at various signal-to-noise ratios (SNR) in order to generate a 50%-point on the psychometric function. The 50% point stands for the SNR at which the subject responded correctly 50% of the time for each test. Only half of the test was given at each SNR: 25 out of 50 lists for the MRT, 25 out of 50 sentences for the SPIN, and 6 out of 12 sentences for the QSIN. This is achieved by changing the signal-to-noise ratio, or amount of background noise coming through the speakers. An interpolation process was used to determine the exact 50% point which was later used as a control during Session 2.

2.3B *Session 2*

Session 2 measured performance under 3 conditions at the S/N determined in Session 1: unprocessed, processed with the default audiogram (Proc-D), and processed with the individual audiograms (Proc-IA). Each test was performed one time for each of the three conditions. The entire test was administrated for each test of speech understanding as opposed to Session 1.

CHAPTER 3

Results

The data collected during Session 1 was used primarily for determining the presentation level used during Session 2 (i.e., the 50%-correct threshold). Speech recognition performance was determined during Session 2 in three conditions for each speech test: unprocessed, processed with the default audiogram (Proc-D), and processed with an individual audiogram (Proc-IA). The percent improvement between the unprocessed and processed conditions was also determined.

3.1 Group Results

Descriptive statistics (means and standard deviations) for each speech test and each condition are presented Table 3.1. As can be seen in Table 3.1, the average recognition performance unprocessed was better than the estimated 50% for the MRT (66.4%), SPIN (61.7%), and QSIN (52.5%). Recognition performance for each of the three intelligibility tests for the unprocessed condition and both processed conditions (Proc-D and Proc-IA) is illustrated in Figure 3.1A. As can be seen in Figure 3.1A, mean recognition performance was better for the default audiogram across all test measures. The difference in recognition performance between processing conditions was small for the MRT (76.5% vs. 69.8%). In contrast, the difference in recognition performance between processing conditions was larger for the SPIN (72.4% vs. 60.6%) and the QSIN (75.1% vs. 49.3%).

	MRT (%)	SPIN (%)	QSIN (%)
Unprocessed			
Mean	66.4	61.7	52.5
SD	7.7	12.1	14.7
Processed-Default			
Mean	76.4	72.4	75.1
SD	8.4	8.5	13.0
Processed-Individual			
Mean	69.8	60.6	49.3
SD	9.5	21.8	13.8

Table 3.1: Means and standard deviations for the MRT, SPIN, and QSIN across the three conditions: unprocessed, processed default audiogram, and processed individual audiogram.

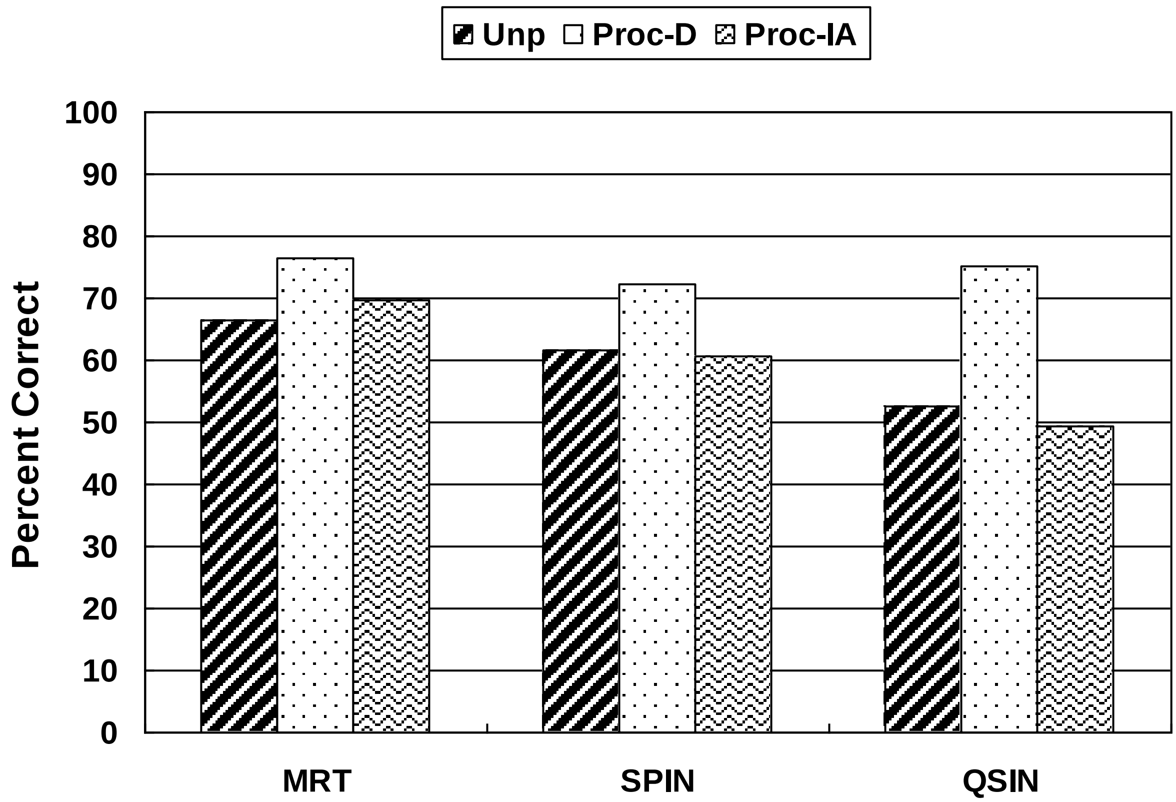


Figure 3.1A: Average recognition performance (in percent) for each speech measure (MRT, SPIN and QSIN) and the three conditions: unprocessed, processed with the default audiogram (Proc-D), and processed with the individual audiograms (Proc-IA).

Overall, the percent correct recognition performance was greater when the TSEA was processed with the default audiogram when compared to the unprocessed condition. In contrast, when the TSEA was processed with the individual audiograms, recognition performance was greater than unprocessed for the MRT only. The application of the TSEA with the default audiogram illustrated greater recognition performance for the SPIN and QSIN than it did with the individual audiograms. The amount of improvement in recognition performance due to processing with the default audiogram for the MRT, SPIN, and QSIN was 10.1%, 10.7%, and 22.6% respectively. The amount of improvement in recognition performance due to processing with the individual audiograms for the MRT, SPIN, and QSIN was 3.4%, -1.1%, and -3.2% respectively. The average percent improvement due to processing with default and individual audiograms is shown in Figure 3.1B.

Prior to statistical analysis, the average scores for the MRT, SPIN, and QSIN were transformed to rationalized arcsine units (rau's) in order to correct for the error variance associated with percentage data (Studebaker, 1985). The transformed data were examined using a series of one-way repeated measures analysis of variance (ANOVA). The ANOVAs revealed a significant main effect of *condition* for the MRT ($F_{2, 27} = 3.6$; $p < .05$) and the QSIN ($F_{2, 27} = 10.7$; $p < .05$). No significant main effect of condition was found for the SPIN ($F_{2, 27} = 1.8$; $p > .05$). Post hoc analysis for the MRT using paired-samples *t*-tests with Bonferonni correction revealed a significant difference in recognition performance between the unprocessed condition and the Proc-D condition ($t_9 = -5.8$; $p < .017$). Specifically, recognition performance on the MRT was signficantly better in the Proc-D condition, reflecting the improvement in performance.

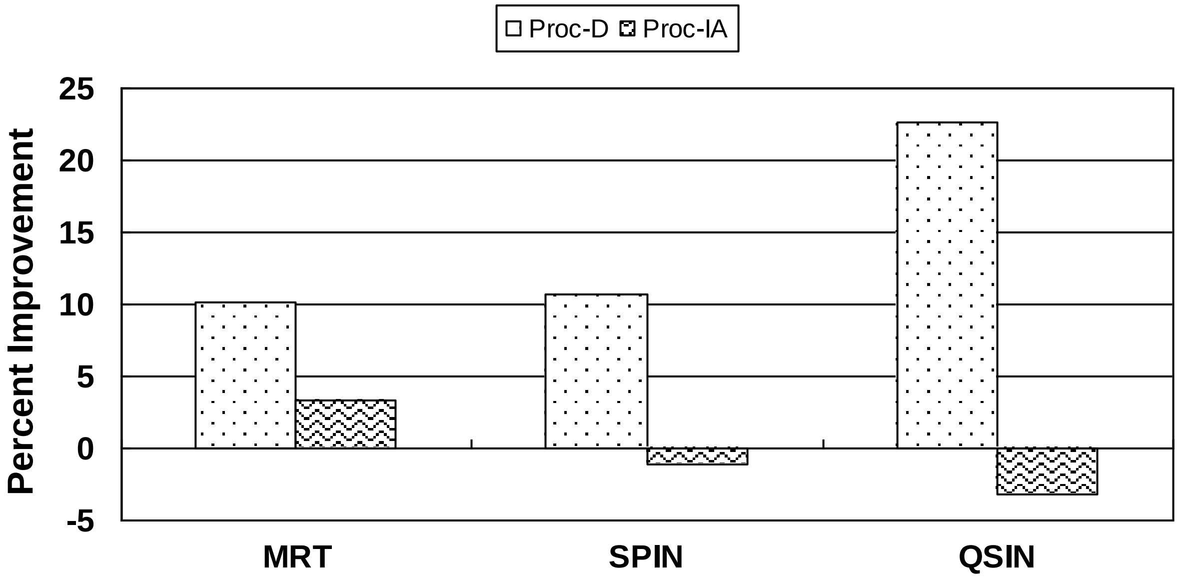


Figure 3.1B: Average percent improvement in recognition performance due to processing with the default audiogram (Proc-D) and with the individual audiograms (Proc-IA) for MRT, SPIN, and QSIN.

due to the TSEA. No significant differences were found for the MRT between the unprocessed and Proc-IA conditions or the Proc-D and Proc-IA conditions. Post hoc analysis was also conducted for the QSIN using paired-samples t-tests with Bonferonni correction. Results revealed signficant differences in recognition performance between the unprocessed and Proc-D conditions ($t_9 = -3.4$; $p < .017$), reflecting the improvement in performance due to the TSEA. A significant difference was also found between the Proc-D and Proc-IA conditions ($t_9 = 5.1$; $p < .017$). In this case, recognition performance on the QSIN was significantly poorer when the TSEA used the individual audiograms than with the default audiogram.

3.2 Individual Subject Results

Overall the data reflected better recognition performance with the default algorithm rather than the algorithm using the individual audiograms. When each subject was analyzed independently, key trends became evident. All subjects had better performance on the QSIN for Proc-D than for Proc-IA. Similarly, most subjects exhibited better performance on the MRT and SPIN for Proc-D than for Proc-IA. Subject 1 and Subject 6 exhibited similar high frequency hearing loss, however, major differences in the low frequency thresholds resulting in differences in performance across the three speech understanding tests for both the Proc-IA and Proc-D.

Subject 6 was an 84 year old male with precipitously sloping normal to profound sensorineural hearing loss in the left ear and precipitously sloping mild to profound sensorineural hearing loss in the right ear. Figure 3.2A illustrates Subject 6's audiogram in comparison to the default audiogram. For the present study, Subject 6 used his left ear throughout the entire study. His percent correct for the individual audiogram algorithm during the MRT was greater than the default algorithm and equal to the default algorithm during the SPIN. For the QSIN, however, the default algorithm percent correct was greater than the individual audiogram algorithm. The MRT, SPIN, and QSIN percent correct for the processing individual algorithm was 78%, 76%, and 63.3% respectively. While for the default audiogram, the MRT, SPIN, and QSIN percent correct was 64%, 76%, and 83.3% respectively. The improvement rate for the individual algorithm when compared to the unprocessed percent correct for the MRT, SPIN, and QSIN was 14%, 18%, and 16.7%, respectively, while for the default algorithm was 0%, 18%, and 36.7%, respectively. Figure 3.2B illustrates the percent correct data and figure 3.2C illustrates the percent improvement for Subject 6.

Subject 1 was a 74 year old male with a normal to mild hearing loss in the low frequencies precipitously sloping to severe sensorineural hearing loss in the high frequencies in the right ear. Figure 3.2D illustrates Subject 1’s audiogram in comparison to the default audiogram. For the present study, Subject 1 used his right ear throughout the entire study. His percent correct for the Proc-IA during the MRT, SPIN, and QSIN was less than the Proc-D. The MRT, SPIN, and QSIN percent correct for the processing individual algorithm was 64%, 32%, and 33.3% respectively. For the default audiogram, the MRT, SPIN, and QSIN percent correct was 66.6%, 60%, and 46.6% respectively. The percent improvement for the individual algorithm when compared to the unprocessed percent correct for the MRT, SPIN, and QSIN was 12%, -4%, and -48.3% respectively, while for the default algorithm was 14.6%, 24%, and -35%, respectively. Figure 3.2E illustrates the percent correct data and figure 3.2F illustrates the improvement rates for Subject 1.

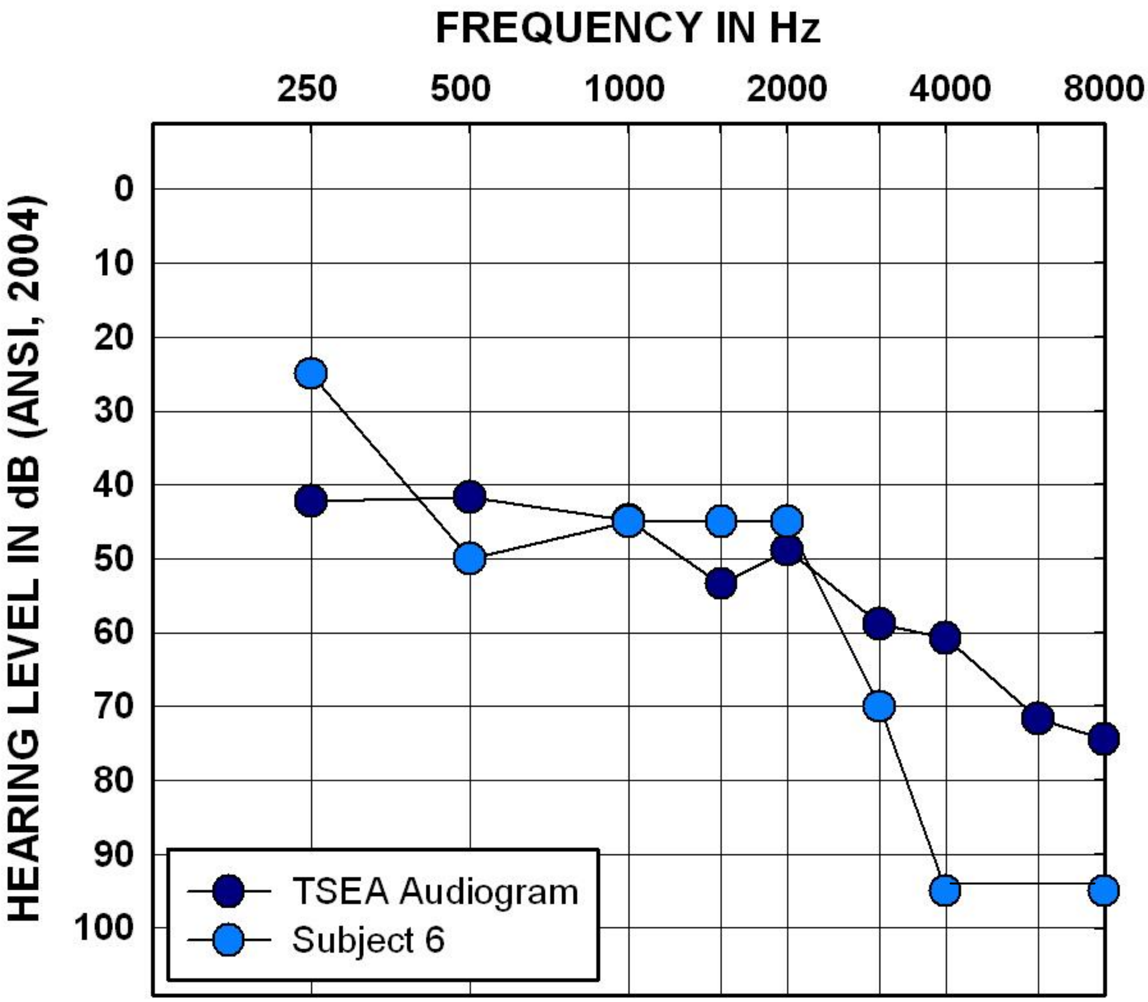


Figure 3.2A: Subject 6 audiogram in comparison to the default audiogram.

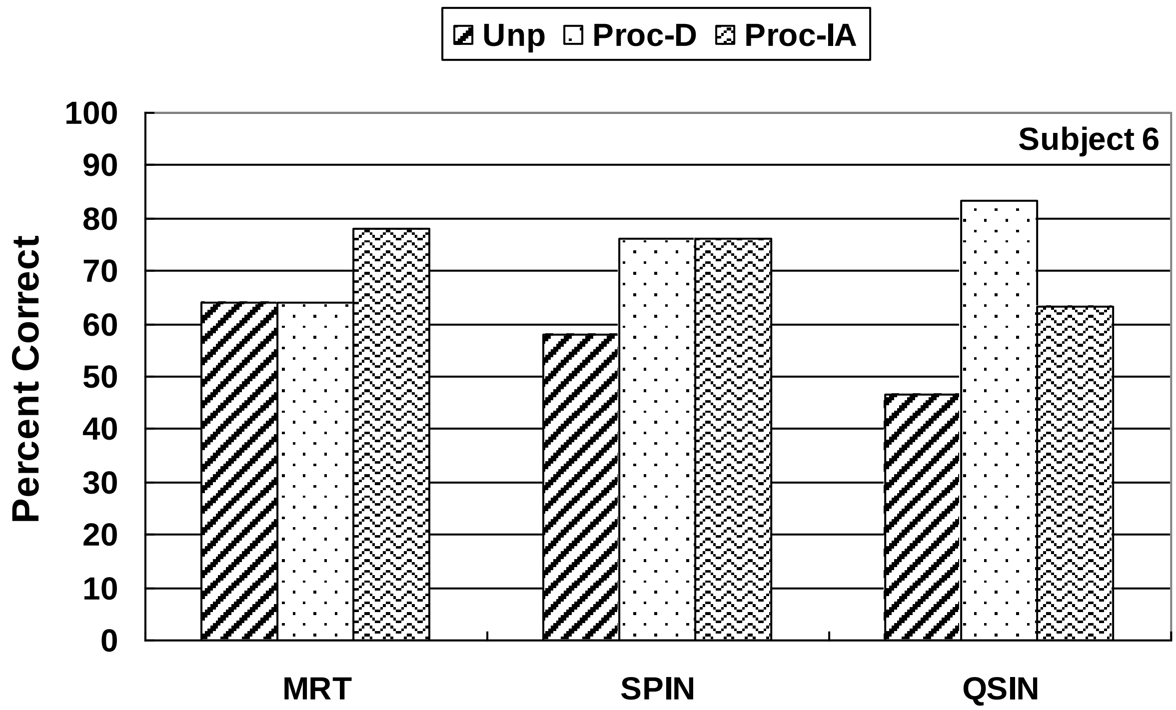


Figure 3.2B: The Percent Correct for the MRT, SPIN, and QSIN for Subject 6 during Session 2

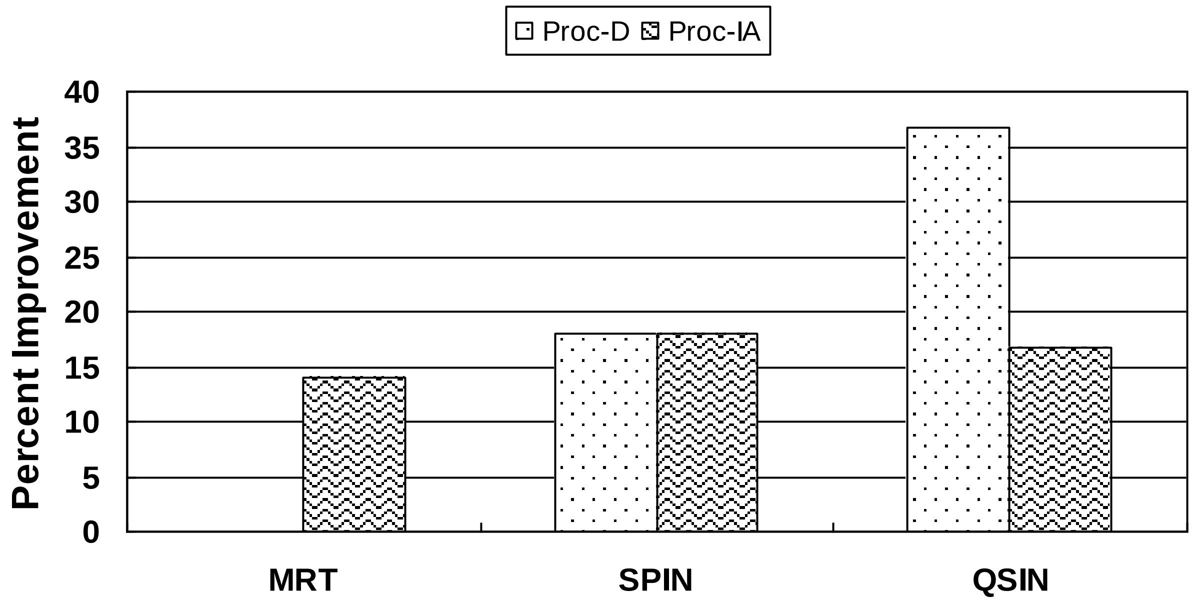


Figure 3.2C: The Improvement Rates for the MRT, SPIN, and QSIN for Subject 6 during Session 2

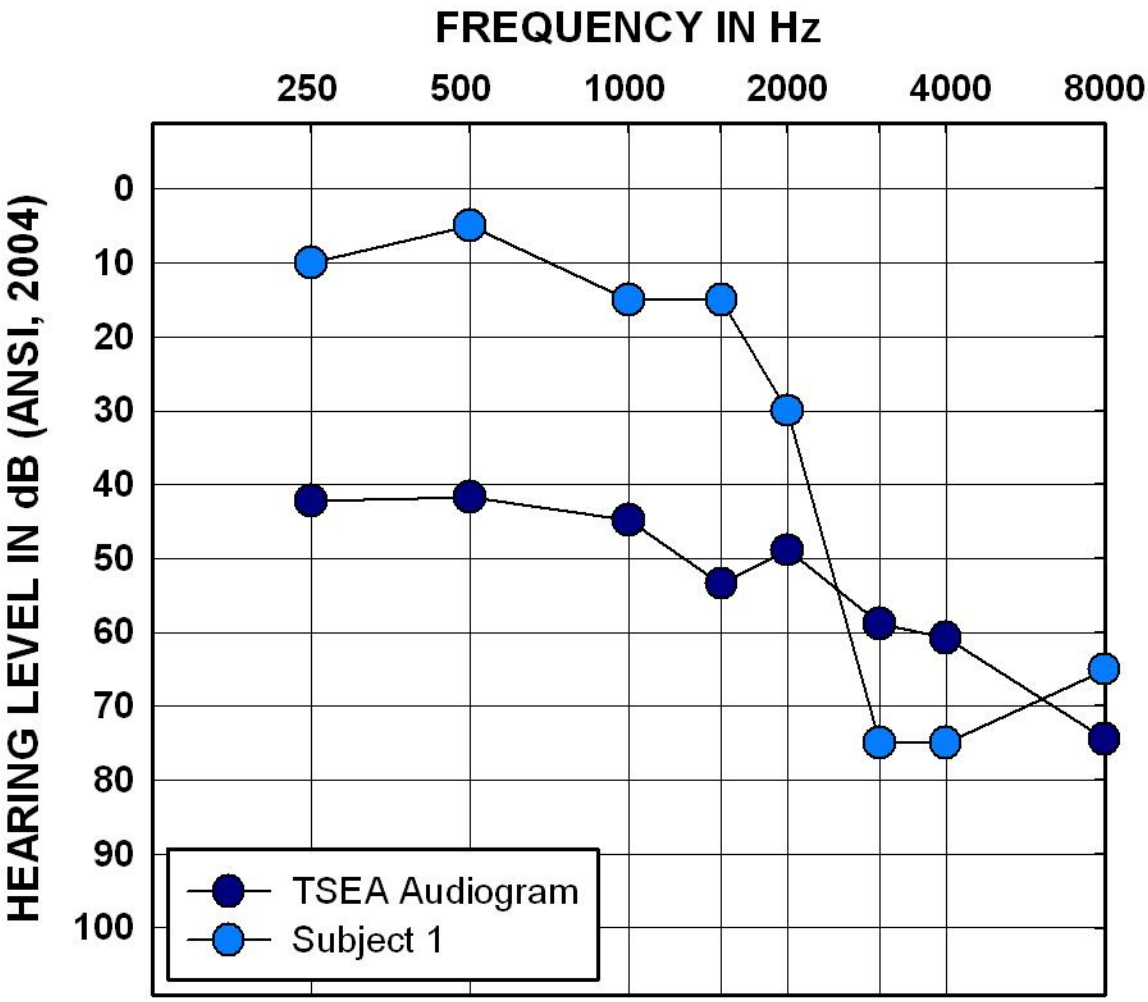


Figure 3.2D: Subject 1 audiogram in comparison to the default audiogram.

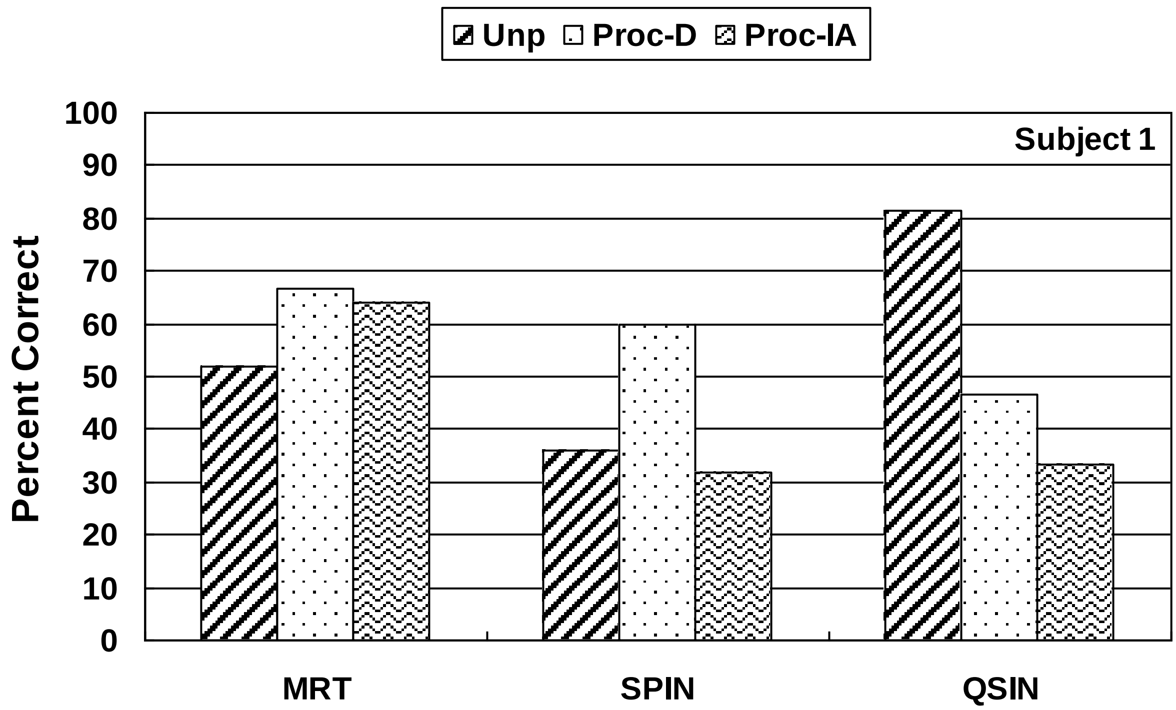


Figure 3.2E: The Percent Correct for the MRT, SPIN, and QSIN for Subject 1 during Session 2

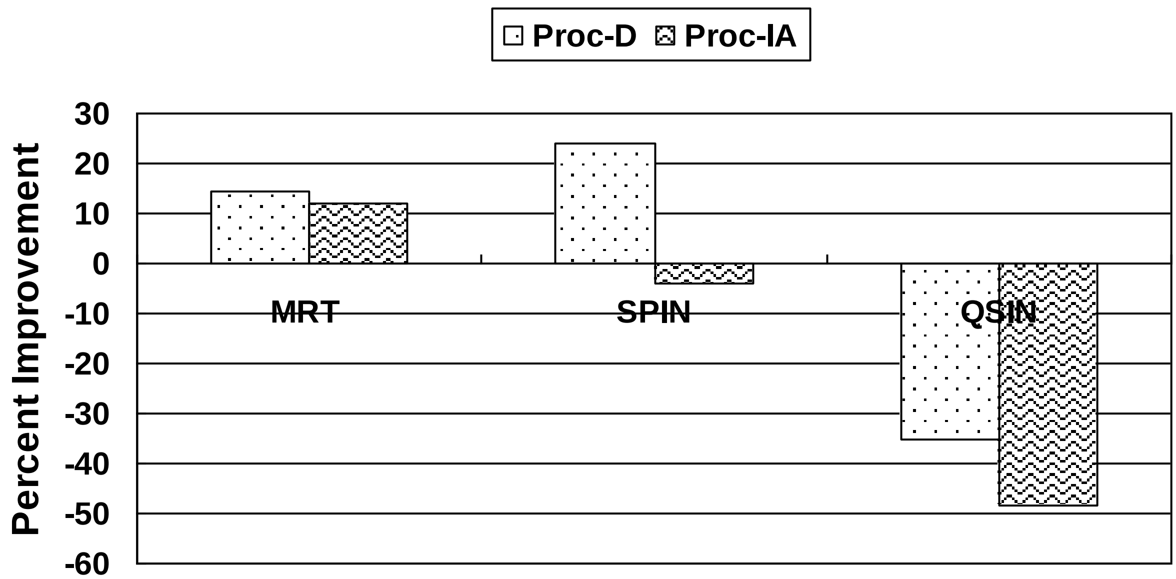


Figure 3.2F: The Improvement Rates for the MRT, SPIN, and QSIN for Subject 1 during Session 2

CHAPTER 4

Discussion and Conclusions

Telephone communication for older adults with hearing loss can be extremely difficult, or in many cases, impossible. Such difficulties can be contributed to the limited bandwidth of the telephone, lack of visual cues, amount of background noise present, as well as the degree, type, and configuration of the individual's hearing loss. The development of the TSEA, however, has been proven to aid in speech understanding over the telephone by using a hearing aid compression strategy that amplifies the soft sounds (such as consonants) without amplifying the higher intensity sounds (such as vowels). The TSEA was created to compensate for the limited audibility experienced by hearing-impaired listeners. The telephone still filters the frequencies to fit within its limited bandwidth (i.e., 300-3000 Hz). Using a hearing aid compression strategy, the TSEA adjusts the amount of gain depending on the intensity level of the speech signals and the reference audiogram, while remaining within the individual's dynamic range. The results from the present study using the individual audiograms as the reference hearing loss for the TSEA suggests that certain characteristics of a presbycusic hearing loss, specifically in the 300-3000 Hz range, may lead to more improvement in performance over the default audiogram used in the present and previous studies.

4.1 Group Results

Overall, the recognition performance was poorer when measured in the unprocessed condition when compared to performance with TSEA with the Proc-D across speech measures. In contrast, the unprocessed condition resulted in equal to better recognition performance than the Proc-IA condition for the QSIN and SPIN. Differences in performance between the two processed conditions (Proc-D and Proc-IA) was likely due to differences in the reference audiogram used by the TSEA. Specifically, the default audiogram employed a mean audiogram with a relatively flat audiometric configuration resulting in amplification across all frequencies, including the lower frequencies. The average group audiogram had a normal sloping to moderately-severe sensorineural hearing loss, characteristic of presbycusis where thresholds at the higher frequencies are worse than thresholds at the lower frequencies (refer to Figure 1.2). Because the TSEA provides more gain for higher frequencies sounds, which contain important speech spectrum's information, the signal via the telephone for the average group audiogram was similar to the default. The average group audiogram, however, has less hearing loss in the lower frequencies, which contain the speech spectrum's intensity, causing the loudness of the signal to be lower. Therefore, the default audiogram is has a greater perception in loudness (refer to Figure 2.1).

Though the TSEA amplified the entire speech spectrum, the telephone still filtered the spectrum in order to fit within its limited bandwidth of 300-3000 Hz. Because the average hearing loss in the 300-3000 Hz frequency range of the present study is less than that in the default audiogram, the default received more low frequency amplification than the study's average hearing loss did. This could explain why the Proc-D improvement in performance was greater than that of the average results of the Proc-IA (refer to Figure 3.1B).

4.2 Individual Results

The same results develop when comparing two particular subjects hearing loss and their performance. Subject 6 also exhibited the common presbycusic hearing loss that slopes from the low to high frequencies. His hearing loss was relatively normal to mild in the low frequencies unlike the default audiogram where there is a moderate loss in the lower frequencies (refer to Figure 3.2A). Because the telephone only allows for a limited frequency range for the outgoing signal, the incoming signal may become distorted when unprocessed. Subject 6's thresholds, specifically in the 300-3000 Hz frequency range, are similar to the default audiogram allowing the TSEA to process each the Proc-IA and Proc-D with relatively similar amplification of the low frequencies and clarity of the high frequencies. Because the 300-3000 Hz thresholds of Subject 6 are similar or worse at frequencies than that of the default's thresholds, it could explain the greater improvement in performance for the Proc-IA for the MRT and SPIN. The TSEA's attention to the low and high frequency thresholds of Subject 6 improved the speech spectrum's loudness and quality perception for both the Proc-IA and the Proc-D.

Subject 1 had a similar high frequency hearing loss, however, this 74 year old male had normal to mild hearing loss up to 2000 Hz. Though his hearing loss was severe in the high frequencies allowing for a clearer speech spectrum when the TSEA was applied, his low frequencies did not receive the same amplification from the TSEA as Subject 6 or the average audiogram did. This resulted in less volume of the signal, meaning the signal sounded as loud as it would if it was unprocessed. Therefore, it was expected that his performance on the Proc-IA would be less than the Proc-D due to the lack of low frequency amplification when the TSEA was applied to his hearing loss. The

TSEA only improved the quality of the high frequency spectrum than is caused by the telephone's roll-off.

Though the TSEA appeared to be less successful overall with the Proc-IA than it did with the Proc-D, as seen with Subject 6, there may be individuals with the types of presbycusis that may be better used as the preferred reference hearing loss for the TSEA. The results suggest that individuals with more hearing loss across the entire frequency range (specifically the 300-3000 Hz range) would have more improvement in performance with the Proc-IA than the Proc-D due to the TSEA's amplification across the speech spectrum. Furthermore, the data collected from the present study illustrated that as long as the individual's hearing in the 300-3000 Hz frequency range is worse than or similar to the default's 300-3000 Hz range, improvement in performance with the Proc-IA is expected. This suggests that the Proc-IA TSEA is more successful when the individual hearing loss is equal or worse than the default audiogram at all frequency thresholds.

4.3 Conclusions and Future Studies

Many older adults, especially those that live independently, are dependent on telephone communication for safety and relationship maintenance. For this reason, it is essential that communication over the telephone becomes easier for the older adults that rely on it. The development of the TSEA suggests an opportunity for improvement for telephone communication. Studies involving the TSEA have been promising. Not only has the algorithm been shown to aid in speech understanding for older adults over the telephone using an average audiogram, it has also shown that improvement in performance is possible if the individual's hearing loss is significant enough across most, if not all, frequencies. The present study illustrated that the use of individualized losses may aid those with more severe cases of presbycusis. As more data is collected about the effects of the TSEA on specific types of presbycusis, it can be expected that the variability of performance will decrease and the reliability of the study will increase. Future research can also aid in determining the population of individuals with such types of presbycusis that can benefit more with the individualized processing of the TSEA.

Other factors that were not measured but potentially affected the results included: 1) the participant's motivation and interest in the test; 2) word recognition ability; and 3) overall communication experience and familiarity with the telephone. Also, the data from the present study represented only 10 listeners varying across a large age range (i.e., 18 years) and a large range of hearing losses.

References

ASHA, (2009). Type, Degree, and Configuration of Hearing Loss. *American Speech and Hearing Association.* Available at: http://www.asha.org/public/hearing/disorders/types.htm

Bess, F & Humes, L. (2003). *Audiology: The Fundamentals*. Baltimore: Lippincott Williams & Wilkins.

Chmiel, R., Jerger, J. (1993) Some factors affecting assessment of hearing handicap in the elderly. *Journal of the American Academy of Audiology,* 4, 249-257.

Connelly, P.E. (2003). Presbycusis – A Look into the Aging Inner Ear. *International Hearing Society.* Last retrieved November-December 2003. Available at: http://www.ihsinfo.org/IhsV2/Hearing_Professional/2003/060_November-December/080_Presbycusis_A_Look_into_the_Aging_Inner_Ear.cfm

Franklin County Office on Aging (FCOA). (n.d.) Keeping Seniors Safe at Home. Available at: http://www.officeonaging.org/common/news/newsletters/2004Jan.pdf

Garstecki, D. (1996). Older Adults: Hearing Handicap and Hearing Aid Management. *American Journal of Audiology*, 5 (3), 25-34.

Gates, G. & Mills, J. (2005). Presbycusis. *The Lancet*, 366, 1111-1120.

Gerber, S. (1974). *Introductory Hearing Science.* Philadelphia: W.B. Saunders.

Harhager, K. (2007). Evaluation of speech processing for telephone use by older adult listeners with hearing loss. *Unpublished Capstone Project, The Ohio State University*.

Heine, C., & Browning, C.J. (2001). Communication and psychosocial consequences of sensory loss in older adults: Overview and rehabilitation directions. *Disability and Rehabilitation*, 24, 763-773.

Jerger, J., Chmiel, R., Wilson, N., & Luchi, R. (1995). Hearing impairment in older adults: New concepts. *American Geriatrics Society*, 43, 928-935.

Lesner, S. (2003). Candidacy and management of assistive listening devices: Special needs of the elderly. *International Journal of Audiology*, 42, 2S68-2S76.

Mack, A. & Mormer, E. (2003). Telephone across the life span: An issue for all. *Seminars in Hearing*, 24, 1-16.

Ross, M. (2004). The Audiogram: Explanation and Significance. *Hearing Loss,* 25 (3), 29-33.

Roup, C.M, Poling, G.L., Harhager, K. et al. (In Review). Evaluation of a Telephone Speech Enhancement Algorithm Among Older Adults with Hearing Loss. *Submitted to Ear and Hearing.*

Schuknecht, H.F. (1964). Further observations on the pathology of presbycusis. *Archives of Otolaryngology*, 80, 369-382.

Smith, I.M. (n.d.). *Hearing Loss and Aging.* Available at: http://www.teresonic.com/pdf/Hearing%20loss%20and%20aging.pdf

Studebaker, G.A. (1985). A rationalized arcsine transform. *J Speech Hear Res,* 28, 455-462

Tejero-Calado, J. C., Rutledge, J.C., & Nelson, P.B. (2001). Preserving Spectral Contrast in Amplitude Compression for Hearing Aids. *Proceedings of the 23rd Annual EMBS International Conference, Istanbul, Turkey.*

Terry,M. et al. (1992). Processing the Telephone Speech Signal for the Hearing Impaired. *Ear and Hearing*, 13, 70-79.

U.S. Census Bureau. (2000). Historical census of housing graphs-telephones. Last revised 02 December 2004. Available at: http://www.census.gov/hhes/www/housing/census/historic/phone.html.

U.S. Census Bureau, (2005). Older Adults in 2005. *Population Profile of the United States.* Available at: http://www.census.gov/population/www/pop-profile/files/dynamic/OLDER.pdf

U.S. Census Bureau, (2006). Selected Highlights from 65+ in the United States: 2005. Last revised 06 June 2006. Available at: http://www.census.gov/Press-Release/www/2006/65appendix.html

Weinstein, B. (2000). *Geriatric Audiology*. New York: Thieme.

Yost, W. (2006). *Fundamentals of Hearing: An Introduction.* Chicago: Academic Press.

Appendix A

MRT

	UNPROCESSED		
Subjects	SN	raw score	% correct
1	-16	26	52.0%
2	-6.8	34	68.0%
3	-7.6	37	74.0%
4	-7.5	34	68.0%
5	0.4	39	78.0%
6	-6.5	32	64.0%
7	-17.6	35	70.0%
8	-11.75	28	56.0%
9	5	34	68.0%
10	6.5	33	66.0%

	Proc-D		
Subjects	SN	raw score	% correct
1	-16	33	66.6%
2	-6.8	42	84.0%
3	-7.6	41	82.0%
4	-7.5	41	82.0%
5	0.4	45	90.0%
6	-6.5	32	64.0%
7	-17.6	40	80.0%
8	-11.75	36	72.0%
9	5	37	74.0%
10	6.5	35	70.0%

	Proc-IA		
Subjects	SN	raw score	% correct
1	-16	32	64.0%
2	-6.8	30	60.0%
3	-7.6	37	74.0%
4	-7.5	34	68.0%
5	0.4	39	78.0%
6	-6.5	39	78.0%
7	-17.6	27	54.0%
8	-11.75	33	66.0%
9	5	35	70.0%
10	6.5	43	86.0%

SPIN

	UNPROCESSED		
Subjects	SN	raw score	% correct
1	-7.37	18	36.0%
2	2.67	36	72.0%
3	12.8	33	66.6%
4	8.6	36	72.0%
5	6.3	34	68.0%
6	4.3	29	58.0%
7	-8.4	23	46.0%
8	3.33	30	60.0%
9	10.5	34	68.0%
10	8.75	35	70.0%

	Proc-D		
Subjects	SN	raw score	% correct
1	-7.37	30	60.0%
2	2.67	29	58.0%
3	12.8	36	72.0%
4	8.6	40	80.0%
5	6.3	41	82.0%
6	4.3	38	76.0%
7	-8.4	33	66.0%
8	3.33	40	80.0%
9	10.5	39	78.0%
10	8.75	36	72.0%

	Proc-IA		
Subjects	SN	raw score	% correct
1	-7.37	16	32.0%
2	2.67	17	34.0%
3	12.8	31	62.0%
4	8.6	42	84.0%
5	6.3	44	88.0%
6	4.3	38	76.0%
7	-8.4	16	32.0%
8	3.33	29	58.0%
9	10.5	30	60.0%
10	8.75	40	80.0%

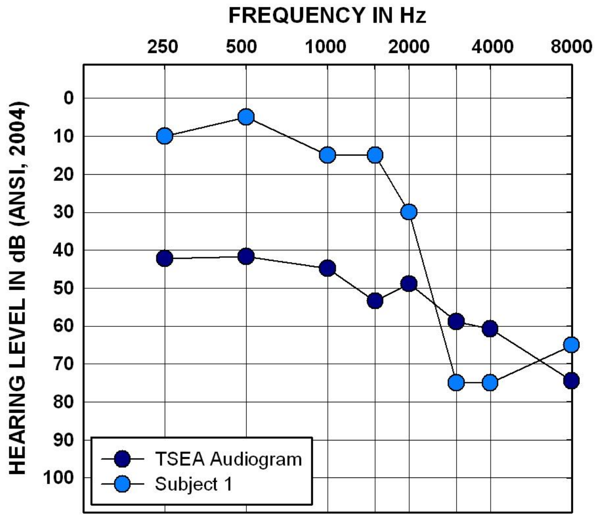
QSIN

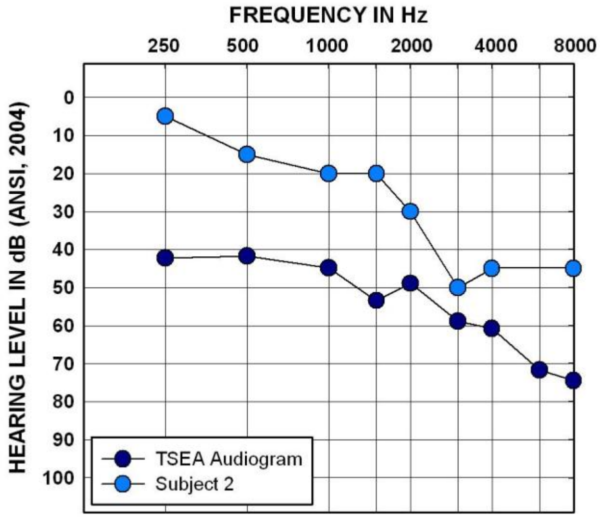
	UNPROCESSED		
Subjects	SN	raw score	% correct
1	-1.4	49	81.6%
2	-4.2	35	58.3%
3	21.6	25	41.6%
4	6.5	32	53.3%
5	1.94	39	65.0%
6	9.4	28	46.6%
7	-4.39	36	60.0%
8	-3.4	17	28.3%
9	16	26	43.3%
10	21.5	28	46.6%

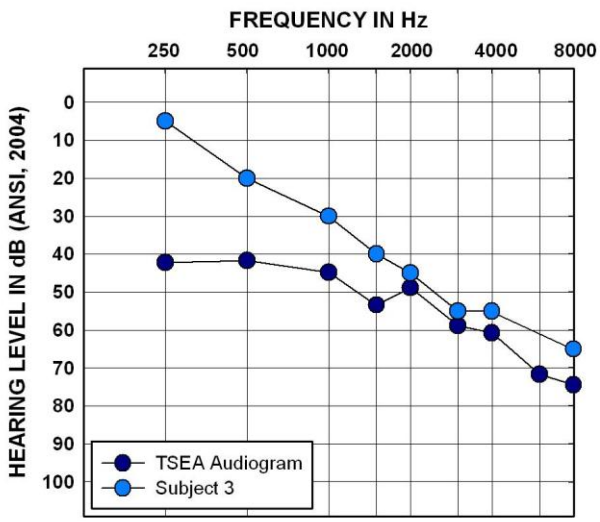
	Proc-D		
Subjects	SN	raw score	% correct
1	-1.4	28	46.6%
2	-4.2	54	90.0%
3	21.6	43	71.6%
4	6.5	50	83.3%
5	1.94	53	88.3%
6	9.4	50	83.3%
7	-4.39	46	76.0%
8	-3.4	40	66.6%
9	16	40	66.6%
10	21.5	47	78.3%

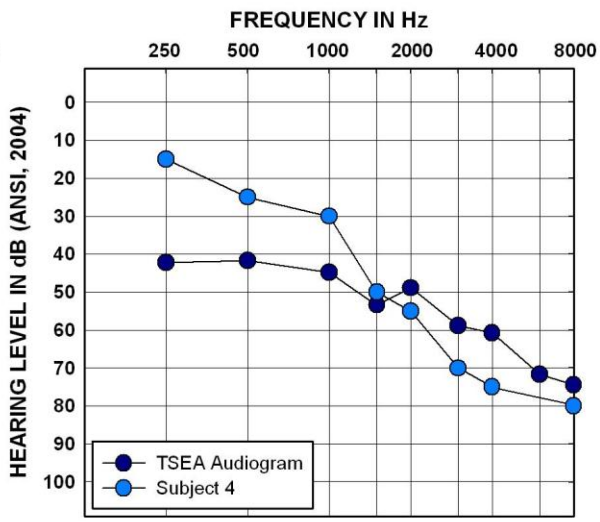
	Proc-IA		
Subjects	SN	raw score	% correct
1	-1.4	20	33.3%
2	-4.2	16	26.6%
3	21.6	38	63.3%
4	6.5	34	56.6%
5	1.94	36	60.0%
6	9.4	38	63.3%
7	-4.39	23	38.0%
8	-3.4	26	43.3%
9	16	27	45.0%
10	21.5	38	63.3%

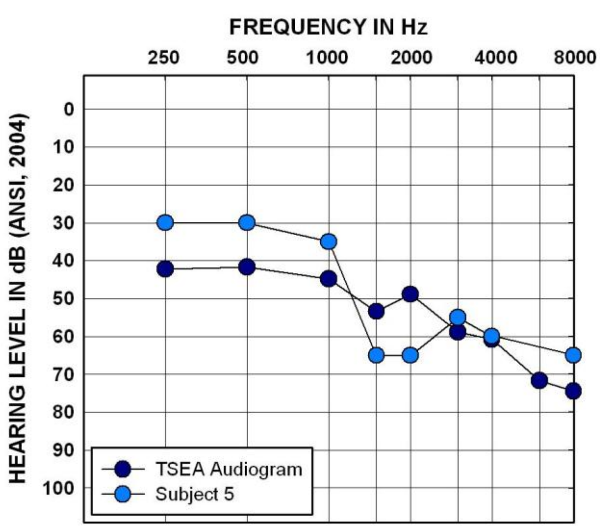
Appendix B

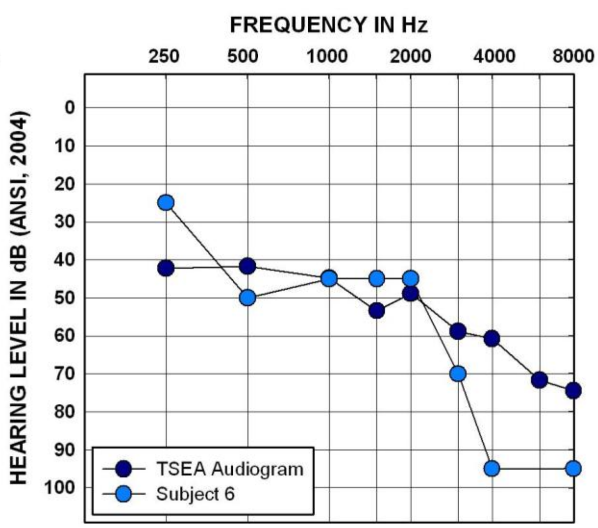

FREQUENCY IN Hz
250 500 1000 2000 4000 8000
HEARING LEVEL IN dB (ANSI, 2004)
0 10 20 30 40 50 60 70 80 90 100
TSEA Audiogram
Subject 1

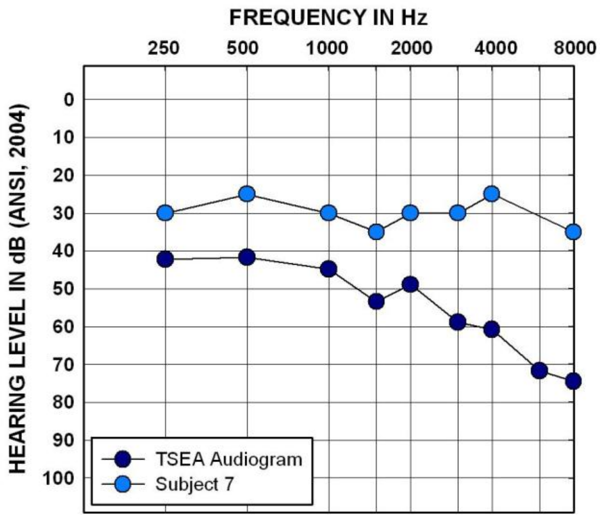

FREQUENCY IN Hz
250 500 1000 2000 4000 8000
HEARING LEVEL IN dB (ANSI, 2004)
0 10 20 30 40 50 60 70 80 90 100
TSEA Audiogram
Subject 2

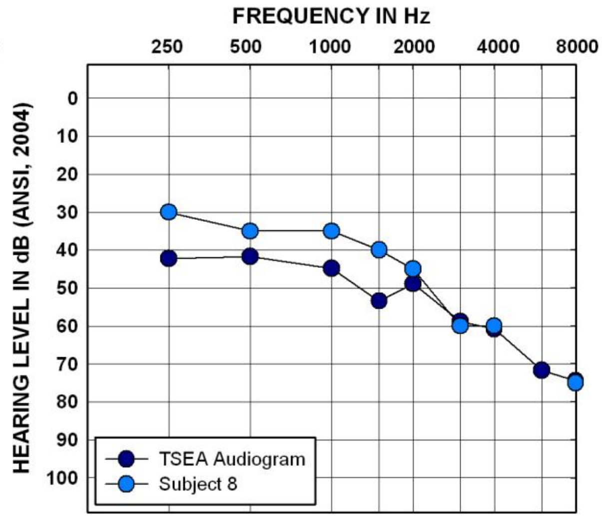

FREQUENCY IN Hz
250 500 1000 2000 4000 8000
HEARING LEVEL IN dB (ANSI, 2004)
0 10 20 30 40 50 60 70 80 90 100
TSEA Audiogram
Subject 3

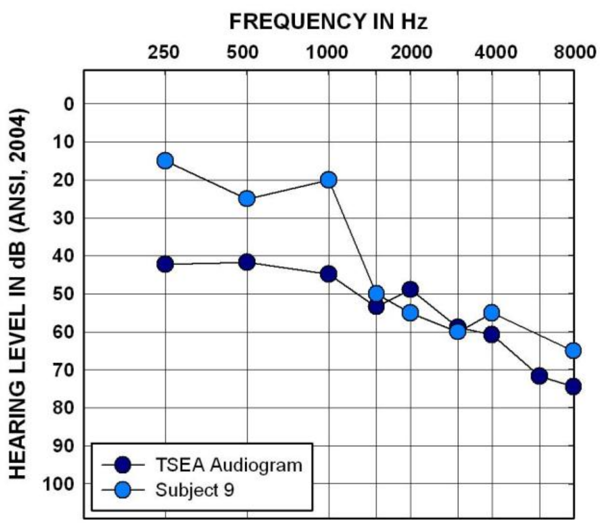

FREQUENCY IN Hz
250 500 1000 2000 4000 8000
HEARING LEVEL IN dB (ANSI, 2004)
0 10 20 30 40 50 60 70 80 90 100
TSEA Audiogram
Subject 4

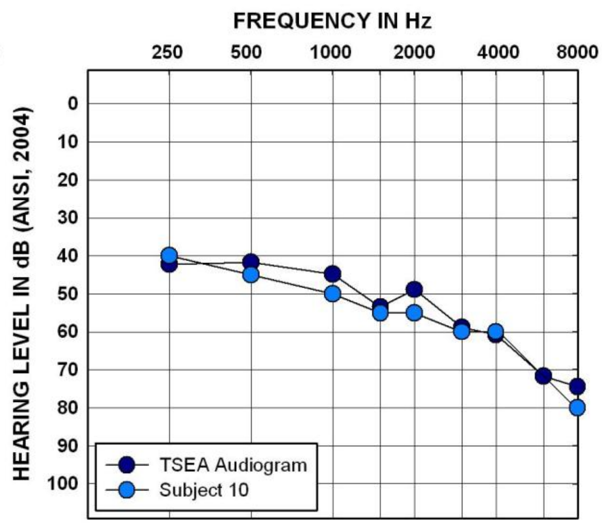

FREQUENCY IN Hz
250 500 1000 2000 4000 8000
HEARING LEVEL IN dB (ANSI, 2004)
0 10 20 30 40 50 60 70 80 90 100
TSEA Audiogram
Subject 5


FREQUENCY IN Hz
250 500 1000 2000 4000 8000
HEARING LEVEL IN dB (ANSI, 2004)
0 10 20 30 40 50 60 70 80 90 100
TSEA Audiogram
Subject 6


FREQUENCY IN Hz
250
500
1000
2000
4000
8000
HEARING LEVEL IN dB (ANSI, 2004)
0
10
20
30
40
50
60
70
80
90
100
TSEA Audiogram
Subject 7


FREQUENCY IN Hz
250
500
1000
2000
4000
8000
HEARING LEVEL IN dB (ANSI, 2004)
0
10
20
30
40
50
60
70
80
90
100
TSEA Audiogram
Subject 8


FREQUENCY IN Hz
250
500
1000
2000
4000
8000
HEARING LEVEL IN dB (ANSI, 2004)
0
10
20
30
40
50
60
70
80
90
100
TSEA Audiogram
Subject 9


FREQUENCY IN Hz
250
500
1000
2000
4000
8000
HEARING LEVEL IN dB (ANSI, 2004)
0
10
20
30
40
50
60
70
80
90
100
TSEA Audiogram
Subject 10