2007

Psychometric functions of clear and conversational speech for young normal hearing listeners in noise

Jane Smart
University of South Florida

Psychometric Functions of Clear and Conversational Speech for Young Normal Hearing Listeners in Noise

by

Jane Smart

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
Department of Communication Sciences and Disorders
College of Arts and Sciences
University of South Florida

Major Professor: Jean C. Krause, Ph.D.
Gail Donaldson, Ph.D.
Stefan Frisch, Ph.D.

Date of Approval:
July 10, 2007

Keywords: speech intelligibility, speaking rates, talker variability, benefit size, intelligibility benefit, study comparisons

Acknowledgments

I would like to thank my thesis committee members, Dr. Jean Krause, Dr. Gail Donaldson and Dr. Stefan Frisch, for their guidance and support. In particular, I would like to acknowledge my thesis advisor, Dr. Jean Krause, for her many hours of assistance, endless patience and constant encouragement.

I would also like to thank all of the listeners for taking the time to participate in this study. Participant screenings and sessions could not have been completed without the help of Ann Siapno, and I thank her for all of her cheerful assistance. Additionally, I would like to thank Ann and John Lum for their efforts with scoring listener responses.

Finally, I would like to thank my family for their love and support in all my endeavors.

Table of Contents

List of Tables

List of Figures

Psychometric Functions of Clear and Conversational Speech for Young Normal Hearing Listeners in Noise

Jane Smart

ABSTRACT

Clear speech is a form of communication that talkers naturally use when speaking in difficult listening conditions or with a person who has a hearing loss. Clear speech, on average, provides listeners with hearing impairments an intelligibility benefit of 17 percentage points (Picheny, Durlach, & Braida, 1985) over conversational speech. In addition, it provides increased intelligibility in various listening conditions (Krause & Braida, 2003, among others), with different stimuli (Bradlow & Bent, 2002; Gagne, Rochette, & Charest, 2002; Helfer, 1997, among others) and across listener populations (Bradlow, Kraus, & Hayes, 2003, among others). Recently, researchers have attempted to compare their findings with clear and conversational speech, at slow and normal rates, with results from other investigators' studies in an effort to determine the relative benefits of clear speech across populations and environments. However, relative intelligibility benefits are difficult to determine unless baseline performance levels can be equated, suggesting that listener psychometric functions with clear speech are needed. The purpose of this study was to determine how speech intelligibility, as measured by percentage key words correct in nonsense sentences by young adults, varies with changes in speaking condition, talker and signal-to-noise ratio (SNR).

Forty young, normal hearing adults were presented with grammatically correct nonsense sentences at five SNRs. Each listener heard a total of 800 sentences in four speaking conditions: clear and conversational styles, at slow and normal rates (i.e., clear/slow, clear/normal, conversational/slow, and conversational/normal). Overall results indicate clear/slow and conversational/slow were the most intelligible conditions, followed by clear/normal and then conversational/normal conditions. Moreover, the average intelligibility benefit for clear/slow, clear/normal and conversational/slow conditions (relative to conversational/normal) was maintained across an SNR range of -4 to 0 dB in the middle, or linear, portion of the psychometric function. However, when results are examined by talker, differences are observed in the benefit provided by each condition and in how the benefit varies across noise levels. In order to counteract talker variability, research with a larger number of talkers is recommended for future studies.

Chapter 1

Introduction

Clear speech is a verbal form of communication that differs from typical "conversational" style speech. It is used naturally by speakers when they are trying to communicate in a noisy environment or with a person who is hearing impaired. One of the first research studies conducted with a "clear" speaking style revealed a 17 percentage point intelligibility benefit of clear speech over conversational speech for hearing impaired listeners (Picheny, Durlach, & Braida, 1985). This landmark study was the first of many research studies to show the substantial gains in intelligibility from clear speech (Bradlow & Bent, 2002; Helfer, 1997; Krause & Braida, 2002; Payton, Uchanski, & Braida, 1994; Uchanski, Choi, Braida, Reed, & Durlach, 1996; among others). Picheny et al.'s findings led other researchers to further explore clear speech with listeners from different populations and in various listening environments. Researchers studied clear speech in quiet (e.g., Payton et al., 1994) and noise (e.g., Bradlow, Kraus, & Hayes, 2003), with normal (e.g., Krause & Braida, 2002), hearing impaired (e.g., Payton et al., 1994), native and non-native listeners (e.g., Bradlow & Bent, 2002), and children (Bradlow et al., 2003). All of these studies found an intelligibility benefit for clear versus conversational speech, demonstrating that the clear speech effect is sizeable and robust.

In addition to the perceptual benefits, clear speech is very easy for talkers to produce. Instructions are minimal and talkers generally become proficient with only 10 to 15 minutes of practice (Schum, 1997). Talkers have been instructed to speak as if they were communicating in a noisy environment, with a listener who has a hearing loss (Helfer, 1997, 1998; Picheny et al., 1985; Schum, 1997), or with a listener whose native language differs from the speaker's language (Bradlow & Bent, 2002; Smiljanic & Bradlow, 2005). Additional instructions may include specific directions for the speaker to articulate each word in a clear and precise manner (Schum, 1996). Beyond these instructions, neither the talker nor the listener requires any special training or abilities. Thus, clear speech provides an economical, practical and beneficial means of communication.

As a result of the many recognized benefits of clear speech, there are several useful applications. Early research focused on determining the clear speech effect and the acoustic differences that contribute to speech intelligibility in order to enhance signal-processing algorithms for hearing aids (Picheny, Durlach, & Braida, 1985, 1986). By learning more about the properties of clear speech, the goal was to develop signal-processing schemes aimed at converting conversational speech to clear speech, improving speech perception and resulting in better overall satisfaction by hearing aid users. Clear speech is also a valuable tool for use in clinical settings. The clinician who uses clear speech can be better assured that the client is receiving the communicative message accurately, with less need for repetitions. In aural rehabilitation settings, education about clear speech is routinely provided to the family members of listeners with hearing loss (Schum, 1996). This practice provides an effective approach to communication for families and caregivers, resulting in fewer communication breakdowns and less frustration for both speaker and listener. Moreover, clear speech can be useful in educational settings, such as with children with learning disabilities (Bradlow et al., 2003), among others. The teacher who uses clear speech in the classroom can convey the day's lessons with greater confidence that the children are receiving the message accurately. Overall, using clear speech in different settings can facilitate more effective communication.

Given the widespread potential and many applications of clear speech, investigation of its benefits remains an active area of research. Recently, investigators have attempted to compare results between studies in order to synthesize the results of multiple intelligibility experiments with clear speech (Krause & Braida, 2002; Liu, Del Rio, Bradlow, & Zeng, 2004; Panagiotopoulos, 2005; among others). These comparisons are necessary to determine if the clear speech benefit received by different listeners is the same across different environments or if some listeners receive a greater benefit in one or more environments. For example, clear speech intelligibility benefits found for older listeners (Panagiotopoulos, 2005) have been compared to benefits found for younger listeners (Krause & Braida, 2002). Unfortunately, comparisons between studies are generally difficult due to differences in presentation levels, environments, listeners and stimuli. When such differences exist, absolute performance levels are affected and confound comparisons of relative performance. Consequently, studies reporting a clear speech benefit relative to conversational speech should not be compared unless absolute performance is controlled. To facilitate such comparisons, psychometric functions, which characterize listener performance (plotted on the y-axis) as a function of stimulus input (plotted on the x-axis), are needed. However, there is currently only one clear speech study with psychometric functions calculated for a limited number of normal hearing and cochlear implant listeners in noise (Liu et al., 2004), and it is unknown to what extent these functions would vary with different types of speech materials. Therefore, in order to assess the validity of comparisons and measurements of the clear speech benefit across populations and environments that have been made to date, and to conduct additional such comparisons, further research is needed to fully characterize the psychometric functions of clear speech for normal hearing listeners.

Chapter 2

Literature Review

Background

Picheny et al.'s (1985) landmark study was conducted to assess the intelligibility benefit of clear speech for hearing impaired listeners. In this study, three male talkers who had some experience with public speaking or knowledge of clear speech recorded nonsense sentences using both clear and conversational speaking styles. These sentences were syntactically comparable to simple English sentences, but semantically anomalous. Five listeners with sensorineural hearing loss heard the sentences via headphones, with two frequency-gain characteristics and three presentation levels. The proportion of key words each listener identified in the clear and conversational conditions were compared. Results showed an average intelligibility benefit of 17 percentage points for the sentences spoken clearly across the different presentation modes. The compelling results from Picheny et al.'s study prompted further research into the attributes of clear speech. Inspection of acoustic features revealed many acoustic differences between clear and conversational speech (Picheny et al., 1986).

Acoustic Properties of Clear Speech

Several acoustic characteristics that distinguish clear from conversational speech have been identified. Temporal envelope modulations, the slowly varying amplitude changes that naturally occur in speech, are measured using a temporal modulation index. Clear speech has been found to have a greater temporal modulation index than conversational speech (Krause & Braida, 2004; Liu et al., 2004) for frequencies less than 3-4 Hz. Phonetic features identified with clear speech include higher average fundamental frequency, a wider frequency range and expanded vowel space (Bradlow et al., 2003; Ferguson & Kewley-Port, 2002; Krause & Braida, 2004; Picheny et al., 1986). In addition, clear speech has some distinguishing phonological characteristics. These include fewer vowel modifications and stop burst eliminations, a decrease in the use of alveolar flaps and an increase in schwa insertions (Bradlow, et al., 2003; Picheny et al., 1986). One of the more striking characteristics associated with clear speech, however, is its reduced rate. Picheny et al. (1986) noted that clear speech, with nonsense sentences, has a markedly slower rate of 91 to 101 words per minute (wpm), a striking contrast to conversational sentences which were 160 to 200 wpm. This noticeable difference led researchers to question the influence of rate on intelligibility.

Role of speaking rate. In order to determine if rate could be manipulated to improve intelligibility, two groups of researchers artificially altered the rates of clear and conversational speech. Whether the time-scale of sentences were adjusted uniformly (Picheny, Durlach, & Braida, 1989) or nonuniformly (Uchanski et al., 1996) by altering the duration of individual phonemes, overall intelligibility decreased for both clear sentences that were sped up to match the conversational rate and conversational sentences slowed down to match the clear rate. While the nonuniform time-scaling did not degrade the intelligibility of speech as much as uniform time-scaling, it nonetheless produced sped up clear speech that was less intelligible than conversational speech. Moreover, when sentences were reprocessed to their original rates, intelligibility levels returned to their previous levels, indicating that the decrease in intelligibility for processed sentences was not due to signal processing artifacts. In a related study, however, Uchanski et al. reported that clear sentences with deleted pauses were more intelligible than both unprocessed conversational sentences and conversational sentences with added pauses. The conversational sentences with added pauses were the least intelligible, and the altered clear sentences were not as intelligible as the original sentences that were produced clearly. So, while pause deletions decreased intelligibility for the clear sentences, the alterations did not completely remove the intelligibility gain, suggesting that the benefits of clear speech are not entirely attributable to a slower speaking rate.

Unsuccessful attempts to achieve clear speech at normal rates through artificial manipulations of conversational speech led researchers to investigate if clear speech could naturally be produced at a normal speaking rate. This was first attempted by Uchanski et al. (1996) with a professional speaker who was instructed to produce sentences as clearly as possible at rates up to 400 wpm. These sentences were presented to normal hearing listeners in quiet and noise, and to hearing impaired listeners. Although the talker was able to produce speech at different rates, no intelligibility benefits were observed. In a subsequent study, five talkers with backgrounds in public speaking received training for producing clear speech at normal (clear/normal) and quick (clear/quick) rates (Krause & Braida, 2002). The talkers also produced clear speech at slow (clear/slow) rates, and conversational sentences at slow (conversational/slow), normal (conversational/normal) and quick (conversational/quick) rates. The sentences were presented to normal hearing listeners in noise. Results showed that listeners received comparable intelligibility benefits with clear/slow and clear/normal speech, but benefits were not found with clear/quick speech. Krause and Braida’s research demonstrated that a slower speaking rate is not necessary for listeners to receive the intelligibility benefits associated with clear/slow speech. Thus, acoustic factors other than rate must be responsible for its increased intelligibility.

The Clear Speech Benefit

Although the specific acoustic characteristics responsible for increased intelligibility with clear speech have yet to be identified, its intelligibility benefits are well established and continue to be extended to additional listening situations. Researchers have measured intelligibility benefits with clear speech in a variety of listening environments, with different stimuli, talkers and for different listener populations. Though the amount of benefit received may vary from one study to another, all have shown some improvement in intelligibility with clear versus conversational speech.

Presentation environments. Of the many listening environments that have been used to assess the benefits of clear speech, several have focused on presentation in noise. By presenting stimuli with noise to normal hearing listeners, researchers attempt to simulate hearing loss in their listeners. As a result, presentation of clear speech in noise is the most frequently used test environment for normal hearing listeners. These environments include speech-shaped (Krause & Braida, 2002; Liu et al., 2004; Payton et al., 1994), broadband (Gagne, Rochette, & Charest, 2002) and Gaussian (Uchanski et al., 1996) noise. Speech-shaped noise, or filtered white noise, is created to match the average long-term spectra of the signal (e.g. Payton et al., 1994). Signal-to-noise ratios have varied, from -14 dB (Gagne et al., 2002) to +20 dB (Liu et al., 2004), but intelligibility benefits for clear speech have been consistently obtained. The benefits of clear speech in noise are generally sizeable, ranging from 4 rationalized arcsine transformation units (RAU) (Bradlow & Bent, 2002) to 25 percentage points (Liu et al., 2004).

Other presentation environments for which a clear speech benefit has been reported include low and high pass filtering. Krause and Braida (2003) designed a study to assess the intelligibility of two talkers' clear speech in a low pass environment (achieved by presenting the stimuli through 1/3 octave band filters with frequencies from 80 to 1000 Hz) and a high pass environment (achieved by presenting the stimuli through a 1/3 octave band filters centered at 3150 Hz). While a benefit in both environments was reported for both talkers' clear/slow speech, each talker achieved a benefit in only one environment with clear/normal speech. In the low pass environment, the female talker's clear/normal speech was more intelligible (15 percentage points) than her conversational/normal speech, while the male talker's clear/normal speech had a greater benefit (19 percentage points) in the high pass environment. Thus, large intelligibility benefits from clear speech are possible in filtered environments, although the benefit may be talker-dependent for clear speech at normal rates.

Another environment in which the intelligibility benefit of clear speech has been established is reverberation. Payton et al. (1994) tested normal hearing listeners in quiet and varying levels of noise in three different reverberant environments. An anechoic (ANEC) environment with no reverberation time (RT), a "living room" (LIVR), with 0.18 second RT and a "conference room" (CONF), with 0.60 second RT, were used. Results showed that as the noise, reverberation, and noise plus reverberation levels increased, clear speech intelligibility benefits increased. The benefit between clear and conversational speech ranged from 15 percentage points in the LIVR environment with no noise, to 34 percentage points in the CONF environment with a 0 dB SNR. Krause and Braida (2003) found similar results for clear/slow speech in the CONF environment as well as a 19 percentage point benefit for clear/normal speech presented in reverberation with their male talker. These results suggest a strong clear speech benefit in reverberant environments.

Presentation environments using audio-visual and visual-only modalities have also shown large benefits with clear speech. In a study with normal hearing listeners in noise, for example, Gagne et al. (2002) found a significant benefit for clear versus conversational speech when syllables were presented with a carrier phrase in audio-visual, visual-only and auditory-only modalities. The clear speech benefit averaged 7, 10 and 13 percentage points, respectively. Additionally, Helfer compared perception of nonsense sentences presented in auditory-only and auditory-visual conditions for both young (1997) and older (1998) listeners. Older listeners received an average clear speech benefit of 15 percentage points in the auditory-only condition and 11 percentage points in the auditory-visual condition. Younger listeners received similar benefits for clear speech, with averages of 14 and 18 points for auditory-only and auditory-visual modes. The total benefit between auditory-only conversational and auditory-visual clear speech for younger listeners was 32 percentage points. This difference equaled the sum of the benefit for clear speech in auditory-only and auditory-visual modes, suggesting that the visual presentation of clear speech provided additive intelligibility benefits.

Stimuli. Not only does clear speech provide an intelligibility improvement in a variety of presentation environments, but its intelligibility benefits are not limited to specific speech materials. A variety of stimuli have been used in clear speech experiments. In the landmark study by Picheny et al. (1985), as well as studies by several other investigators, grammatically correct but semantically anomalous sentences consisting of three to four key words were used with intelligibility benefits ranging from 15 to 34 percentage points (Helfer, 1997, 1998; Krause & Braida, 2002, 2003; Payton et al., 1994; Picheny et al., 1989; Smiljanic & Bradlow, 2005; Uchanski et al., 1996). Nonsense sentences were chosen to allow the talkers to vary their prosody as in natural speech, but without the benefit of semantic context for the listener. Meaningful sentences have also been used, from the Bamford-Kowel-Bench sentences (Bradlow & Bent, 2002; Bradlow et al., 2003, Liu et al., 2004) to the Johns Hopkins Lipreading Corpus (Schum, 1996). These meaningful stimuli have shown a clear speech benefit ranging from 4 (Bradlow & Bent, 2002) to 22 RAU (Schum, 1996). And, Gagne et al. (2002) found clear speech intelligibility gains of 7 to 13 RAU when CV and VCV syllables were presented with a carrier phrase. In sum, regardless of the diversity of stimuli presented in multiple studies, all listeners received an intelligibility benefit with clear speech.

Talkers. Intelligibility benefits with clear speech have also been obtained for a variety of talkers. Although some studies used recordings from talkers with some knowledge of clear speech (Payton et al., 1994; Picheny et al., 1985; Uchanski et al., 1996) and specifically selected talkers with a background in public speaking for eliciting clear speech at normal and quick rates (Krause & Braida, 2002), special training or a special population of talkers is not required for talkers to achieve a form of clear speech that is beneficial to intelligibility (Helfer, 1997, 1998; Picheny et al., 1985; Schum, 1997). In a study conducted to assess the intelligibility of older and younger talkers' clear and conversational speech, Schum (1996) discovered no statistically significant difference between the groups' ability to produce clear speech.

In general, it has also been reported that the benefit of clear speech is largely independent of talker, at least when the listener population and listening environment are held constant (Krause & Braida, 2002; Picheny et al., 1985). While the talker main effect was significant in Picheny et al.'s study, analysis of variance (ANOVA) revealed talker and talker x speaking mode accounted for only 7% and 2% of the variance, respectively. Similarly, Krause and Braida found the talker x speaking mode interaction was responsible for only a small percentage of the intelligibility variance, although the talker factor was statistically significant. Despite differences in talkers' overall intelligibility, a comparable clear speech benefit can be received.

In some cases, however, variability between different talkers' clear speech intelligibility must be considered. Even with closely matched acoustic features, different intelligibility outcomes have been observed between talkers (Goy, Pichora-Fuller, van Lieshout, Singh, & Schneider, 2007). Goy et al. presented identical high and low context sentences, recorded by two male talkers, to younger and older listeners in noise. Significant differences in intelligibility between the two talkers were noted for both "clear" and "normal" speech, despite the talkers' comparable average FO and rate. Acoustic characteristics of the more intelligible talker's target words included longer duration, higher FO and slightly increased intensity. A further source of variability is that talkers may use different strategies to produce clear speech, particularly when constrained to normal speaking rates. For example, two talkers made opposing adjustments in voice onset time and stop releases for clear/normal speech, and while these individual strategies resulted in similar intelligibility benefits in noise (Krause & Braida, 2002), substantial differences in the amount of benefit were observed in other environments (Krause & Braida, 2003). In a low-pass environment, a female talker's clear/normal speech was more intelligible, while a male's was more intelligible in high-pass and reverberant environments. These findings demonstrate how benefits with clear speech can be dependent on the strategies employed by the talker.

Populations. Perhaps the most notable advantage of clear speech is that it benefits a variety of listener populations. These include normal hearing listeners in quiet and noise (Gagne et al., 2002; Krause & Braida, 2002; Liu et al., 2004; Payton et al., 1994; Uchanski et al., 1996), hearing impaired listeners (Payton et al., 1994; Picheny et al., 1985; Uchanski et al., 1996), older listeners (Helfer, 1998; Panagiotopoulos, 2005), and listeners with cochlear implants (Liu et al., 2004). In addition, both children with learning disabilities (LD) and normally developing children (Bradlow et al., 2003) show higher intelligibility scores with clear versus conversational speech. However, because children with LD have lower overall intelligibility scores and their speech perception is affected by noise more than children without LD, with clear speech, children with LD obtain intelligibility scores that are comparable to the average non-LD children's scores for conversational speech.

Non-native English speakers are another population shown to benefit from clear speech (Bradlow & Bent, 2002; Krause & Braida, 2003). Although both non-native and native speakers of English receive intelligibility benefits with clear speech, only one study found the benefit to be smaller for non-native listeners (Bradlow and Bent, 2002). This difference was attributed to the amount of time that non-native listeners had been exposed to English, and to the possibility that native listeners may have received a contextual advantage from the meaningful sentence stimuli. Without the availability of context, Krause and Braida found no difference in the amount of benefit between the two groups for nonsense sentences in noise. They attributed the discrepancy between the two studies to the differences in stimuli and length of time the listeners had been exposed to English.

It is also worth noting that the clear speech benefit is not restricted to English. One study compared perception of clear speech nonsense sentences in noise, produced in Croatian and English for their respective native listeners (Smiljanic & Bradlow, 2005). The average intelligibility benefit with clear speech was 16 RAU for English and 15 RAU for the Croatian listeners. There was no significant effect of language, and the authors reported that talkers in both languages demonstrated a reduced rate of speech, expanded vowel space and an increase in pitch range. These findings suggest that applications of clear speech in noise or with a hearing impaired listener would be beneficial, regardless of specific language.

Summary

Clear speech provides benefits to listeners in many different environments. Even with a variety of stimuli, clear speech intelligibility is greater than intelligibility for conversational speech. Talkers may vary the strategies they employ when producing clear speech, but the benefit remains. However, the extent to which the benefit varies across diverse listening situations has not been well characterized. More research is needed to determine how the clear speech benefit for each listener population changes with environment, stimuli, and talker.

The Need for Psychometric Functions

Given that the clear speech benefit applies to so many listening situations, questions have arisen regarding which populations and environments benefit most. To answer these questions, researchers have attempted to compare their findings, with clear and conversational speech at slow and normal rates, with results in other investigators' studies for different populations or environments. These comparisons are complicated by inconsistent methods in the form of participants, stimuli and presentation environments across studies. Therefore, comparisons are at times poorly justified due to insufficient matching of listener performance at baseline (i.e., conversational speech intelligibility) and psychometric functions with clear speech, at normal and slow rates, are needed.

Psychometric functions for speech perception are represented by a graphical plot of performance (e.g., percent key words correct), on the y-axis, in relation to stimulus level or SNR, on the x-axis. Psychometric functions are available for conversational speech in noise, and have been used to help predict performance for normal and hearing impaired listeners (Wilson & Strouse, 1999). These same functions are needed for listeners with clear speech to determine relative intelligibility benefits of clear speech across various levels of degradation.

Many factors contribute to the intelligibility of speech and as a result, the psychometric functions for clear speech could be substantially different than those for conversational speech. Individual speech sounds vary in length, intensity and frequency, with the changing patterns of the sounds contributing to perception and recognition by the listener (French & Steinberg, 1947). One predictor of speech intelligibility is the articulation index (AI), which determines intelligibility of speech sounds with a given frequency, intensity and noise level (French & Steinberg, 1947). However, the AI has not been shown to be an accurate predictor of intelligibility for clear speech (Payton et al., 1994), suggesting that psychometric functions for clear speech are likely to be different from those of conversational speech.

Even if the psychometric functions for clear and conversational speech are similar, the psychometric functions that have been obtained previously for conversational speech with word lists (Wilson & Strouse, 1999) and CV/VC syllables (Miller & Nicely, 1955) may not be appropriate for nonsense sentence materials frequently used in clear speech experiments. One factor affecting slope of a psychometric function is the type of test materials (Wilson & Strouse, 1999). Average listener response for easier or more homogeneous stimuli, such as spondaic words, will have a steeper slope than sentences, which present a more difficult perception task. In addition, large intersubject variability results in a flatter slope, small variability between subjects provides a steeper slope and a greater number of test subjects help to lessen the influence of intersubject variability. Previous studies may not have controlled for intersubject variability by using a small number of listeners and/or stimuli. Therefore, complete psychometric functions for both conversational and clear speech, at normal and slow rates, are needed. The increased knowledge that could be gained from listener psychometric functions with clear speech would provide a foundation for researchers to make more legitimate comparisons between studies, taking baseline performance into account.

Comparisons across studies

Researchers currently tend to compare the relative benefits of clear speech with results reported in other studies, without regard to absolute levels of performance. The validity of such comparisons relies on the assumption that underlying psychometric functions have similar slopes. Most comparisons that have been reported are for normal hearing and hearing impaired listeners. Multiple studies with hearing impaired listeners have shown comparable clear speech intelligibility benefits, from 15 to 17 percentage points (Payton et al., 1994; Picheny et al., 1985; Uchanski et al., 1996). This level of consistency is somewhat unexpected, since listeners with hearing impairments have greater variability with word recognition tasks (Wilson & Strouse, 1999) and baseline performance was not controlled, but the studies were conducted with similar listeners, and used the same stimuli and presentation environments (Payton et al., 1994; Picheny et al., 1985; Uchanski et al., 1996). Comparing studies between normal hearing and hearing impaired listeners has also been undertaken, but results are less straightforward. In an effort to simulate hearing loss or difficult listening situations, normal hearing listeners have been presented different types of stimuli in varying environments. The number of listeners per condition has ranged from 5 (Uchanski et al., 1996) to 12 (Gagne et al., 2002). To accurately compare the results for normal hearing listeners to individuals with hearing loss, basic psychometric functions for a large number of subjects with multiple signal-to-noise ratios and similar stimuli are needed.

Research comparisons with clear speech have also been made between older and younger listeners. Panagiotopoulos (2005) presented nonsense sentences at 0 dB SNR to older listeners to measure intelligibility of clear and conversational speech at different rates. Results were compared to data from Krause and Braida's (2002) study of young normal hearing listeners, which used the same sentence lists with presentation at -2 dB SNR. Intelligibility for conversational/normal and clear/normal speech was the same in both studies, and clear/slow speech intelligibility was comparable. An unexpected finding in Panagiotopoulos' research was a 21 percentage point benefit for the older listeners with conversational/slow speech. This was larger than the clear/normal speech benefit and roughly equal to the benefit observed with clear/slow speech. The conversational/slow intelligibility benefit may be due to acoustic properties of this condition that are more beneficial to older listeners than younger listeners. Another explanation may be that there is a ceiling effect for older listeners with slow speech at 0 dB SNR, which could also occur with younger listeners. Due to the differences in signal-to-noise ratio between Panagiotopoulos' and Krause and Braida's studies, direct comparisons of the results are difficult.

Other research comparisons have been conducted between listeners within a study. For example, Bradlow and Bent (2002) compared native and non-native listener performance. They reported a decrease in the clear speech benefit for non-native listeners and proposed that this was due to the listener's length of exposure to the language. In contrast, Krause and Braida (2003) found the same clear speech benefit for both native and non-native listener groups. Neither study attempted to control baseline performance in order to ensure that the across-group comparisons were valid. Moreover, differences in signal-to-noise ratio and stimuli (meaningful vs. nonsense sentences) complicate comparisons between the two studies. To truly compare intelligibility benefits for native and non-native listeners, similar stimuli and performance levels should be used.

Statement of the Problem

Comparisons of clear speech data obtained in different studies must be interpreted with caution until basic psychometric functions are known. Researchers need to be able to compare results across studies, where superior or inferior performance between subjects has been controlled. The influence of intersubject variability can be reduced with a larger pool of participants and greater number of stimuli (Wilson & Strouse, 1999). Characterization of basic functions for young normal hearing listeners will provide a foundation for comparing intelligibility with older listeners, hearing impaired listeners and other populations. These basic functions are needed to provide information about listener performance with clear speech at different rates and noise levels. Finally, listener functions with nonsense sentences are necessary to control for any influence of semantic context (Picheny et al., 1985).

There is some research providing psychometric functions for clear speech. Liu et al., (2004) reported intelligibility of clear and conversational speech as a function of SNR for normal hearing listeners and for cochlear implant users. In one of four experiments, five listeners heard clear and conversational speech recorded by a female talker, and six listeners heard a male's speech. The stimuli consisted of meaningful sentences in signal-to-noise ratios of -20 to +20 dB in 5 dB presentation steps. Each listener heard only eight clear and eight conversational sentences at each SNR. When psychometric functions were graphically plotted, the slope for clear speech stimuli was steeper than the slope for conversational speech stimuli, with an average intelligibility benefit of 29 percentage points. Although these psychometric functions contribute important information regarding the relative intelligibility benefits of clear speech at various SNRs, to what extent these functions are applicable to other types of speech materials is not known. The 29 point benefit with clear speech, larger than the benefit found in previous studies with nonsense sentences, may be due to contextual cues received by the listeners or to the range of SNRs tested. In addition, intersubject variability may have skewed the results due to the small number of participants and limited stimuli presented.

Since nonsense sentences are frequently used in clear speech research, the purpose of this study is to determine how speech intelligibility, as measured by percentage key words correct in nonsense sentences by young adults with normal hearing, varies with changes in speaking condition, talker and varying signal-to-noise ratios.

Chapter 3

Methods

Participants

A total of 40 participants (31 females, 9 males; ages 18 to 38 years) were selected for the study. Demographic information for participants is listed in Appendix A. For inclusion in the study, participants were required to be between 18 and 40 years of age and to pass an audiological hearing screening. Eligible participants had pure tone thresholds of 20 dB HL or below at 250, 500, 1000, 2000 and 4000 Hz in at least one ear. Participants who were not native English speakers or who did not have a high school diploma or its equivalent were excluded from the study.

Materials

The sentence materials used for the listening sessions consisted of grammatically correct, but semantically anomalous, sentences created by Picheny et al. (1985) and previously recorded by Krause and Braida (2002). These nonsense sentences allowed the listeners to receive prosodic information without any contextual cues (Picheny et al., 1985). An example of a nonsense sentence is *"His right cane could guard an edge."*

The sentence lists were previously recorded by four talkers from Krause and Braida's (2002) study. All of the talkers had experience with public speaking.

Of the five talkers in that study, one male (T5) and three female (T1, T3, T4) talkers were selected based on their ability to manipulate both rate (slow vs. normal) and intelligibility (clear vs. conversational).

A total of nine distinct 50-sentence lists were selected from the Krause and Braida (2002) database. One list was used for practice. The remaining eight lists were presented twice, once in each of the two speaking styles. Specifically, for each of the four talkers, one list was presented in clear/slow and conversational/slow styles, and one list in clear/normal and conversational/normal styles. Thus, 800 utterances (8 lists x 50 sentences x 2 recordings) were divided evenly between four different speaking styles, and 200 utterances (50 utterances per talker) were presented in each of the following modes: conversational at a normal rate (conversational/normal), conversational at a slow rate (conversational/slow), clear at a normal rate (clear/normal) and clear at a slow rate (clear/slow). For presentation, the eight lists were sorted into blocks of four, as shown in Table 1.

Table 1. Sentence Blocks

Block	Talker/Sentence Mode	List
A	T1 – Conversational/normal	List 1
	T1 – Clear/slow	List 2
	T4 – Clear/normal	List 3
	T4 – Conversational/slow	List 4
B	T3 – Clear/normal	List 5
	T3 – Conversational/slow	List 6
	T5 – Conversational/normal	List 7
	T5 – Clear/slow	List 8
C	T1 – Clear/normal	List 1
	T1 – Conversational/slow	List 2
	T4 – Conversational/normal	List 3
	T4 – Clear/slow	List 4
D	T3 – Conversational/normal	List 5
	T3 – Clear/slow	List 6
	T5 – Clear/normal	List 7
	T5 – Conversational slow	List 8

Presentation Sessions

During the presentation sessions, each participant was assigned a workstation with a keyboard, monitor and headphones in the group lab of the Communication Sciences and Disorders department of the University of South Florida. Because this lab is designed to accommodate four participants at one time, there were a total of ten groups of participants, with four participants per group. Due to scheduling conflicts, however, some participants were alone in the lab during their sessions. The tester monitored all sessions from an adjacent lab that contained a window providing a view of the participant(s), as well as audio feedback via microphones placed in the group lab.

Each group or individual participant was scheduled to attend four weekly two-hour presentation sessions. In each session, participants heard one block of four 50-sentence lists, presented monaurally over Sennheiser HD265 headphones. After listening to the first and third lists, participants received a five minute break, and a ten minute break was given after the second list. The default presentation level was set to 85 dB SPL, and each participant had the opportunity to select a preferred listening level (which was then fixed for the duration of the experiment) within 6 dB of the default level, adjusted in 3 dB increments.

Sentences were presented in the presence of speech-shaped noise taken from the *Hearing In Noise Test* (Nilsson, Soli, & Sullivan, 1994). As shown in Table 2, a total of five signal-to-noise ratios were used (-4, -2, 0, +2, +4 dB), but due to limitations in the number of sentences, each listener group heard sentences presented at only four of the noise levels. The signal-to-noise ratio was varied across listening sessions, with one noise level for each of the four sessions (i.e. sentence blocks). The order of presentation for signal-to-noise ratio, also shown in Table 2, was from the most difficult in session one to the easiest in session four. This order was chosen to minimize fatigue and to maintain the participant's interest in the study.

Finally, Table 2 shows the presentation schedule for the sentence lists. In order to minimize learning effects within a list during the experiment, listeners received a one-week break in between sessions two and three. This break provided a three week time period between repeat lists by the same talker, at slow or normal rates (i.e. Blocks A and C; Blocks B and D).

Table 2. Sentence Presentation Schedule

	Session 1		Session 2		Session 3		Session 4	
Group	SNR	Block	SNR	Block	SNR	Block	SNR	Block
1	-4	A	-2	B	0	C	2	D
2	-2	A	0	B	2	C	4	D
3	-4	D	0	A	2	B	4	C
4	-4	C	-2	D	2	A	4	B
5	-4	B	-2	C	0	D	4	A
6	-4	C	-2	D	0	A	2	B
7	-2	C	0	D	2	A	4	B
8	-4	B	0	C	2	D	4	A
9	-4	A	-2	B	2	C	4	D
10	-4	D	-2	A	0	B	4	C

For each group session, the tester sat at a personal computer and presented one sentence at a time via a Matlab software program, and the participants typed their responses in an Excel spreadsheet. Each participant signaled when they were ready for the next sentence by turning a page on a flip chart at his/her workstation that was visible to the tester. The tester did not present the next sentence until all participants signaled they were ready to continue. Participants were also given the option of verbally requesting a pause, asking a question or raising their hand to signal a request for the tester to pause at any time during the session. For single participant sessions, participants controlled the rate of presentation of sentences from their workstation. In all sessions, each participant’s responses were saved as a text file after each sentence list was completed, and a new spreadsheet was generated for the next sentence list.

At the first session, participants received instructions and listened to 10 to 15 practice sentences to become familiar with the task. They were informed that the sentences would not make any sense and that the background noise might make it difficult to understand the words. They were encouraged to make their best attempt at typing a complete sentence, guessing if necessary. They were informed that their responses provided valuable information, whether their answers were correct or not, as there was no “right” or “wrong” response. For subsequent sessions, participants received brief instructions reminding them to do their best to type what they heard.

Participants typed their responses to the practice sentences in the Excel spreadsheet and turned the pages on their flip charts to become accustomed to the session procedures. The participants did not receive any feedback from the tester regarding the correctness of their responses, but were reminded to make their “best guess.” The first five practice sentences were presented in quiet. This format was intended to allow participants to hear the nonsense sentences without any competing noise to adjust to the unfamiliar sentence structure. The next group of sentences was presented with speech-shaped noise, and participants then provided feedback about the presentation level. The tester adjusted the gain plus or minus 3 dB or 6 dB, per participant request, and recorded each adjustment for future sessions. After the participants completed the practice sentences and the tester addressed any questions that were posed, the first sentence list was presented.

Scoring Procedures

Each sentence was scored for three to four key words, consisting of all nouns, verbs and adjectives in the sentence. Responses were scored up to two times, once with an autoscoring program in Matlab, which graded each sentence as “correct” if the participant’s response contained all of the key words with exact spelling. A second text file containing any sentence with an error was then generated from Matlab. This file was hand graded by one of three graders, and credit was given for homophones, nouns with a plural added or deleted, changes in verb (- ed) past tense, or spelling errors. For example, if the target sentence was *“Her blond shore grins at her manner”* and the response was *“The wand shores grinned at her manor”,* Matlab placed this sentence in an error file and scored the entire sentence as incorrect. The key words in this sentence are “blond,” “shore,” “grins” and “manner.” The tester credited the respondent’s sentence with three correct key words (“shores,” “grinned” and “manor”). The number of key words correct were totaled and then divided by the total possible key words in the sentence list to obtain the percentage of key words correct for each list.

Data Analysis

The percent-correct scores from the listening sessions were then compiled to construct 16 psychometric functions for clear and conversational speech at five signal-to-noise ratios. Data from eight listeners comprised each data point on each psychometric function. Due to the limited number of sentence lists, the same eight listeners did not make up the data points in a particular function. However, based on previous studies, responses from normal hearing listeners have minimal variability and should be relatively interchangeable (Wilson & Strouse, 1999). There were four functions per talker: clear/slow, clear/normal, conversational/slow, conversational/normal. Functions were compared between the clear and conversational conditions of the same rate to see how the size of the clear speech benefit varied with talker and/or signal-to-noise ratio.

Chapter 4

Results

The purpose of this study was to examine how speech intelligibility, measured as percent key words correct, varied with condition, talker and signal-to-noise ratio for young normal hearing listeners. In addition to assessing the relative intelligibility of each condition on average, the data collected were examined for talker differences by constructing 16 graphs (4 talkers x 4 conditions) depicting psychometric functions, with intelligibility as a function of signal-to-noise ratio. A three-way analysis of variance (ANOVA) was also performed on key word scores after an arcsine transformation ($\sqrt{I_j}/100$) was applied to equalize the variances. Between-subject factors were condition (clear/slow, clear/normal, conversational/slow and conversational/normal), SNR (five levels) and talker (four levels). As shown in Table 3, all main effects and most interactions were statistically significant ($p < 0.01$); each of these will be discussed in detail in the following sections.

Table 3. Between-subject Effects and Variables

Source	Type III Sum of Squares	df	Mean Squares	F	Sig.	Partial Eta Squared
talker	4.167	3	1.389	263.785	0.000*	0.586
SNR	18.960	4	4.740	900.161	0.000*	0.865
condition	3.500	3	1.167	221.554	0.000*	0.543
talker x SNR	0.220	12	0.018	3.486	0.000*	0.070
talker x condition	2.470	9	0.274	52.116	0.000*	0.456
SNR x condition	0.119	12	0.010	1.877	0.034	0.039
talker x SNR x condition	0.419	36	0.012	2.208	0.000*	0.124

* $p < 0.01$

Tables B1-B4 in Appendix B list key word scores obtained by all listeners in all conditions for each of the four talkers. These results are summarized in Figure 1, which shows a large effect (partial $\eta^2 = .543$, $p = 0.000$) of condition. As expected, conversational/normal was the least intelligible condition (53%), since this speaking style represents “typical” speech (normal rate, with no particularly emphasis on clarity). Therefore, this condition was considered the baseline for measuring intelligibility benefits for the other three conditions. Post-hoc t-tests were conducted to evaluate pairwise comparisons of all conditions (see Table C1, Appendix C). Slow conditions were the most intelligible, with benefits of 15 ($p = 0.000$) and 17 ($p = 0.000$) percentage points for conversational/slow and clear/slow, respectively. Clear/normal speech also provided an intelligibility benefit of 11 percentage points ($p = 0.000$).

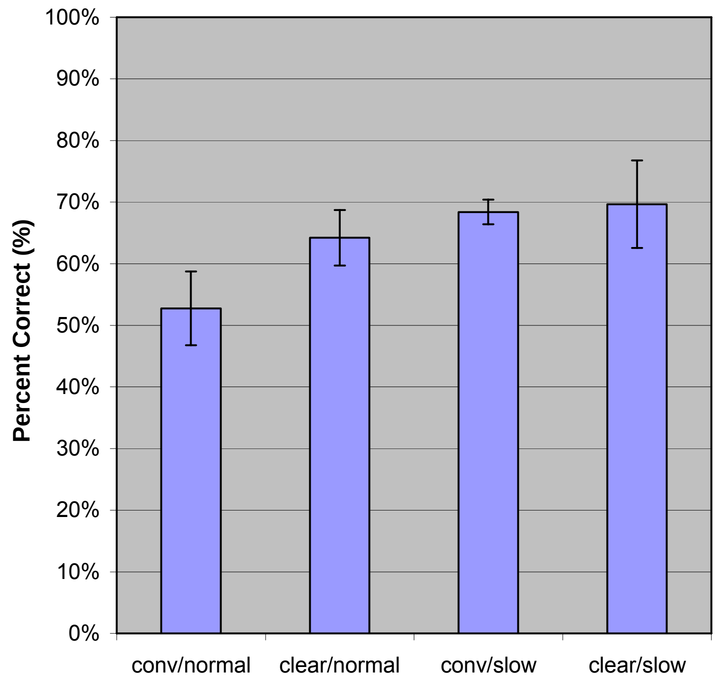


Figure 1. Average intelligibility, in percent key words correct, for each test condition.

For clear/slow and clear/normal speech, intelligibility benefits of this size are consistent with benefits reported in previous studies of young, normal hearing listeners tested under similar conditions (nonsense sentences presented in noise). For example, the average intelligibility benefit of clear/slow speech in the current study, 17 percentage points, is comparable to benefits reported for three previous studies that utilized nonsense sentences in noise. Helfer (1997) and Krause and Braida (2002) reported listener intelligibility benefits of 14 and 18 percentage points, respectively, and Smiljanic and Bradlow's (2005) listeners received benefits of 16 RAUs. Similarly, with clear/normal speech, the current group of listeners received an 11 point benefit, which is consistent with the 14 point benefit reported by Krause and Braida. For conversational/slow speech, however, the average benefit of 15 percentage points was much larger than the 6 point benefit reported for Krause and Braida's listeners. This difference is particularly surprising given that the stimuli used in the current study were obtained from the same database used in Krause and Braida's research.

One explanation that might account for this discrepancy is that the average intelligibility benefit across all five SNRs employed in this study is not the most valid metric for comparing to previous studies that employed only one SNR. Therefore, comparisons were further narrowed to benefits obtained at the same SNR. Using this method, consistency of results with previous research remains strong for clear/slow speech: at -2 dB SNR, the present listeners received intelligibility benefits of 20 points, versus Krause and Braida's (2002) 18 points; at 0 dB SNR, the benefit was 19 points, compared to Smiljanic and Bradlow's (2005) listener benefit of 16 RAUs, and at +2 dB SNR, listeners in the current study received a 14 point benefit, the same benefit reported for Helfer's (1997) listeners. Another strong comparison is the clear/normal speech benefit, at -2 dB SNR, which was 14 percentage points for listeners in the both the present and Krause and Braida's study. Yet, the inconsistency for conversational/slow speech benefits remains between the current study (19 points) and Krause and Braida's research (6 points), even when benefits are compared only at the SNR used in both studies, -2 dB. In fact, at 0 dB, the 19 point intelligibility benefit listeners received in this study is closer to benefits received by older normal hearing listeners (23 points) at the same presentation level (Panagiotopoulos, 2005).

Psychometric Functions

Figure 2 shows the psychometric functions for the four conditions, averaged across talkers. As expected, a large effect of SNR was observed (partial $\eta^2 = .865$, $p = 0.000$), with listener performance improving with SNR. However, the psychometric functions were constructed to assess intelligibility benefits of each condition as a function of signal-to-noise ratio, or the SNR x condition interaction. Again, intelligibility for conversational/normal speech was used as the baseline for measuring benefits for the other three conditions at each SNR. The broad pattern of the overall results, seen in Figure 1, was not affected by changing signal-to-noise ratio (Figure 2). Measurements between conditions, as well as visual inspection, indicate a roughly constant benefit from -4 to 0 dB SNR. However, the interaction of SNR with condition approached significance ($p = .034$) because the size of intelligibility benefits for all conditions decreased for the higher SNRs (+2 and +4 dB). This change in benefit can be observed for the clear/slow condition. At +4 dB SNR, there is an 11 percentage point intelligibility benefit with clear/slow speech that increases to 21 points at -4 dB SNR. Similarly, with clear/normal speech, the largest benefit (14 points) was observed at -4 dB and the smallest benefit (8 points) at +4 dB. The reason for the decrease in benefits at +2 and +4 dB SNR may be due to a ceiling effect, where intelligibility is asymptotically approaching maximum levels. For the SNR range of -4 to 0 dB, on the other hand, the intelligibility of each condition increases linearly with SNR, suggesting that this region represents the “middle” of the psychometric function.

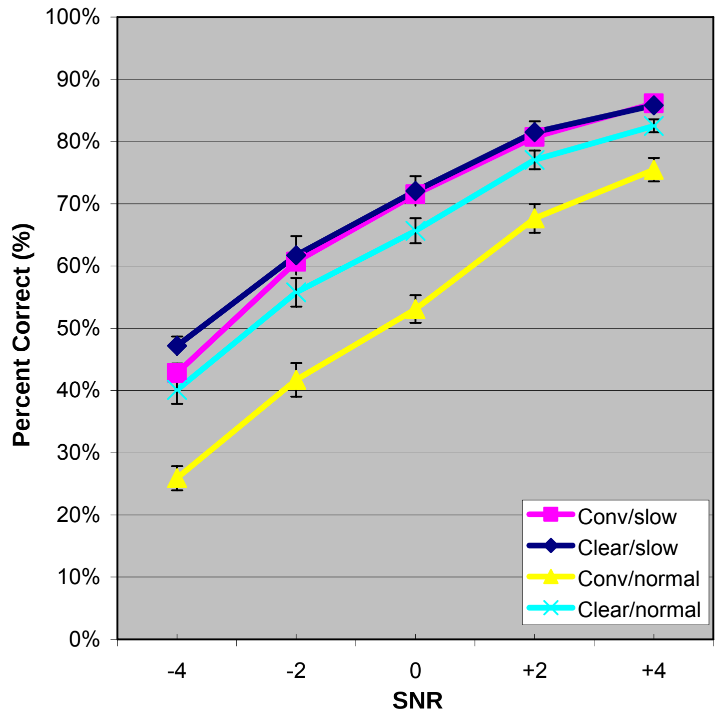


Figure 2. Average psychometric functions, in percent key words correct, across signal-to-noise ratio.

Defining the "middle" or linear portion of the psychometric function and determining whether the relative intelligibility of each condition varies in this region is one of the central purposes of this study. The data in Figure 2 and individual talker data in Figure 4 suggest that the middle of the psychometric function for young, normal hearing listeners presented with nonsense sentences in noise corresponds to the region for which intelligibility for the conversational/normal condition ranges from 20% to 65%. Over this region, the psychometric functions for all conditions are approximately linear. Because the slopes for all conditions were relatively similar in this region, the relative difference in intelligibility between the psychometric functions for each condition remained fairly constant, with a change of 2 percentage points or less in the benefit size for all conditions. For instance, the intelligibility benefit for conversational/slow speech was 19 points at 0 dB SNR, and 17 points at -4 dB SNR. Excluding the ceiling effect, these results demonstrate a small change in the average size of intelligibility benefits across signal-to-noise ratio. This finding suggests that young, normal hearing listeners can receive consistent and predictable benefits for a given range of noise levels.

Effect of Talker

Not surprisingly, some talkers were more intelligible overall than other talkers, and the ANOVA (Table 3) showed a large effect of talker (partial η^2 = 0.586). Post-hoc t-tests (see Appendix C, Table C2) confirmed that T3, a female, was the least intelligible talker at 53% ($p = 0.000$) on average and the most intelligible talker overall was T5, a male, at 73% ($p = 0.000$) on average across all conditions.

For the purposes of this thesis, however, the goal was to compare how the intelligibility benefits of each condition varied across talkers. Therefore, the talker x condition interaction was of primary interest. This interaction was significant ($p = 0.000$) and showed a large effect size (partial η^2 =0.456), nearly as large as that of talker alone. To assess this interaction, Figure 3 shows the average key words scores for each talker in each condition. T1's data generally followed the overall pattern obtained for all talkers, with clear/slow and conversational/slow speech as the most intelligible conditions, followed by clear/normal speech, and with conversational/normal speech as the least intelligible condition. Even so, the intelligibility benefit of T1's clear/normal speech was somewhat smaller (6 points) than the benefit of this condition on average across all talkers (11 points). Patterns for the other three talkers differed more substantially from the overall results. For example, neither T4 nor T5 produced conversational/slow speech that provided as large an intelligibility advantage as clear/slow speech. Moreover, T5's clear/normal speech provided a substantially larger benefit (13 points) than his conversational/slow speech (5 points), nearly as large a benefit as that provided by his clear/slow speech (17 points).

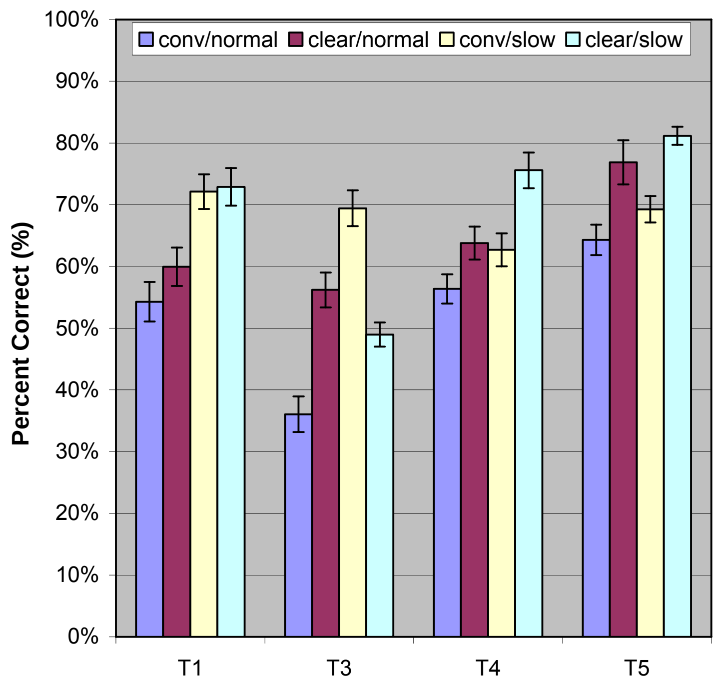


Figure 3. Average intelligibility, in percent key words correct, for each talker in each condition.

On the whole, however, it was T3’s configuration that differed from the average more than any talker, with intelligibility benefits of 20 points in the clear/normal and 33 points in the conversational/slow conditions, compared to 11 points and 15 points for these conditions on average. Moreover, it is worth noting that T3 was the only talker who did not have the largest intelligibility benefit with the clear/slow condition. While any definitive reasons for T3’s relatively poor clear/slow intelligibility are unknown at this time, there are some possible explanations. In Krause and Braida’s (2002) study, clear/slow was the first condition that talkers produced, with no training, and the first condition recorded. It is possible that T3’s production of clear speech improved with subsequent tasks, resulting in greater intelligibility with her clear/normal condition, which involved intense training and listener feedback. Additionally, all of the talkers had public speaking experience, ranging from five to seven years, with the exception of T3, who had been a broadcasting student for two years and was the least experienced in the group.

Despite these variations, some similarities across talkers can be observed. Intelligibility for all talkers improved with speaking clearly, at a slow or normal rate, and/or with speaking slowly, in a conversational or clear mode. For three talkers, T1, T4 and T5, clear/slow was the most intelligible condition. And, for T3, T4 and T5, clear/normal was the second most intelligible condition. Therefore, clear speech provided an intelligibility benefit for listeners at both slow and normal rates.

Across talkers, the benefit of conversational/slow speech was the most variable of any condition (indicating that it was least correlated with a talker's conversational/normal intelligibility), with the size of the benefit varying as much as 28 points between different talkers (T3 - 33 points, T5 - 5 points). One explanation for this variability could be that this speaking style may represent an unnatural task for talkers. In the original study where the conversational/slow stimuli were recorded, Krause and Braida (2002) allowed their talkers to produce this particular speaking condition without any specific instructions other than to use a "normal" style for speaking slowly without any particular emphasis on clarity. This method of elicitation may have resulted in an unreliable mode of communication because it may not be realistic to expect talkers to have a "normal" style at slow rates that does not involve increased clarity. Thus, talkers may have found it difficult to slow their speech without adopting any of the acoustic characteristics of clear speech. Other talkers may have used atypical prosody that distorted the acoustic cues that listeners typically rely upon with speech. Note that with this degree of inter-talker variability, experiments that base conclusions on outcomes for single talkers are likely to obtain substantially different results, from each other and from studies using multiple talkers, regarding the benefits of conversational/slow speech.

In the clear/normal condition, for which talkers received training (see Krause & Braida, 2002), there was less variation in intelligibility benefit across talkers. The size of the benefit varied 14 points between T3 (20 points), and T1 (6 points). Even though this variability is somewhat smaller than that observed for the conversational/slow benefit, it is still sizeable and likely to be a factor in experiments investigating the benefits of clear/normal speech that involve only a single talker. The least variable condition was clear/slow, with the size of the intelligibility benefit varying just 7 points between T3 (13 points) and T4 (20 points) and only 3 points across three of the four talkers (T1 - 19, T4 - 20, T5 -17 points). The reduction in inter-talker variability for clear/slow speech is good news for researchers who employ only one or two talkers, as more reliable results can be expected. However, a 7 point difference in intelligibility is not trivial, and further work is needed to determine whether the variability introduced by T3 is typical of talkers, or whether she is an outlier and the 3 point variability is representative of most talkers.

Psychometric functions by talker. Further analysis of performance for each talker was conducted as a psychometric function of SNR in order to evaluate the combined effects of talker and SNR on the relative intelligibility of each condition. A medium size effect (partial $\eta^2 = 0.124$; $p = 0.000$) for the talker x SNR x condition interaction was observed (see Table 3). Figure 4 displays each talker's results, all of which reflect variations from the average psychometric functions shown in Figure 2. As with the average results for talker x condition, T1 (Figure 4a) was the only talker who followed the overall pattern relatively closely. But, in the middle of the psychometric function (described earlier as the linear region corresponding to conversational/normal intelligibility of 20% to 65%), from -4 to 0 dB SNR, benefits for T1 in all conditions were smaller at -2 dB than at 0 dB, rather than remaining constant. This change in benefit size with SNR may be explained by the conversational/normal psychometric function, which is not completely linear in this region. There appears to be a “drop” in intelligibility for conversational/normal speech at 0 dB and then a local “peak” at -2 dB, resulting in a larger benefit at 0 dB and a smaller benefit at -2 dB for all three conditions. For example, with clear/normal speech, T1’s maximal benefit is 11 points at 0 dB SNR, which decreases to 3 points at -2 dB and then increases to 8 points at -4 dB. Similarly, T3’s clear/normal benefit (Figure 4b) varies as much as 8 points (from a 16 point benefit at +4 dB to a 24 point benefit at -2 dB) in the middle of the psychometric function from -2 to +4 dB SNR (note that the linear region of T3’s psychometric function spans a different range of SNRs due to her lower overall intelligibility). In contrast, however, T3 shows a fairly constant benefit for the clear/slow (13 -17 points) condition in this region but even less consistency (benefits ranging from 26 - 39 points) for conversational/slow speech. In addition, the relative benefits for all conditions decreases at -4 dB SNR for T3, likely due to a floor effect, as her conversational/normal speech intelligibility is less than 20% at this presentation level. Although results for T4’s (Figure 4c) clear/slow condition are consistent with the average psychometric function, suggesting a relatively constant benefit from -4 to 0 dB SNR, this talker’s intelligibility varied considerably from the average pattern for both clear/normal and conversational/slow speech, which both showed larger benefits (12 – 13 points) at -2 dB, than at either 0 dB (6 points in both conditions) or -4 dB (3 and 5 points for clear/normal and conversational/slow speech, respectively).

Perhaps the most notable differences from the average psychometric functions were demonstrated by T5 (Figure 4d). In the middle of the psychometric function from -4 to 0 dB SNR, conversational/slow benefits decreased with worsening SNR while the intelligibility benefits of clear/slow and clear/normal speech increased over the same range. Moreover, the change in benefit over this range was sizeable: the clear/slow advantage increased from 19 to 33 points, and the clear/normal advantage increased from 16 to 23 points. So, while the benefits for each condition are preserved across SNR for talkers on average, this was not the case for individual talkers.

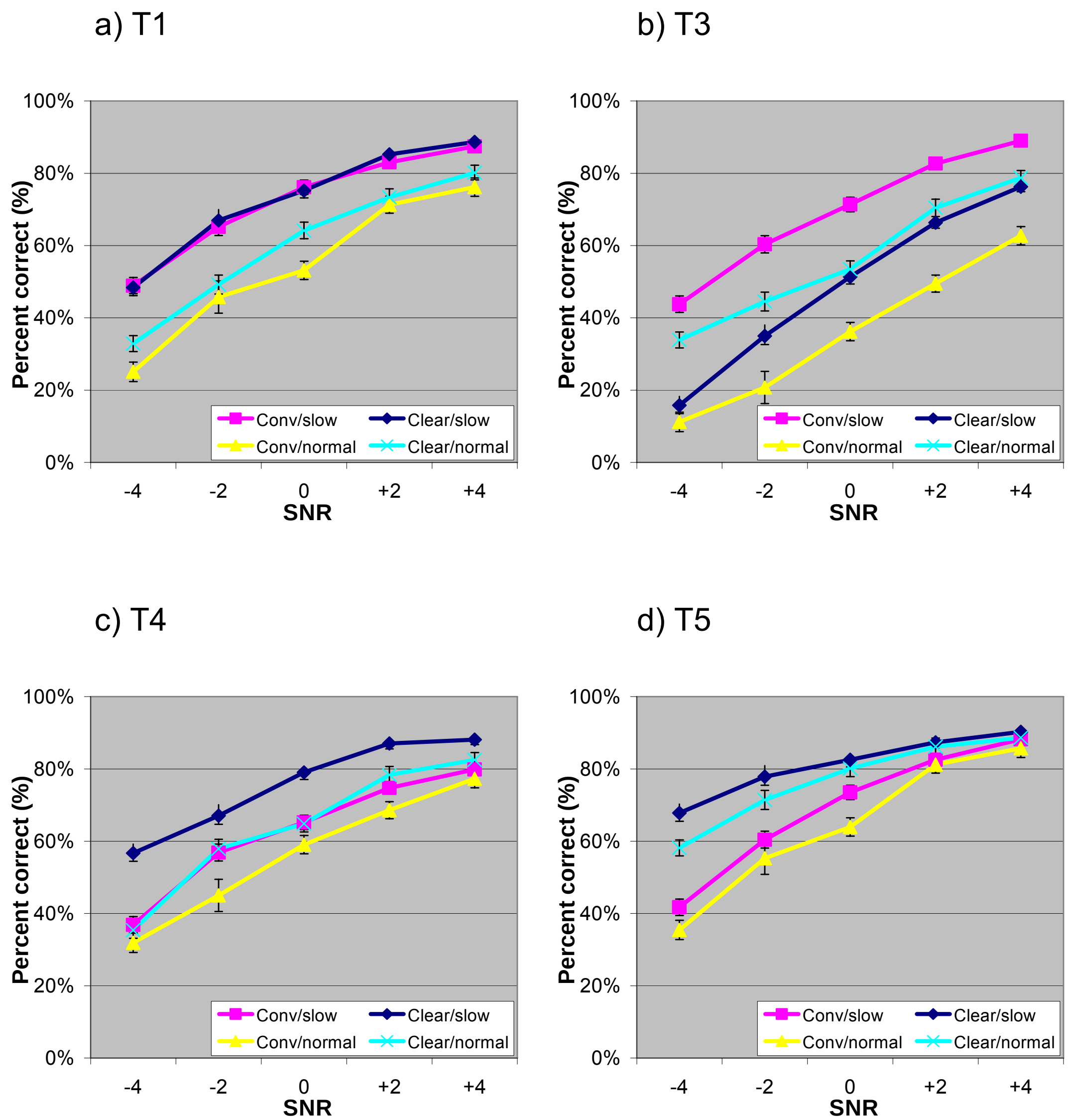


Figure 4. Average intelligibility, in percent key words correct, for each talker across signal-to-noise ratio.

Inspection of the psychometric functions for individual talkers also revealed substantial differences in the effect of SNR on the size of intelligibility benefits for each condition, represented in Table 4. To illustrate, with clear/slow speech, T5 had the greatest change in the size of benefit across SNRs. At -4 dB, his clear/slow speech provided a benefit of 33 points relative to the conversational/normal baseline of 35%, but it provided a benefit of only 4 points at 4 dB, where baseline intelligibility was 86%. Thus, the benefit of clear/slow speech varied up to 29 percentage points for T5, although his intelligibility at the easier listening conditions likely represents a ceiling effect.

Table 4. Individual talker intelligibility benefits by condition and SNR

Talker	SNR (dB)	Conv/normal (baseline) intelligibility	Clear/normal benefit	Conv/slow benefit	Clear/slow benefit
T1	**-4**	**25**	**8**	**24**	**23**
	-2	**46**	**3**	**19**	**21**
	0	**53**	**11**	**23**	**22**
	+2	73	0	10	12
	+4	76	4	11	13
Benefit range*	-	-	8	5	2
T3	-4	11	23	33	5
	-2	**21**	**24**	**39**	**14**
	0	**36**	**17**	**35**	**15**
	+2	**49**	**21**	**34**	**17**
	+4	**63**	**16**	**26**	**13**
Benefit range	-	-	8	13	4
T4	**-4**	**32**	**3**	**5**	**25**
	-2	**45**	**13**	**12**	**22**
	0	**59**	**6**	**6**	**20**
	+2	69	9	6	18
	+4	77	6	3	11
Benefit range	-	-	10	7	5
T5	**-4**	**35**	**23**	**7**	**33**
	-2	**55**	**16**	**5**	**23**
	0	**64**	**16**	**10**	**19**
	+2	81	5	2	6
	+4	86	3	2	4
Benefit range	-	-	7	5	14

* Benefit range is calculated over the middle of the psychometric function, where conversational/normal intelligibility ranges from 20% to 65%. Data points in the middle of the psychometric function are indicated in bold.

More importantly, how the size of the intelligibility benefit for each condition changed in the middle of the psychometric function (i.e. avoiding ceiling effects) for individual talkers was also measured. As shown in Table 4 (data points corresponding to the middle of the psychometric function are indicated in bold), there were substantial changes in intelligibility benefits with SNR for individual talkers. For 10 out of 12 talker-condition combinations (4 talkers x 3 conditions), the benefit varied by 5 or more percentage points across SNR. The two talker x condition combinations that varied by less than 5 percentage points (T1 and T3 in the clear/slow condition), suggest that for some talkers clear/slow speech can be somewhat less variable across a range of SNRs. Although the benefit of clear/slow speech for these talkers was only affected a small amount by SNR, such a result did not occur for all talkers. In fact, the largest range of benefits observed in any condition was with T5's clear/slow speech (14 points).

Conversational/slow speech had similar differences across SNR for individual talkers, but the SNR x condition difference between talkers fell within a somewhat smaller range of 8 points. Specifically, the conversational/slow benefit was least affected by SNR for talkers T1 and T5, whose benefits each changed 5 points across SNRs, and most affected for T3, whose intelligibility benefit varied by 13 points across SNR. The clear/normal condition x SNR had the smallest difference between talkers (3 points). However, the range of benefits across SNR still varied substantially, changing anywhere from 7 points across SNR for T5 to 10 points for T4.

Talker variability within condition. A further analysis of variability across talkers within condition was conducted. For each of the three conditions, a scatter plot in Figure 5 represents each talker's average intelligibility benefit at each SNR, in relation to the conversational/normal (baseline) intelligibility for that SNR. This depiction of the data allowed for comparisons between talkers and across SNRs while equating baseline intelligibility, so that the variability in intelligibility benefits for each condition can be visually examined. As can be seen in Figure 5c, there is little variation overall with clear/slow speech benefits for all talkers and SNRs. In the baseline range of 20 to 65%, benefits are fairly consistent. There is more variability with the clear/normal condition (Figure 5a) and conversational/slow speech (Figure 5b) appears to be highly variable. Moreover, conversational/slow speech shows a strong effect of talker, with T3 obtaining larger benefits than other talkers, even when baseline intelligibility is controlled. Thus, while the changes in intelligibility benefits for individual talkers across SNRs are considerable, benefits across talkers are generally predictable for clear/slow speech, and somewhat predictable for clear/normal speech.

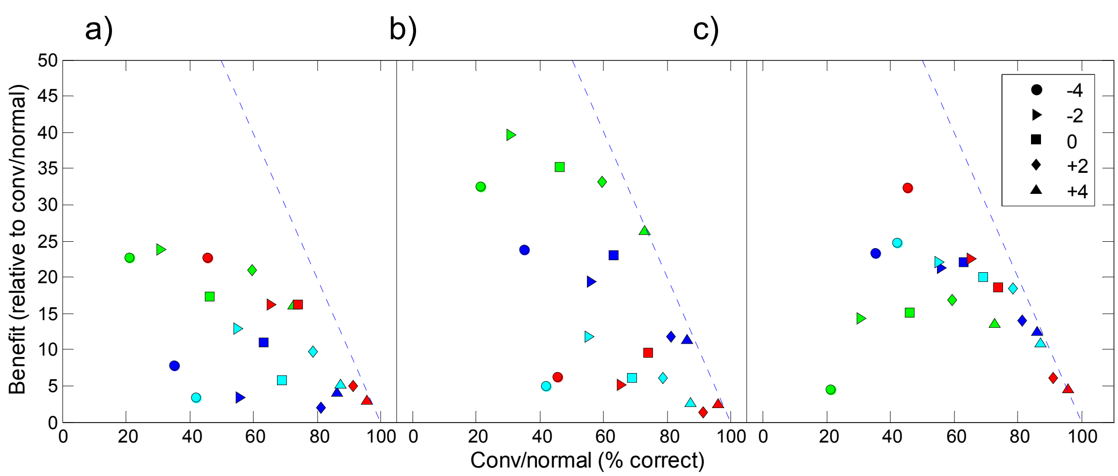


Figure 5. Intelligibility benefits by talker (T1 - blue, T3 - green, T4 - aqua, T5 - red) and SNR, relative to baseline performance, for each condition. Dotted line indicates maximum possible benefit for each baseline performance level.

However, even in the clear/slow condition, which was least variable, noticeable talker differences were observed. For example, with this condition, T3's conversational/normal speech was 36% intelligible at 0 dB SNR and listeners benefited by 17 percentage points. At this same baseline measure, but at a more difficult SNR of -4 dB, T5's listeners received an average intelligibility benefit of 23 points. This difference suggests that some talkers can achieve a greater intelligibility improvement with clear/slow speech than others, even when baseline performance is equated and all other experimental conditions (e.g. stimuli, listener population, etc.) are held constant.

Chapter 5

Discussion

Results indicate that, in noise, young normal hearing listeners benefit (relative to conversational speech at normal rates) when talkers speak clearly and/or when talkers speak slowly. More specifically, clear/slow and conversational/slow conditions provide the greatest intelligibility benefits on average, followed by clear/normal speech. Although intelligibility in all conditions generally decreases with worsening SNR, the average intelligibility *benefit* of the three conditions is maintained in the middle of the psychometric function, across signal-to-noise ratios of -4 to 0 dB. However, when individual talker data within a condition are examined, differences in intelligibility benefits are observed between talkers, on average and across noise levels.

When the changes in intelligibility benefits within condition are analyzed, the range of benefits across SNRs can be quite variable between different talkers, even in the middle of the psychometric function. T5, for example, had the largest change in benefit size across SNRs (14 points, i.e. benefits ranged from 19 to 33 points) for clear/slow speech. The clear/slow speech benefit of the other three talkers, however, was much less affected by SNR. In fact, without T5, the benefit change for each of the three talkers would have ranged from just 2 points (T1) to 5 points (T4) across SNRs. Similarly, for conversational/slow speech, T3 had the largest benefit range across SNRs (13 points), but without T3, the effect of SNR on this condition would have been a relatively small change in benefits across talkers, from 5 points (T1, T5) to 7 points (T4) in the middle of the psychometric function. So, the effect of condition and SNR contributed to the variability of benefits within and across talkers, suggesting that another group of talkers may produce very different results.

Reliability

Reliability for scoring of listener responses was examined to assess variability with sentence list scorers. As noted previously, a program in Matlab automatically scored each response list for key words correct, generating an error file for lists that contained one or more errors. Each error file was then re-scored and credit was given for misspelled words, homophones, and plural and tense errors. Three graders shared responsibility for re-scoring these files, creating the possibility for inter-rater reliability issues. Although the grading was largely objective in nature, particular concerns pertaining to reliability included atypical typing and/or tense errors produced by some listeners because these errors required some judgment on the part of the rater to determine if credit should be given for a key word. To reduce inconsistencies between raters, an "ambiguous response" word list was created, where raters indicated the listener, sentence list, target word, listener's response and whether or not credit was received. For example, if the listener typed "trimbs" for the key word "trims", credit was given and this response was added to the ambiguous response list. The list was updated by all three raters and then referenced whenever a listener’s response was questionable, improving both inter- and intra-rater reliability.

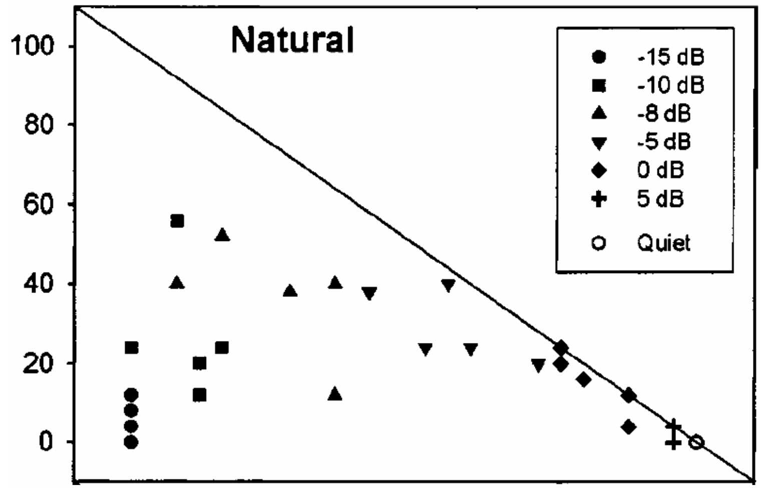
As a rough estimate of inter-rater reliability in scoring, four sentence lists were randomly selected for reliability analysis. Each list was examined to check for accuracy of scoring, and ambiguous responses were tabulated. Two lists contained a total of four ambiguous listener responses, one of which was judged to be scored incorrectly. Overall scoring agreement ranged from 98.7% to 100%, indicating good reliability between raters.

Effect of Stimuli

The psychometric functions for clear/slow and conversational/normal speech (obtained with nonsense sentences) were compared to Liu et al.’s (2004) corresponding psychometric functions for meaningful sentences, to determine any effect of stimuli. For meaningful sentences, Liu at al.’s normal hearing listeners received an average intelligibility benefit of 29 percentage points (when baseline intelligibility was 50%), for clear/slow speech presented by one talker. When compared to the current study’s average benefit (at the same baseline intelligibility) of roughly 19 percentage points for four talkers, the 10 point difference is most likely attributable to the dissimilar stimuli, as contextual cues in meaningful sentences can improve listener responses. For Liu et al.’s listeners, the advantage of semantic context is likely to have further enhanced the advantage provided by the acoustic properties that increase the intelligibility of clear speech. The increased clear speech benefit for meaningful stimuli appears to be greater for conditions that produce very low (<15%) baseline intelligibility.

Figure 6a shows that Liu et al.’s listeners obtained benefits up to nearly 60 points in this region, whereas the benefits in the current study (Figure 6b) were much smaller (0 -15 points). However, only one talker in the current study (T3) produced intelligibility scores that fell below 15% for the more difficult SNRs. Given the level of talker variability observed in the current study, it is possible that different results would have been achieved with the other three talkers (if smaller SNRs had been used to decrease their baseline intelligibility to <15%).

a) Liu et al. (2004)



b) Current study

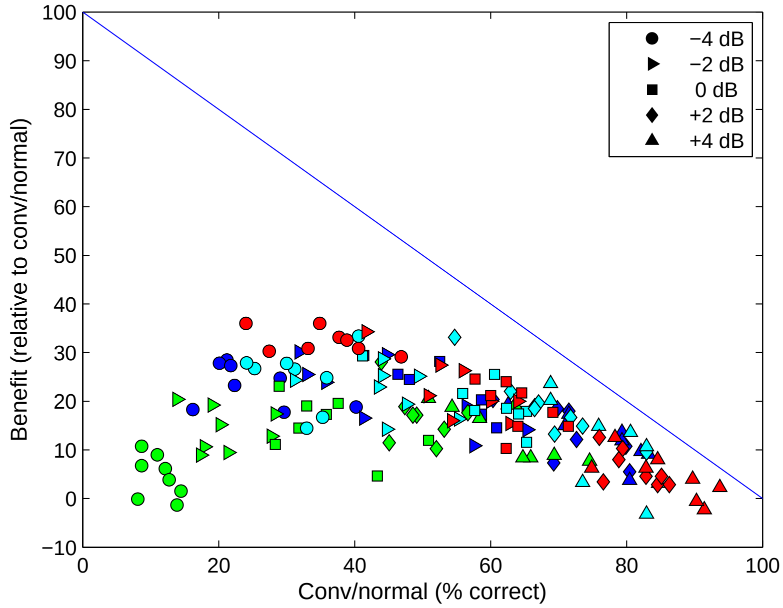


Figure 6. Intelligibility benefits of clear/slow speech by talker for Liu et al. (a) versus the current study (b), (T1 - blue, T3 - green, T4 - aqua, T5 - red) and SNR, relative to baseline performance, for all listeners. Diagonal line indicates maximum possible benefit for each baseline performance level.

Talker variability

Talker variability merits further consideration. Of the four talkers in this study, T3 was most different from the other talkers in at least two respects. Her overall intelligibility was the lowest of any talker in the study, and she was the only talker for whom clear/slow was not the most intelligible condition. If results for T3 were omitted from the analysis, different outcomes would be observed. Across talkers, there would be smaller differences in average intelligibility benefits for all conditions. With four talkers, conversational/slow was the most variable condition, with a difference of 28 percentage points between talkers with the smallest and largest intelligibility benefits. Without T3, this difference would reduce to 13 points. For clear/normal speech, the variability in benefits across talkers would decrease from 14 to 7 points. And, for the clear/slow condition, there would be only a 3 percentage point difference between intelligibility benefits for T1, T4 and T5 (as compared to a 7 point difference when T3 is included). Therefore, the variations in the size of benefits across talkers for all conditions would have been much smaller, with greater overall predictability, without T3.

A possible explanation for the variability between talkers involves training. Most talkers in previous clear speech studies were not professionals (Schum, 1996; among others) and were only instructed to carefully enunciate their words and/or speak as if they are talking to a person with a hearing loss or in a noisy environment. Minimal training and effort on behalf of the talker has been proclaimed as one of the major benefits of clear speech (Schum). Yet, to further counteract variability across talkers and avoid effects of a talker like T3, additional instruction and/or listener feedback may be required for more reliable intelligibility results. If additional training does reduce variability, more talker training may be warranted in all future studies of clear speech.

Regardless, talker variability was a significant factor in this study, which suggests that it may also have been a factor in previous clear speech studies. Among similar studies that presented nonsense sentences in noise to young normal hearing listeners, the number of talkers ranges from one (Helfer, 1997, among others) to five (Krause & Braida, 2002; Smiljanic & Bradlow, 2005). If any of these studies had used a single talker with similar intelligibility to T3, their findings would have been considerably different. If Krause and Braida had only used T3 in their research, for example, they would have found greater average intelligibility benefits with clear/normal (20 points) and conversational/slow conditions (33 points) than with clear/slow speech (13 points). In order to counteract any effects of talker variability, a much larger number of talkers should be used in future clear speech research. Only then can it be determined if T3 is an exception to the average talker, or if researchers should expect one out of every four talkers to exhibit the characteristics seen with T3.

Listener Variability

Using Figure 6b, listener variability in the current study can be estimated. If listener variability is small, any two data points from the same talker (i.e. the same color) which have comparable baseline intelligibility should have comparable benefit from clear/slow speech. While there are many examples where this is the case, there are at least as many examples where listener performance differed substantially. For example, when T5 (red) produced roughly 62% baseline conversational speech intelligibility, one listener obtained a benefit from clear/slow speech of less than 10 points while another listener had a benefit of more than 20 points.

Listener variability with the current study can also be compared to Liu et al.'s (2004) listeners (Figure 6a). Liu et al. noted that listener intelligibility gains resembled an inverted "U" shaped curve and maximum benefits were received in the middle of the psychometric function. This same phenomenon is visible in Figure 6b for the present study's listeners, with the largest intelligibility benefits obtained in the linear portion of the psychometric function, where conversational/normal speech intelligibility is between 20 and 65%. While talker variability appears to increase as baseline intelligibility decreases, listener variability is relatively constant (between 0 and roughly 10 points) across all baseline intelligibility levels. This suggests that listener variability is more "random," while talker variability is more likely associated with individual talker characteristics.

Other listener issues. One factor that may affect listener variability is the method of scoring used in this study. Response lists were scored on a key word basis, so listeners may still have received some benefit of context. Listeners were informed that the sentences were grammatically correct, containing true words. When listeners heard only a portion of a word or sentence (e. g. "His right cane could ard an edge"), they may have tried to "fill in" the missing word ("guard"). With the knowledge that the word must be a verb, many listeners may have guessed correctly. However, other listeners' guesses (e.g., "card") would be scored as completely incorrect, even when their guesses included some of the correct phonemes. To reduce variability between listeners, responses could be scored on a phoneme-by-phoneme basis to increase consistency of the results in these cases.

Another factor that could have affected listener variability was the effects of group dynamics on listener responses, an unexpected difficulty encountered in collecting the data. Most listeners completed the study as part of a group of four participants. During administration of the sentence lists, some indications of competitiveness in typing of responses were noted between participants. This "rush" to complete a sentence may have compromised typing accuracy and/or created poor listening habits. Additionally, there were participants who attempted to start typing before a sentence presentation was completed, possibly impairing other listeners' ability to hear the stimulus. These issues were addressed by reminding the listeners to take as much time as needed to type their responses. In addition, the listeners were instructed that the sentences had to be separated by a minimum interstimulus interval, so that hurrying would not translate to finishing the list or the session any more quickly. To enforce this notion, when a group appeared to be rushing, the delay between sentence presentations was increased by the tester, who manually controlled the rate of presentation of the sentences. For the participants who stated that they needed to start typing as soon as each sentence began in order to remember what to type, they were urged to type quietly in order to be considerate of the other participants in the room.

Several recommendations are offered to eliminate or reduce the effects of individual listener or group issues for future studies. First, more instructions and reminders regarding sentence presentation limitations and typing guidelines should be provided for each group session. As previously noted, there were some listeners who did not participate as part of a group. These listeners controlled the presentation of the sentences, rather than the tester, and all expressed a high level of comfort and satisfaction with this testing arrangement. There was no pressure to keep up or work ahead of the group and the listeners could type during the sentence without interfering with other participants' listening. Preferably, experiments should be conducted with individual listeners, or with multiple listeners in separate rooms, eliminating the problems observed with group sessions.

Future Work

Future research involving the current data will include more detailed examination of the slopes of the psychometric functions for each talker. Although it appears that slopes are fairly constant within talkers, there are some exceptions. For example, T5 appears to have one slope for conversational conditions, and a different slope for clear conditions. Statistical analysis of the slopes will help predict if this phenomenon holds at more difficult SNRs. And, comparing slopes between talkers will determine whether one talker (e.g. T4) has a flatter slope than the other talkers. This will help to establish if talker's slopes are truly different or if they represent different places on the psychometric function.

It is also possible that slopes may differ between populations. Therefore it is important to characterize the psychometric functions for a given population. Once these psychometric functions are established, comparisons can be conducted to determine if clear speech intelligibility levels for a particular population have the same slope as for young normal hearing listeners. For example, it cannot be assumed that children with learning disabilities receive the same intelligibility benefits with clear speech as young adults, even when baseline (conversational/normal) levels are equated. The slope of the psychometric function for children may be flatter or steeper than for young adults, so performance at a given presentation level will vary.

For these reasons, future clear speech research should, ideally, use full psychometric functions for a given population. Studies should be conducted with multiple talkers and similar stimuli to minimize the effects of talker and stimuli. However, conducting research with a large number of talkers to reduce variability is not always a realistic process due to limited time and financial resources. A possible solution is for researchers to share their talker databases, as well as stimuli and type of degradation used (i.e., speech-shaped noise, filtering). If researchers use identical stimuli in the same environments, comparable psychometric functions can be established for multiple populations and differences in intelligibility benefits can be attributed to differences in those populations. By controlling their methods, researchers can make valid comparisons across studies.

Furthermore, if researchers conduct pilot experiments to determine the stimulus condition (e.g., SNR) producing approximately 50% intelligibility for a given population, a practical range of signal-to-noise ratios can be established. This will reduce floor and ceiling effects and concentrate resources on the linear portion of the psychometric function where listeners receive maximum and consistent benefits. This would reduce or possibly eliminate the need for a full psychometric function while still providing baseline performance levels for comparisons.

Clinical Implications

Clinicians should keep in mind that intelligibility benefits with clear speech are dependent on the listening environment and the talker. This point needs to be emphasized when educating family members and clients about the benefits of clear speech. Enhanced training and practice, preferably in different listening situations, should be conducted. Clients should be informed that there are some noise situations where acoustic cues of clear speech will not benefit the listener. And, that the benefits of clear speech are dependent on the talker's ability to speak clearly as well as the listener's hearing acuity. These variability factors should be fully explained to clients and their families to minimize frustration in the event that clear speech does not appear to improve communication.

Despite the variability issues identified in this thesis, the many clinical implications of clear speech as an effective and practical mode of communication remain. By using clear speech with all populations, clinicians can be assured their communicative message is enhanced and that clients will benefit. In a clinical setting, clients are frequently presented with information in a technical format or using medical jargon. Presenting that information in a clear, more intelligible form may enable the client to focus more on the communicative meaning by lightening their “cognitive load.” Additionally, clear speech requires minimal effort but offers immeasurable rewards through improved communication and relationships with clients.

References

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America, 112*, 272-284.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with disabilities: Sentence perception in noise. *Journal of Speech, Language and Hearing Research 46,* 80-97.

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 112,* 259-271.

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

Gagne, J., Rochette, A., & Charest, M. (2002). Auditory, visual and audiovisual clear speech. *Speech Communication, 37*, 213-230.

Goy, H., Pichora-Fuller, K., van Lieshout, P., Singh, G., & Schneider, B. (2007, August). *Effect of within- and between-talker variability on word identification in noise by younger and older adults.* Paper accepted for presentation at the meeting of Interspeech 2007, Antwerp, Belgium.

Helfer, K. S. (1997). Auditory and auditory-visual perception of clear and conversational speech. *Journal of Speech, Language and Hearing Research, 40*, 432-443.

Helfer, K. S. (1998). Auditory and auditory-visual recognition of clear and conversational speech by older adults. *Journal of the American Academy of Audiology, 9*, 234-242.

Krause, J. C., & Braida, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America, 112,* 2165-2172.

Krause, J. C., & Braida, L. D. (2003). Effects of listening environment on intelligibility of clear speech at normal speaking rates. *Iranian Audiology, 2*, 39-47.

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America, 115,* 362-378.

Liu, S., Del Rio, E., Bradlow, A. R., & Zeng, F. (2004). Clear speech perception in acoustic and electric hearing. *Journal of the Acoustical Society of America, 116,* 2374-2383.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusion among some English consonants. *The Journal of the Acoustical Society of America, 2,* 338-352.

Nillson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America, 95,* 1085-1099.

Panagiotopoulos, A. P. (2005). *The benefits of clear speech at normal rates for older listeners with normal hearing*. Unpublished master's thesis, University of South Florida, Tampa.

Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America, 95,* 1581-1592.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research, 28,* 96-103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research, 29,* 434-446.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing III: An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research, 32,* 600-603.

Schum, D. J. (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology, 7*, 212-218.

Schum, D. J. (1997). Beyond hearing aids: Clear speech training as an intervention strategy. *The Hearing Journal, 50*(10), 36-40.

Smiljanic, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America, 118,* 1677-1688.

Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech and Hearing Research, 39*, 494-509.

Wilson, R. H., & Strouse, A. L. (1999). Auditory measures with speech signals. In F. E. Musiek & W. F. Rintlemann (Eds.), *Contemporary perspectives in hearing assessment* (pp. 21-66). Needham Heights, MA: Allyn & Bacon.

Appendices

Appendix A: Participant Demographics

Participant	Gender	Age	Participant	Gender	Age
L1	F	23	L21	F	23
L2	F	19	L22	F	25
L3	F	38	L23	F	22
L4	F	21	L24	F	29
L5	F	22	L25	F	23
L6	M	21	L26	M	37
L7	F	22	L27	F	25
L8	M	24	L28	M	21
L9	F	20	L29	F	18
L10	F	21	L30	F	20
L11	F	27	L31	M	28
L12	F	26	L32	F	37
L13	M	18	L33	F	18
L14	F	20	L34	F	19
L15	F	21	L35	F	21
L16	F	33	L36	M	18
L17	M	19	L37	F	19
L18	F	33	L38	M	19
L19	F	21	L39	F	34
L20	F	19	L40	F	24

Appendix B: Key Word Scores by Talker

Table B1. T1

		% Correct			
SNR	Listener	Conv/normal	Clear/normal	Conv/slow	Clear/slow
-4	L1	29.1	34.1	49.7	53.8
-4	L2	21.2	43.9	56.1	49.7
-4	L3	22.3	25.4	40.4	45.6
-4	L4	20.1	30.6	45.6	48.0
-4	L33	21.8	32.4	49.1	49.1
-4	L34	29.6	32.4	51.5	47.4
-4	L35	40.2	38.7	57.9	59.1
-4	L36	16.2	25.4	40.9	34.5
-2	L5	33.0	34.1	52.6	58.5
-2	L6	41.3	49.7	64.3	57.9
-2	L7	35.8	49.7	73.1	59.6
-2	L8	31.8	57.8	71.3	62.0
-2	L37	56.4	44.5	62.6	75.4
-2	L38	57.5	56.1	61.4	68.4
-2	L39	44.7	49.1	70.8	74.3
-2	L40	65.4	52.6	64.9	79.5
0	L9	58.7	74.6	81.9	75.4
0	L10	48.0	60.1	80.7	72.5
0	L11	58.7	63.0	77.8	76.0
0	L12	52.5	63.6	80.1	80.7
0	L21	58.7	61.8	71.9	78.9
0	L22	46.4	53.8	64.9	71.9
0	L23	60.9	71.7	77.8	75.4
0	L24	41.3	64.7	74.3	70.8
2	L13	72.6	76.9	85.4	84.8
2	L14	79.9	79.8	86.5	90.6
2	L15	60.3	70.5	80.1	80.7
2	L16	66.5	71.1	87.1	85.4
2	L25	69.3	68.2	77.8	76.6
2	L26	80.4	77.5	78.9	86.0
2	L27	71.5	81.5	88.9	89.5
2	L28	69.8	61.3	79.5	88.3
4	L17	62.6	86.1	90.1	81.9
4	L18	71.5	72.8	80.7	88.9
4	L19	82.1	82.1	89.5	91.8
4	L20	83.2	82.7	92.4	92.4
4	L29	80.4	85.5	85.4	84.2
4	L30	70.9	75.7	84.8	86.0
4	L31	79.3	72.3	86.0	93.0
4	L32	79.3	84.4	90.6	91.2

Appendix B (continued)

Table B2. T3

		% Correct			
SNR	Listener	Conv/normal	Clear/normal	Conv/slow	Clear/slow
-4	L9	12.1	20.8	33.1	18.3
-4	L10	13.9	32.9	42.3	12.6
-4	L11	14.5	34.1	41.7	16.0
-4	L12	11.0	43.9	54.3	20.0
-4	L37	8.7	43.4	50.3	15.4
-4	L38	12.7	24.3	34.9	16.6
-4	L39	8.1	37.6	46.3	8.0
-4	L40	8.7	34.1	47.4	19.4
-2	L13	20.2	54.3	58.9	35.4
-2	L14	28.3	43.4	58.9	45.7
-2	L15	17.9	41.6	66.9	28.6
-2	L16	13.9	51.4	57.7	34.3
-2	L21	21.4	35.8	55.4	30.9
-2	L22	19.1	45.1	55.4	38.3
-2	L23	27.7	54.9	70.3	40.6
-2	L24	17.3	29.5	59.4	26.3
0	L17	28.3	52.0	76.0	39.4
0	L18	43.4	58.4	73.7	48.0
0	L19	37.6	47.4	70.9	57.1
0	L20	50.9	48.6	66.3	62.9
0	L25	28.9	56.6	73.1	52.0
0	L26	31.8	53.8	65.7	46.3
0	L27	32.9	54.3	69.1	52.0
0	L28	35.8	56.6	76.0	53.1
2	L1	52.0	74.0	85.7	62.3
2	L2	49.1	69.9	75.4	66.3
2	L3	43.9	70.5	86.9	72.0
2	L4	53.2	71.1	84.6	67.4
2	L29	47.4	63.6	82.9	66.3
2	L30	45.1	72.8	83.4	56.6
2	L31	48.6	80.9	85.1	65.7
2	L32	56.6	60.7	77.7	74.3
4	L5	65.9	79.2	84.6	74.3
4	L6	69.4	82.1	94.9	78.3
4	L7	58.4	75.7	87.4	74.9
4	L8	54.3	73.4	89.1	73.1
4	L33	50.9	74.6	82.9	71.4
4	L34	64.7	78.6	89.7	73.1
4	L35	74.6	80.9	93.1	82.3
4	L36	63.6	85.5	90.3	82.9

Appendix B (continued)

Table B3. T4

		% Correct			
SNR	Listener	Conv/normal	Clear/normal	Conv/slow	Clear/slow
-4	L13	31.2	40.0	40.5	57.8
-4	L14	40.6	28.8	31.8	74.0
-4	L15	25.3	34.1	39.3	52.0
-4	L16	35.3	29.4	37.0	52.0
-4	L21	32.9	29.4	22.0	47.4
-4	L22	30.0	38.8	42.2	57.8
-4	L23	35.9	52.4	52.0	60.7
-4	L24	24.1	30.0	30.1	52.0
-2	L17	31.2	54.7	57.2	55.5
-2	L18	47.6	56.5	54.9	67.1
-2	L19	44.1	45.9	39.3	72.8
-2	L20	55.3	53.5	53.8	71.7
-2	L25	44.7	66.5	60.1	59.0
-2	L26	43.5	52.4	62.4	66.5
-2	L27	49.4	58.8	55.5	74.6
-2	L28	44.1	75.3	71.7	69.4
0	L1	65.3	68.8	74.0	83.2
0	L2	64.1	51.2	52.6	81.5
0	L3	60.6	68.8	65.9	86.1
0	L4	65.3	68.2	67.6	76.9
0	L29	55.9	66.5	66.5	77.5
0	L30	41.2	62.9	67.6	70.5
0	L31	62.4	74.1	72.8	80.9
0	L32	57.6	58.2	54.3	75.7
2	L5	69.4	76.5	77.5	82.7
2	L6	71.8	90.6	83.2	88.4
2	L7	67.1	70.0	61.8	86.7
2	L8	66.5	68.8	72.3	85.0
2	L33	54.7	74.1	73.4	87.9
2	L34	73.5	78.8	78.0	88.4
2	L35	82.9	85.3	74.6	92.5
2	L36	62.9	82.4	76.9	85.0
4	L9	68.8	77.1	71.7	92.5
4	L10	73.5	83.5	82.1	76.9
4	L11	80.6	82.9	83.8	94.2
4	L12	82.9	85.3	83.2	93.6
4	L37	84.7	83.5	80.3	87.9
4	L38	68.8	75.3	71.1	89.0
4	L39	75.9	85.3	81.5	90.8
4	L40	82.9	87.1	85.0	79.8

Appendix B (continued)

Table B4. T5

		% Correct			
SNR	Listener	Conv/normal	Clear/normal	Conv/slow	Clear/slow
-4	L17	27.4	59.4	42.3	57.7
-4	L18	33.1	54.9	36.0	64.0
-4	L19	40.6	64.0	40.0	71.4
-4	L20	46.9	56.0	49.1	76.0
-4	L29	37.7	63.4	40.6	70.9
-4	L30	24.0	50.9	30.3	60.0
-4	L31	34.9	54.9	38.9	70.9
-4	L32	38.9	61.7	56.6	71.4
-2	L1	62.9	73.1	60.0	78.3
-2	L2	54.3	80.0	72.6	70.3
-2	L3	52.6	70.3	46.3	80.0
-2	L4	60.0	72.0	62.3	80.0
-2	L33	41.7	66.9	56.6	76.0
-2	L34	56.0	70.9	58.3	82.3
-2	L35	64.0	77.7	70.9	84.0
-2	L36	50.9	60.6	56.6	72.0
0	L5	64.0	64.6	62.3	78.9
0	L6	64.6	77.7	73.1	86.3
0	L7	60.0	77.7	77.1	81.1
0	L8	69.1	84.0	81.1	86.9
0	L37	62.3	78.9	64.6	86.3
0	L38	62.3	88.6	77.7	72.6
0	L39	57.7	87.4	73.7	82.3
0	L40	71.4	82.3	78.3	86.3
2	L9	84.6	84.6	84.6	87.4
2	L10	76.6	84.6	85.1	80.0
2	L11	79.4	86.9	88.6	89.7
2	L12	86.3	91.4	90.3	89.1
2	L21	85.1	87.4	82.9	89.7
2	L22	78.9	77.1	61.1	86.9
2	L23	82.9	85.7	84.0	87.4
2	L24	76.0	91.4	84.0	88.6
4	L13	82.9	89.1	86.3	89.1
4	L14	89.7	93.1	90.9	93.7
4	L15	74.9	82.9	85.7	81.1
4	L16	78.3	85.1	86.9	90.9
4	L25	84.6	86.3	84.0	92.6
4	L26	91.4	92.6	90.3	89.1
4	L27	93.7	93.7	96.0	96.0
4	L28	90.3	86.3	85.7	89.7

Appendix C: Pairwise Comparisons

Table C1. Pairwise Comparisons by Condition

Dependent Variable: RAU
Tukey HSD

(I) Condition	(J) Condition	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Clear/normal	Clear/slow	-.062312*	.0081131	.000	-.083218	-.041407
	Conv/normal	.126107*	.0081131	.000	.105201	.147012
	Conv/slow	-.046514	.0081131	.000	-.067420	-.025609
Clear/slow	Clear/normal	.062312*	.0081131	.000	.041407	.083218
	Conv/normal	.188419*	.0081131	.000	.167514	.209325
	Conv/slow	.015798	.0081131	.210	-.005107	.036704
Conv/normal	Clear/normal	-.126107*	.0081131	.000	-.147012	-.105201
	Clear/slow	-.188419*	.0081131	.000	-.209325	-.167514
	Conv.slow	-.172621*	.0081131	.000	-.193527	-.151716
Conv/slow	Clear/normal	.046514*	.0081131	.000	.025609	.067420
	Clear/slow	-.015798	.0081131	.210	-.036704	.005107
	Conv/normal	.172621*	.0081131	.000	.151716	.193527

Based on observed means.
*The mean difference is significant at the 0.05 level.

Appendix C (continued)

Table C2. Pairwise Comparisons by Talker

Dependent Variable: RAU
Tukey HSD

(I) Talker	(J) Talker	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
T1	T3	.133969*	.0081131	.000	.113063	.154874
	T4	.002895	.0081131	.984	-.018010	.023801
	T5	-.092616*	.0081131	.000	-.113522	-.071711
T3	T1	-.133969*	.0081131	.000	-.154874	-.113063
	T4	-.131074*	.0081131	.000	-.151979	-.110168
	T5	-.226585*	.0081131	.000	-.247490	-.205679
T4	T1	-.002895	.0081131	.984	-.023801	.018010
	T3	.131074*	.0081131	.000	.110168	.151979
	T5	-.095511*	.0081131	.000	-.116417	-.074606
T5	T1	.092616*	.0081131	.000	.071711	.113522
	T3	.226585*	.0081131	.000	.205679	.247490
	T4	.095511*	.0081131	.000	.074606	.116417

Based on observed means.
*The mean difference is significant at the 0.05 level.