The Dissertation Committee for Rachael Celia Gilbert Certifies that this is the approved version of the following dissertation

Environment- and listener-oriented speaking style adaptations across the lifespan

Committee:

____

Rajka Smiljanic, Supervisor

____

Scott Myers

____

Harvey Sussman

____

Chang Liu

____

Bharath Chandrasekaran

Environment- and listener-oriented speaking style adaptations across the lifespan

by

Rachael Celia Gilbert, B.A.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2014

Acknowledgements

The author would like to acknowledge the UTSoundLab undergraduate research assistants for their contributions to this project.

Environment- and listener-oriented speaking style adaptations across the lifespan

Rachael Celia Gilbert, Ph.D.

The University of Texas at Austin, 2014

Supervisor: Rajka Smiljanic

This dissertation examines how age affects the ability to produce intelligibility-enhancing speaking style adaptations in response to environment-related difficulties (noise-adapted speech) and in response to listeners' perceptual difficulties (clear speech). Materials consisted of conversational and clear speech sentences produced in quiet and in response to noise by children (11-13 years), young adults (18-29 years), and older adults (60-84 years). Acoustic measures of global, segmental, and voice characteristics were obtained. Young adult listeners participated in word-recognition-in-noise and perceived age tasks. The study also examined relative talker intelligibility as well as the relationship between the acoustic measurements and intelligibility results.

Several age-related differences in speaking style adaptation strategies were found. Children increased mean F0 and F1 more than adults in response to noise, and exhibited greater changes to voice quality when producing clear speech (increased HNR, decreased shimmer). Older adults lengthened pause duration more in clear speech compared to younger talkers. Word recognition in noise results revealed no age-related differences in the intelligibility of conversational speech. Noise-adapted and clear speech modifications increased intelligibility for all talker groups. However, the acoustic changes implemented by children when producing noise-adapted and clear speech were less efficient in enhancing intelligibility compared to the young adult talkers. Children were also less intelligible than older adults for speech produced in quiet. Results confirmed that the talkers formed 3 perceptually-distinct age groups. Correlation analyses revealed that relative talker intelligibility was consistent for conversational and clear speech in quiet. However, relative talker intelligibility was found to be more variable with the inclusion of additional speaking style adaptations. 1-3 kHz energy, speaking rate, vowel and pause durations all emerged as significant acoustic-phonetic predictors of intelligibility.

This is the first study to investigate how clear speech and noise-adapted speech benefits interact with each other across multiple talker groups. The findings enhance our understanding of intelligibility variation across the lifespan and have implications for a number of applied realms, from audiologic rehabilitation to speech synthesis.

Table of Contents

List of Tables

List of Figures

Introduction

Although a talker's intrinsic speech clarity is a large determinant of intelligibility, talkers are capable of enhancing their intelligibility via speech modifications elicited in response to adverse communicative situations. While such intelligibility-enhancing speaking style adaptations have been well researched for healthy young adult talkers, our knowledge of how they develop across the lifespan is sparse. This series of experiments seeks to evaluate the production and perception of speaking style adaptations as produced by children, young adults, and older adults. Understanding the effects of age and communicative intent on speech production (and how this variation shapes speech intelligibility) is a pressing issue given the prevalence of communicative difficulties in daily interactions (talking to hearing-impaired listeners, communicating in noisy classrooms and health care clinics, etc.), and knowledge of which acoustic-phonetic changes improve speech intelligibility remains limited.

ENHANCING INTELLIGIBILITY THROUGH SPEAKING STYLE ADAPTATIONS

Research on speaking style adaptations has been generally split into two main fields: environment-oriented speaking style adaptations (noise-adapted speech) and listener-oriented speaking style adaptations (clear speech). Findings from both fields yield numerous insights into how within-talker speech intelligibility may be enhanced. However, the findings are extremely varied and few studies have compared the two types of speech adaptations side by side. Little is known about the combined effects of noise-adapted speech and clear speech. Given that speaking style adaptations are often simultaneously elicited (e.g. talking to a hearing-impaired individual in a noisy environment), it is important to examine these speaking style adaptations in conjunction.

Noise-adapted speech (Lombard, 1911) is an automatic response to noise, resulting in speech that is more resistant to its masker. Compared to speech produced in quiet, noise-adapted speech typically exhibits a decrease in speech rate, an increase in vocal levels, longer vowel duration, a higher average F0, a higher peak F0, as well as increased energy at higher frequencies (Lombard, 1911; Pittman and Wiley, 2001; Summers et al., 1988; Junqua, 1993; Lane and Tranel, 1971; Navarro, 1996; Cooke and Lu, 2010). Perceptually, speech produced in response to noise and then mixed with the noise for subsequent listening tests is significantly more intelligible and better recognized from memory than speech produced in quiet and then mixed with noise (Pittman and Wiley, 2001; Lu and Cooke, 2008; Dreher and O'Neill, 1957; Summers et al., 1988; Gilbert et al., 2014). Similar involuntary responses, e.g., rise in call amplitude, aimed at increasing the signal-to-noise ratio, thus facilitating signal transmission have been found in birds and mammals, including bats (Hage et al., 2013; Brumm and Todt, 2002; Brumm et al., 2004).

Clear speech, on the other hand, is a talker's (perhaps more intentional) adaptation to a speech perception difficulty on the part of the listener (e.g. low proficiency or hearing impairment). This term has often been used as an umbrella term, but in line with the definition most commonly used in clear speech research, the scope of this term will be restricted to read laboratory speech elicited by instructions given to talkers rather than to the spontaneous speech occurring in a more natural setting (for a discussion of terminology, see Smiljanic and Bradlow, 2009). Relative to conversational speech, clear speech is most often characterized by a decrease in speaking rate (longer segments as well as longer and more frequent pauses), an increase in vocal levels, a wider F0 range, more salient stop releases, an expanded vowel space, greater obstruent RMS energy, and increased energy at higher frequencies (Smiljanic and Bradlow, 2005; Picheny et al., 1986; Krause and Braida, 2004; Bradlow et al., 2003; Liu et al., 2004, Ferguson and Kewley-Port, 2002). Perceptually, it is well established that clear speech improves word recognition and recognition memory relative to conversational speech (Smiljanic and Bradlow, 2005; Van Engen et al., 2012; Gilbert et al., 2014).

While the two speaking style adaptations share a number of features, the specific articulatory-acoustic modifications vary considerably speaker-to-speaker. Tartter and colleagues (1993), for instance, found one talker increased F0 in response to noise while a second talker showed the opposite pattern. In another study, two talkers produced a higher F0 in clear speech than in conversational speech, two produced a lower F0, and one showed no change (Krause and Braida, 2004). This variability extends to a number of other features, e.g. vowel space expansion (Ferguson and Kewley-Port, 2007; Ferguson et al., 2010).

Similarly, the perceptual benefits of quiet-to-noise and conversational-to-clear speech modifications are highly variable. Some studies have found noise-adapted speech to improve intelligibility by up to 35% (Pittman and Wiley, 2001) while others have found only a minor 2% improvement (Goy et al., 2007), despite similarities in methodology (native speakers, normal-hearing native listeners, meaningful sentences). For clear speech, Liu et al. (2004) found an average improvement of 30%, while Ferguson (2004) only found an average increase of 8.5% (which, when examined across all 41 talkers, ranged from -12% to 33%)[1]. Ferguson and Kewley-Port (2007), when examining 6 talkers who produced a larger clear speech benefit against 6 talkers who produced no clear speech benefit, found that the variability in the intelligibility gain was

[1] Both studies cited here used native talkers and normal-hearing native listeners, although Liu et al. (2004) used sentences while Ferguson (2004) used vowels.

mirrored by an equally large amount of variability in the acoustic-phonetic features the talkers implemented to produce clear speech.

The direct mapping between specific acoustic changes and increased intelligibility is still rather elusive; while research has shown both noise-adapted and clear speech adaptations improve speech intelligibility, the exact acoustic-phonetic modifications that are responsible have not been reliably identified (Lu and Cooke, 2009; Smiljanic and Bradlow, 2009). Additionally, very few studies have examined the impact of these acoustic changes on speech perception tasks aside from word recognition (aside from Van Engen et al., 2012 and Gilbert et al., 2014 who both found speaking style adaptations to improve sentence recognition memory).

Although noise-adapted speech and clear speech often share many acoustic-articulatory features (e.g. slower speaking rate, an increase in F0), they differ both acoustically and perceptually (Cooke, King, Garnier, Aubanel, 2014; Godoy et al., 2014; Gilbert et al., 2014). For instance, vowel space expansion has been shown in the clear speech literature much more reliably than in the noise-adapted speech literature (Godoy et al., 2014). Findings also indicate the two styles may differ in terms of F0 range and the long-term average spectrum (Gilbert et al., 2014; Cooke, King, Garnier, Aubanel, 2014). However, many analyses are not equally distributed across the two fields of research; according to Pichora-Fuller et al. (2010), measures of vowel space and F0 variability were proportionally more common in clear speech studies than noise-adapted speech studies, whereas measures of intensity and spectra were more common in noise-adapted speech studies than in clear speech studies.

One of the goals of this dissertation is to address this gap in our knowledge by providing a direct and comprehensive examination of noise-adapted speech and clear speech, both separately and in conjunction. In what manner do the acoustic-phonetic correlates of noise-adapted clear speech resemble those of noise-adapted speech and clear speech separately? Is noise-adapted clear speech as intelligible as the sum of its parts, or is there an interaction between listener-oriented and environment-oriented speaking style adaptations? How are clear speech and noise-adapted speech different? Can some of the acoustic features reflect more intentional vs. automatic changes further delineating the difference between the two types of adaptations?

SPEAKING STYLE ADAPTATIONS IN DIFFERENT TALKER POPULATIONS

Work on these speaking style adaptations has been largely limited to normal-hearing, adult speakers of English (e.g. Pittman and Wiley, 2001; Ferguson, 2004). However, research on speaking style adaptations in different talker populations has established striking variability. For example, late learners of English produce a smaller clear speech benefit compared to early learners and native speakers of English (Rogers et al., 2010; Smiljanic and Bradlow, 2007, 2011). Female talkers provide a larger clear speech intelligibility gain than male talkers (Ferguson, 2004). Even within a talker population, the articulatory-acoustic features of these speaking style adaptations vary considerably speaker-to-speaker (Tartter et al., 1993; Krause and Braida, 2004; Kang and Guion, 2008; Ferguson and Kewley-Port, 2007; Ferguson et al., 2010).

While some work has examined non-native clear speech production and gender-related differences, age-related changes in the production of intelligibility-enhancing speaking styles have been largely ignored. Given that children, young adults, and older adults significantly differ in their speech production systems (e.g. vocal tract length, speech-motor control), speaking style adaptations are likely to differ across lifespan (Lee et al., 1999; Benjamin, 1982). There is little work on the baseline intelligibility of these talker populations, and even less is known about the extent to which children and older adults are able to enhance their intelligibility. The primary goal of this dissertation is to examine noise-adapted and clear speech across talkers of different ages. Are children and older adult talkers able to increase their intelligibility to the same degree as young adult speakers? Do children and older adult talkers implement the same acoustic-articulatory changes common in young adult noise-adapted and clear speech?

Children

Although there is little research on speaking style adaptations in children, it is well documented that children produce speech in a different manner than adults. Research has shown that many aspects of language are slow to develop in children, e.g. even by age 13, children do not yet show an adult-like level of co-articulation (Gerosa et al., 2006). Children also exhibit several physiological differences in their vocal systems, both laryngeal and respiratory in nature (Tang and Stathopoulos, 1995; Stathopoulos and Sapienza, 1993). These effects culminate in speech production differences such as longer segmental durations, higher and more variable F0 and vowel formants, and lower harmonics-to-noise ratios (an index of the degree of hoarseness, quantifying the amount of additive noise in the voice signal) (Ferrand, 2000; Lee et al., 1999). Although older children (13 year-olds) can produce as intelligible speech as adults, they exhibit greater variance in intelligibility rates (Markham and Hazan, 2004).

Only a few studies have examined whether children are able to produce modifications in response to an adverse communicative situation and in what way they achieve these changes. Children are aware of the adverse communicative situations and are able to adjust their speech in response to environment and to the listener at a very young age. Weeks (1971) found that a 1:7 year-old would increase pauses in response to the listener's perceptual difficulty. Nicoladis and Genesee (1997) found that bilingual 2 year-olds were able to shift their language depending on the language of their listener. Brinton et al. (1986) found a 3 year-old would repeat, revise, and supplement information in response to listener difficulty. Shatz and Gelman (1973) found 4 year-old children used shorter and less complex utterances when speaking to younger children than when speaking to their peers and adults. Andersen and colleagues found 4 year-old children were also able to adjust their speech to imitate various social roles, e.g. “doctor”, “teacher”, “mommy”, “daddy” (Andersen, 1996; Andersen, Brizuela, DePuy, & Gonnerman, 1999).

However, children’s command of speaking styles is slow to mature; at the lexical level, children do not produce adult-like nuances of polite speech until late grade school (Pedlow, Sanson, and Wales 2001; 2004). Both temporally and spectrally, the characteristics of children’s speech also show increased within- and across-talker variability up to 15 years of age (Lee et al., 1999).

In terms of intelligibility-enhancing speech adaptations to the environment and to the listener (noise-adapted and clear speech), research has focused on children under 6 years of age. Three to four year-old children’s noise-adapted speech is characterized by an increase in intensity similar to adults (Siegel, Pick, Olsen, and Sawin, 1976; Amazi and Garber, 1982). The quiet-to-noise-adapted change in vocal intensity has been shown to be larger in 5 year-old children than in adults (Garber, Speidel, & Siegel, 1980; Garber, Speidel, Siegel, Miller, & Glass, 1980). It is unknown what other acoustic changes accompany this increase in vocal intensity, and how this changes over time.

Only three studies have examined children’s productions of clear speech (Redford and Gildersleeve-Neumann, 2009; Syrett and Kawahara, 2013; Pettinato and Hazan, 2013). Redford and Gildersleeve-Neumann (2009) found that children under 5 were capable of producing distinct speaking styles, although they were unable to produce adult-like conversational-to-clear speech adaptations—for instance, 4 and 5 year-olds' clear speech involved shorter vowels and lower F0. Syrett and Kawahara (2013) found that 3 to 5 year-old children were able to produce a distinct form of clear speech with expanded F0 range as well as longer, louder, and more dispersed vowels. Pettinato and Hazan (2013) compared clear speech produced by 9 to 10 year-olds, 13 to 14 year-olds, and young adults. They found that, while both children and teens slowed their speech rate, neither group produced the vowel hyperarticulation found in young adult clear speech. It appears that many features of adult-like speaking style adaptations continue to develop in adolescence. Given that only one of these speaking adaptation studies has examined children over age 6, it is important to provide additional findings on older children who exhibit more developed articulatory control and planning. Furthermore, it is crucial to examine whether these listener-oriented changes produced by children result in enhanced intelligibility, which none of the above studies addressed.

Older Adults

It is well established that older adults have greater speech processing difficulties arising from a combination of cognitive and peripheral-auditory declines (Pichora-Fuller et al., 1995; Schneider et al., 2002; Gordon-Salant and Fitzgibbons, 1997). Physiologically, aging affects the vocal system via the degeneration of laryngeal mechanisms (atrophy of musculature, degeneration of nerve fibers, ossification of cartilages, etc.) and reductions in the respiratory processes underlying production (reduction in breath support, pulmonary recoil pressures, lung cavity size, muscle force, lung elasticity, etc.) (Huber, 2008; Awan, 2006).

The speech production system is affected by both these types of age-related differences. Research has shown that older adults tend to increase the use of filler words, decrease speaking rate, and show increased F0 variability. Age-related changed also often include changes in voice quality, with decreased harmonics-to-noise ratio (an index of hoarseness, quantifying the ratio of acoustic periodicity to noise components), increased jitter (an index of roughness, quantifying the percentage of cycle-to-cycle irregularity in F0), and increased shimmer (quantifying the percentage of cycle-to-cycle irregularity in amplitude) (Au et al., 1995; Spieler et al., 2004; Halle and Myerson, 1996; Gorham-Rowan and Laures-Gore, 2006; Higgins and Saxman, 1991; Benjamin, 1982; Yumoto et al., 1982; Ferrand, 2002). However, findings are not often in accord, e.g. while many studies have found jitter to increase with age (Linville and Fisher, 1985; Wilcox and Horii, 1980), several others have not (Ferrand, 2002; Linville, 1987).

In terms of intelligibility-enhancing speaking styles, there are no studies that have examined noise-adapted speech as produced by older adults. Few studies have examined the production of clear speech in older adult talkers, and only 2 studies have examined the extent to which this improves intelligibility (Kang and Guion, 2008; Schum, 1996; Smiljanic, 2013). Kang and Guion (2008) found that Korean-speaking younger and older adults produce different enhancement patterns in clear speech; older adult talkers enhanced VOT differences for the aspirated-lenis stop contrast, while younger talkers primarily enhanced F0 (reflecting an ongoing sound change). While both Schum (1996) and Smiljanic (2013) found that older adults produced listener-oriented modifications which increased intelligibility, it is debated to what extent the overall intelligibility and clear speech intelligibility gain resemble that of young adults; Schum found both talker groups to produce comparable benefits while Smiljanic found the older adult clear speech benefit to be smaller than that of young adults'. This is perhaps due in part to the difference in stimuli; Schum used simple meaningful sentences while Smiljanic used more taxing, anomalous sentences. Furthermore, Schum did not report overall intelligibility levels, so it is difficult to assess the clear speech gain across the 2 talker groups. Given that both studies used only a small number of talkers, and that very few acoustic measurements were reported, it is important to explore older adult clear speech productions further. It also remains to be seen whether this pattern extends to noise-adapted speech, what are the acoustic-articulatory changes characterizing these speaking style adaptations, and how these influence listener perception compared to speech produced by children and young adults.

THE EFFECT OF SPEAKING STYLE ADAPTATIONS ON RELATIVE INTELLIGIBILITY

It is well established that in addition to listener- and environment-related factors, talker-related factors play an enormous role in speech understanding. According to Bradlow et al. (1996), "a substantial portion of variability in normal speech intelligibility is traceable to specific acoustic-phonetic characteristics of the talker" (p. 255). That talker-specific features determine speech intelligibility has been widely demonstrated (Hazan and Markham, 2004; Green et al., 2007; Bent et al., 2009; van Dommelen and Hazan, 2012). Even controlling for environment- and listener-related factors, findings have shown the intelligibility of 41 native adult speakers can range from 25% to 83% in conversational speech, and from 29% to 94% in clear speech (Ferguson, 2004).

Despite the large talker-related variability in intelligibility, the relative intelligibility of different talkers is remarkably consistent across different listener populations and environments. This suggests that a talker's speech clarity is an inherent quality independent of environment or listener background. That is to say, out of a group of talkers, the ranking of most and least intelligible talkers is consistent for different listener populations, such as adult and children listeners (Hazan and Markham, 2004), normal-hearing and cochlear implant listeners (Green et al., 2007), or native and nonnative listeners (van Dommelen and Hazan, 2012). This notion of relative intrinsic talker clarity has also been shown to hold across different types of communicative environments, e.g. when speech is vocoded vs. masked by multi-talker babble (Bent et al., 2009).

The extent to which this relative intelligibility of individual talkers holds across different speaking style adaptations is not well known. Is the most intelligible talker in quiet also the most intelligible talker in noise? Individual talkers vary largely in the extent to which they are able to enhance their intelligibility via speaking style adaptations, thus suggesting that relative talker intelligibility may not hold across different speaking style conditions (Gagne et al., 2002; Bradlow et al., 2003). Ferguson (2004), on the other hand, found that talker intelligibility in conversational and clear speech was highly correlated for vowels produced by adult talkers ($r=0.74$, $p<0.001$). It remains to be seen if relative talker intelligibility across speaking styles is consistent for sentence-level materials across a more diverse group of talkers (children, young adults, and older adults) producing 2 different lines of intelligibility-enhancing speaking style adaptations (noise-adapted and clear speech). This is an additional goal of the dissertation.

THE RELATIONSHIP BETWEEN PRODUCTION AND PERCEPTION

Despite a large body of work identifying the specific acoustic-phonetic enhancements and intelligibility benefits of speaking style adaptations, a deeper understanding of the link between production and perception is still lacking. Several studies have sought to identify the acoustic-phonetic correlates that shape speech intelligibility. In one line of research, studies have assessed the intelligibility of a range of different talkers in order to find ‘‘intrinsically clear’’ talkers, i.e., talkers who are relatively more intelligible than other talkers. Another line of research has focused on the intelligibility of speaking style adaptations, or "deliberately clear" speech (Markham and Hazan, 2004). Bond and Moore (1994) argued that the acoustic-phonetic characteristics of "intrinsically clear" and "deliberately clear" speech are the same, although they did not directly compare the two.

Bradlow et al. (1996) found F0 range and vowel measures (range in F1, vowel space dispersion, F2-F1 distance for /i/ and F2-F1 distance for /a/) to be more correlated with intelligibility compared to speaking rate and mean F0 when examining intelligibility of "intrinsically clear" talkers. Bond and Moore (1994) found longer word and vowel durations, differentiated vowel space, maximal cues for consonantal contrasts, and low variation in stressed vowel amplitude to characterize intrinsically more intelligible speakers. Other studies of this type found 1-3 kHz energy and word duration to significantly correlate with intelligibility, more so than the long-term average spectrum slope, F0 measures, CV ratios, and vowel formant measures (Hazan and Markham, 2004; Green et al., 2007; van Dommelen and Hazan, 2012).

For studies that have looked at "deliberately clear" speech, i.e. speaking style adaptations intended to enhance intelligibility, it has been common to discuss the set of acoustic-phonetic changes that characterize the adaptations as a whole (slower speaking rate, wider F0 range, expanded vowel space, etc.); few have sought to examine the individual contributions of the different acoustic-phonetic features (i.e. modifications not in conjunction). Some studies (Picheny et al. 1986; Picheny et al., 1989; Uchanski et al., 1996) have manipulated durational cues of conversational and clear speech, showing that longer durations and slower speaking rates do not necessarily lead to greater intelligibility. Krause and Braida (2002) found that deliberately produced clear speech enhanced intelligibility independent of speaking rate. A follow-up study by Krause and Braida (2004) suggested the clear speech increase in 1-3 kHz energy correlated with increased intelligibility (Krause and Braida, 2004). Ferguson and Kewley-Port (2002, 2007) found that talkers who produced a larger clear speech benefit than others showed significantly longer vowel duration and larger vowel space expansion. Ferguson and Quené (2014) found that the contribution of certain acoustic cues to vowel intelligibility in both conversational and clear speech (vowel duration, steady-state formant frequencies) differed depending on whether the listener had hearing loss or not.

There are very few studies that have examined the impact of individual quiet-to-noise modification on intelligibility. Studies have shown that the noise-adapted speech benefit is not solely due to increased intensity (Summers et al., 1988; Pittman and Wiley, 2001; Goy et al., 2007). Others have shown spectral cues to be more important than durational cues for the intelligibility gain of noise-adapted speech (Cooke, Mayo, Villegas, 2014). However, no specific acoustic property has reliably emerged as a significant correlate of intelligibility for noise-adapted speech (Pittman and Wiley, 2001). Previous studies have suggested that there may not be a simple direct relationship between individual acoustic features and word recognition scores, and that intelligibility is likely the result of complex interactions between several acoustic features.

Although some acoustic-correlates of intelligibility have been identified, there is significant variability in the characteristics of intelligible speech. As noted by Hazan and Markham (2004), there is a need to examine talker intelligibility by including: 1) a set of talkers differing in intrinsic clarity, 2) a set of speaking styles differing in deliberate clarity, and 3) a comprehensive list of acoustic-phonetic features, including additional measures that have previously been excluded from such analyses (e.g. voice quality measures). This is the final goal of this dissertation.

GOALS

The major goals of this dissertation are as follows: 1) examine age- and adaptation-related variation in speech production, 2) examine the extent to which age- and adaptation-related changes shape speech intelligibility for young adult listeners (while confirming that age is accurately perceived), 3) examine the extent to which relative talker intelligibility varies across speaking styles, and 4) identify the acoustic-phonetic predictors of speech intelligibility.

Specifically, I will examine variation in speech production and perception as related to adaptation-specific (noise-adapted speech, clear speech) and talker-specific (children, young adults, older adults) factors. By comparing the two styles of intelligibility-enhancing speech directly for these three groups of talkers, this dissertation will reveal: 1) the manner in which children and older adult talkers produce noise-adapted speech and clear speech compared to young adults (Exp. 1), 2) how these factors impact word recognition in noise for young adult listeners (Exp. 2A), and whether these talkers are perceptually representative of their age groups (Exp. 2B), 3) what these variations in speech intelligibility reveal about relative talker intelligibility, and 4) how specific acoustic-phonetic changes account for variation in intelligibility.

In Experiment 1, a battery of acoustic analyses will be performed on conversational and clear speech produced in quiet and in the presence of noise by children, young adults, and older adults. The findings will reveal the acoustic-phonetic correlates of clear speech and noise-adapted speech for talkers of different ages. Experiments 2A and 2B will examine the perceptual impacts of these acoustic modifications. Experiment 2A will examine the effects of age and speaking style on word recognition in noise. The results will provide insight into the combined benefits of listener- and environment-oriented adaptations, and the interactions between age and communicative intent. The results will also be analyzed to further our understanding of factors that may impact relative talker intelligibility. Experiment 2B will examine the talkers' perceived ages, seeking to demonstrate that the talkers used in the study are representative of their age groups. Finally, a number of regressions examining the variance of acoustic measures compared to that of word recognition in noise scores will provide a more comprehensive picture of the perception-production link.

IMPLICATIONS

This dissertation stands to make important contributions to our understanding of intelligibility variability by providing a bigger-picture account of within-talker variability in response to adverse communicative situations for different talker populations, and its effects on speech perception. This is the first study to investigate the extent to which clear speech and noise-adapted speech benefits interact with each other across the lifespan. The results address our need to better understand relative talker intelligibility, acoustic predictors of speech intelligibility, and how age affects the ability to enhance intelligibility. It is important to specify the attributes underlying why some talkers are more intelligible than others.

Knowledge of the inherent variations in speaking style adaptations across age is also crucial to a number of applied realms, from clinical to computational. Audiologic practices and rehabilitation strategies have traditionally assumed a large degree of speaker heterogeneity. Expanding our knowledge of talker variability in production and its effects on perception has the potential to contribute to a wide range of clinical applications such as fitting hearing aids and defining rehabilitative criteria. For example, an understanding of how speech intelligibility varies across age and speaking style is useful for tailoring rehabilitative strategies to the speaker (e.g. spouse vs. child of an adult with severe hearing loss). This knowledge might also help those who struggle to produce intelligible speech (e.g. those afflicted with Parkinson's disease, second language learners, etc.).

Similarly, knowledge of the inherent variations in speech production can be applied to speech technology: speech synthesizers could be designed to incorporate some intelligibility-enhancing features common in clear speech production, while speech recognition systems could be enhanced by taking into account these natural variations in human speech production across talkers and environments. The discovery of which acoustic properties of speech result in greater intelligibility could also be extremely useful in developing new signal processing algorithms in nonlinear hearing aids. Recent work has cited the need for these algorithms to take greater inspiration from speech adaptation research (Godoy et al., 2014).

Experiment 1: Production

RESEARCH AIMS

The goal of this experiment was to examine the specific acoustic-articulatory changes that characterize intelligibility-enhancing speech modifications, both as a function of 1) age (children, young adults, older adults) and 2) communicative intent (speaking conversationally vs. clearly in quiet vs. in response to noise). Very little is known about the acoustic-phonetic modifications underlying these adaptations for children and older adult talkers. Furthermore, while there is a large body of research on both noise-adapted and clear speech modifications, many of the findings are not in accord, and few studies have examined the extent to which the 2 lines of modifications cumulate (i.e. noise-adapted clear speech) (Pichora-Fuller et al., 2010; Gilbert et al., 2014).

METHODS

Talkers

Ten talkers from each of the following populations were recorded: children (CH; 11-13 years old, mean 12.3 years), young adults (YA; 18-29 years old, mean 21.0 years), and older adults (OA; 60-84 years old, mean 70.2 years). Young adult talkers were UT students recruited from the Linguistics Department subject pool (students participating in the course 'Introduction to Linguistics'). Children and older adults were recruited via word-of-mouth and flyers posted in the Austin community.

The age range for children was selected for both its underrepresentation in speech production research and practical purposes. The ability to produce environment- and listener-oriented speaking style adaptations has not been studied extensively in this age group. Additionally, examining older children with more developed reading skills and attention span allowed for elicitation of the full range of stimuli and a direct comparison with the two other talker groups.

All 30 talkers were native monolingual speakers of English, and balanced for sex within each group. Children and young adults were normal-hearing (thresholds below 25 dB SPL at .5, 1, 2, and 4 kHz). Older adults did not rely on the use of hearing aids, although several exhibited a mild degree of sloping hearing loss (thresholds at or below 30 dB at 0.5 and 1 kHz; thresholds at or below 60 dB SPL at 2 kHz; thresholds at or below 85 dB SPL at 4 kHz). Given that high-frequency hearing loss is a common result of aging, these subjects were not excluded (see Table 1 for thresholds). Young adults were all University of Texas at Austin undergraduate students. Children and older adults were recruited from the Austin community. Participants provided written informed consent and were either paid for their participation or received course credit.

Table 1: Hearing thresholds for older adult talkers. Thresholds above 25 dB (threshold used to determine normal hearing) are highlighted in blue.

Talker	L Ear Thresholds (Hz)				R Ear Thresholds (Hz)			
	500	1000	2000	4000	500	1000	2000	4000
OA01	25	25	25	25	25	25	25	25
OA02	25	25	30	45	25	25	40	40
OA03	25	25	25	25	25	25	25	25
OA04	25	25	25	55	25	25	40	45
OA05	25	25	60	85	25	25	45	75
OA06	25	25	25	40	25	25	25	60
OA07	25	25	25	25	25	25	25	25
OA08	35	30	35	50	30	30	45	55
OA09	25	25	25	40	25	25	25	30
OA10	25	25	35	40	25	25	40	40

Materials

Each talker was recorded producing 60 monosyllabic target words embedded in high predictability sentences (e.g. *Farm animals stay in a barn*) (Fallon et al., 2002). These sentences were chosen as Fallon and colleagues showed them to be appropriate for use with children as young as 5 years old. That Fallon found the target words to be predicted from the sentence context enables a test of word recognition that is dependent on information from the entire sentence (see Table 23 in Appendix for the complete list of sentences). Each speaker produced the sentences in conversational speech (CO) and clear speech (CL). These two sets of sentences were recorded first in quiet (Q) and then in response to speech-shaped noise interference (N) presented via headphones (80 dB SPL). Speech-shaped noise (SSN) was generated by superimposing the spectral shape of 6-talker babble onto noise to provide a consistent level of masking across keywords (i.e. no temporal glimpsing windows as there are in babble). The babble consisted of 6 native talkers of American English (3M, 3F)(c.f. Van Engen and Bradlow, 2007 for additional information; recordings were those used to generate the 6-talker English babble masker in Van Engen and Bradlow, 2007). See Figure 1 for a spectral slice of the SSN, and Tables 2 and 3 for the instructions used to elicit the speaking style adaptations.

Figure 1: Spectral slice of the SSN, with frequency along the x axis and intensity on the y axis

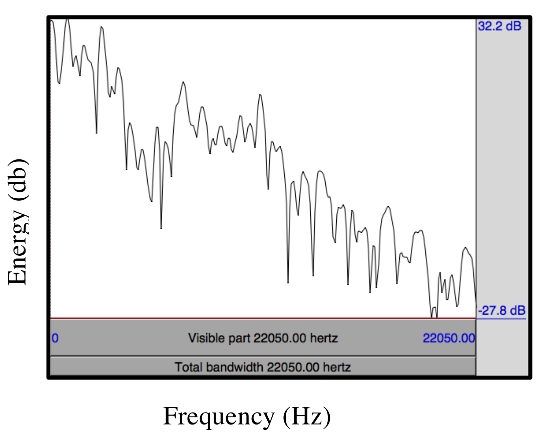


Table 2: Elicitation instructions for adults

Quiet conversational speech (QCO)	"*Speak normally and conversationally, as if you are talking to a friend or family member.*"
Quiet clear speech (QCL)	"*Speak as if you are trying to communicate with someone who has a low proficiency in English, and does not follow you conversationally.*"
Noise-adapted conversational speech (NCO)	*"You will hear some background noise. Try to speak as if you are communicating in this noisy environment. Speak normally and conversationally, as if you are talking to a friend or family member in a noisy place."*
Noise-adapted clear speech (NCL)	*"You will again hear some background noise. Try to speak as if you are communicating in this noisy environment to someone who has a low proficiency in English, and does not follow you conversationally in a noisy place.*"

Table 3: Elicitation instructions for children

Quiet conversational speech (QCO)	"*Speak like you normally do when you talk to your friends.*"
Quiet clear speech (QCL)	"*Speak like you are trying to talk to someone really old, or someone who doesn't know English very well.*"
Noise-adapted conversational speech (NCO)	*"Now I am going to play some noise through these headphones. Pretend you are in a noisy place. Speak like you normally do when you talk to your friends in a noisy place."*
Noise-adapted clear speech (NCL)	*"Pretend you are in a noisy place again. Speak like you are trying to talk to someone really old, or someone who doesn't know English very well.*"

These elicitation instructions are in line with those used in previous research (Pittman and Wiley, 2001; Hanley and Steer, 1949; Picheny et al. 1986; Schum 1996; Krause and Braida 2002; Ferguson and Kewley-Port 2002; Ferguson 2004; Smiljanic and Bradlow 2005; Smiljanic and Bradlow, 2009). While spontaneously-elicited clear and noise-adapted speech may differ from the productions obtained in the lab, it is likely that the results reported here underestimate the intelligibility-enhancing modifications that occur in spontaneous speech. Using speech recorded in a lab under specific instructions allowed a more controlled examination of the relationship between age and within-talker variability.

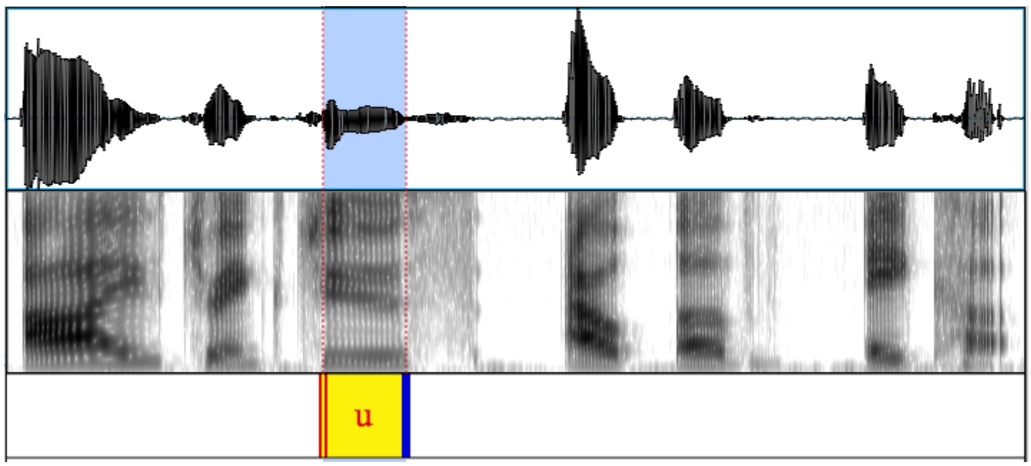
Acoustic analysis

Eleven separate acoustic analyses were performed on all sentences (60 sentences x 30 talkers x 4 speaking styles=7200 recorded sentences total) in order to assess the extent to which the 3 talker groups (children, young adults, older adults) differed in their implementation of noise-adapted and clear speech (see Table 4 for the list of acoustic measurements). These acoustic features have been indicated in previous studies to be affected by age or by speaking style. Importantly, several of the acoustic features common in developmental and aging research have never been included in the analysis of speaking style modifications (e.g. jitter, shimmer, harmonics-to-noise ratio). Given that the speaking style adaptations examined here incur a large degree of vocal effort, examining these novel measures may provide new insight into acoustic-articulatory correlates of intelligibility-enhancing modifications.

Table 4: List of acoustic features analyzed for Experiment 1

No.	Type	Acoustic feature
1	Global	F0 range
2		F0 mean
3		1-3 kHz energy
4		Speaking rate
5		Pause duration
6	Segmental	Vowel duration
7		F1
8		F2
9	Voice	Harmonics-to-noise ratio (HNR)
10		Jitter
11		Shimmer

Figure 2: Example vowel labeling for /u/ in “I drink juice out of a cup”



In order to perform analyses, sentences were manually annotated using Praat textgrids (Boersma and Weenink, 2007). Praat scripts were then run in order to obtain acoustic values. The first 5 measurements listed in Table 4 were obtained from all sentences (pauses were defined with a minimum duration of 100 ms). Measurements 6 through 8 were obtained from a subset of vowel tokens. These consisted of corner vowels (/i, ɑ, ae, u/) embedded in monosyllabic words between 2 obstruents. Measurements were based on 2 tokens per vowel per style per speaker (960 total), 1 from a sentence-final content word and 1 from a mid-sentence content word. See Table 24 in the Appendix for the list of tokens. Measurements 9 to 11 were obtained from all /ɑ/ tokens (2 per style per speaker, 240 total). HNR was analyzed using the ‘to Harmonicity (cc)’ command, jitter with the ‘jitter (local)’ command, and shimmer with the ‘shimmer (local)’ command. Measurements obtained in Hz values were analyzed as such, as each age group was balanced for gender, and the focus of this analysis was on the articulatory characteristics of the speaking style changes in each age group.

Statistical Analysis

Each acoustic feature was analyzed using a mixed effects linear regression in SPSS with Talker Age Group (children, young adults, or older adults), Environment-Oriented Speaking Style (produced in quiet or in response to noise), Listener-Oriented Speaking Style (conversational or clear), and their interactions as fixed effects. To account for talker and item variability, random intercepts for Talker and Sentence were included as well. Random slopes were included in the model for both Listener- and Environment- Oriented Speaking Style at the level of Talker, since this level showed the greatest variance. For F1 and F2 analyses, Vowel Type (/i, ɑ, ae, u/) was added as an additional fixed effect in order to examine how formant frequencies differed for each of the 4 corner vowels. These models determined the impact of age and communicative intent on production.

Hypotheses

As this was the first study to compare the acoustic-phonetic modifications that characterize noise-adapted and clear speech adaptations in children, young adults, and older adults, I did not hold specific predictions for every interaction. There still remained a number of unknowns, debated findings, and results that had yet to be replicated. However, given patterns that had been established in previous research, I outlined several hypotheses below.

Global

Regarding the effects of age, I held no specific predictions for F0 range. In terms of mean F0, however, I expected children to exhibit a higher mean F0 than adults. Some studies have found F0 to be slightly higher in older adults compared to young adults (Stathopoulos et al., 2011) while other studies have found F0 to be lower in older adults

(Ramig et al., 2001). I therefore held no specific predictions regarding the variation of mean F0 in adult populations. Given age-related decreases to the long-term average spectrum found in Linville and Rens (2001), I hypothesized that 1-3 kHz energy would decrease with age. I also expected older adults to show a slower speaking rate than young adults (c.f. Smiljanic, 2013). There was no evidence showing that speaking rate would differ between 11-13 year olds and young adults (Lee et al., 1999; Pettinato and Hazan, 2013). I also held no specific predictions regarding pause duration, although I expected this to highly co-vary with speaking rate (longer pauses in conjunction with slower speaking rate).

Regarding the effects of speaking style adaptations, I hypothesized that (compared to quiet and conversational speech, respectively) noise-adapted and clear speech would exhibit an increased mean F0, a wider F0 range, and greater energy in the 1-3 kHz region. They would also show a slower speaking rate with increased pause duration. Previous work in our lab (Gilbert et al., 2014) showed no interaction between the 2 styles for F0 range, but found that noise-adapted clear speech had a mean F0 much higher than the increases of noise-adapted and clear speech separately. It is important to note that this study was based on one talker, so it was not clear whether these findings would be replicated by the current study, which includes a large number of talkers of varying ages. Previous work has also shown this clear speech-induced speaking rate decrease to comparably affect children and young adults, while older adults made smaller changes to speaking rate (Pettinato and Hazan, 2013; Smiljanic, 2013).

Segmental

Regarding the effects of age, I expected vowel duration to show the same patterns as the other durational measures of speaking rate and pause duration (longer vowels in conjunction with slower speaking rate and longer pauses). That is to say, I expected older adults to show longer vowels than young adults, but I held no specific predictions regarding vowel duration for children vs. adults. However, differences between vowel duration and speaking rate could arise given that vowel duration is just an indication of lengthened segment durations, while speaking rate is both an index of lengthened segment durations as well as greater articulatory precision (e.g. more frequent stop burst releases). I expected formant frequencies to lower with age, given that previous work found older adults to show lower formant frequencies than young adults (Xue and Hao, 2003; Linville and Fisher, 1985; Endres et al., 1971; Scukanec et al., 1991), and young adults to show lower formant frequencies than children (Lee et al., 1999).

In terms of the effects of speaking style adaptations, I expected noise-adapted and clear speech adaptations to show longer vowel durations. Given that the clear speech-induced speaking rate decrease was comparable for children and young adults, but smaller for older adults, I expected vowel duration to show similar results. Given that intelligibility-enhancing speaking style adaptations often exhibit an increase in vowel space area, I hypothesized that F1 for the low vowels and F2 for the front vowels should increase in clear and noise-adapted speech. However, I expected a potentially smaller effect for children than for young adults, given that only young adults (and not children) hyperarticulated their clear speech vowels in Pettinato and Hazan (2013).

Voice

In terms of age-related changes, studies have shown the harmonics-to-noise ratio (HNR) decreases and shimmer increases with old age, although findings on jitter remain debated (Ferrand, 2002). Ramig and Ringel (1983) suggested this is heavily due to physical condition more so than chronological age. Data on these measures for children compared to adults is scarce; studies have shown HNR to be higher in young adults than in children (Ferrand, 2000; Stathopoulos et al., 2011) while another showed jitter to be independent of age in children 7 to 15 years old (Linders et al., 1995). Another study showed that, when speaking loudly, children increased HNR and decreased jitter and shimmer (Glaze et al., 1990). This suggests that, at least for children, the increased vocal effort (characteristic of speaking style adaptations) may reduce the amount of perturbations and noise in these measures. Given the paucity of established findings, I did not hold specific predictions regarding age-related effects on voice quality, though I suspected perhaps older adults might show increased noise in terms of reduced HNR and increased shimmer (compared to young adults).

In terms of speaking style adaptations, I suspected that the increased vocal effort necessary to produce noise-adapted and clear speech might reduce perturbations in voice quality (i.e. increased HNR, decreased jitter and shimmer). However, this has never been shown.

RESULTS

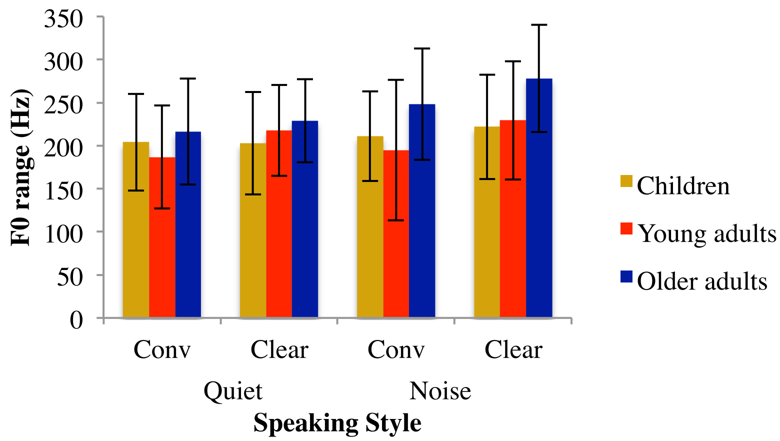
An overview of the results for each of the 11 acoustic features can be found in Tables 25 and 26[2] in the Appendix. Estimates and pairwise comparisons can be found in Tables 27 to 45 in the Appendix.

Global

F0 range

[2] For HNR, jitter, and shimmer, the final Hessian matrix was not positive definite although all convergence criteria were satisfied.

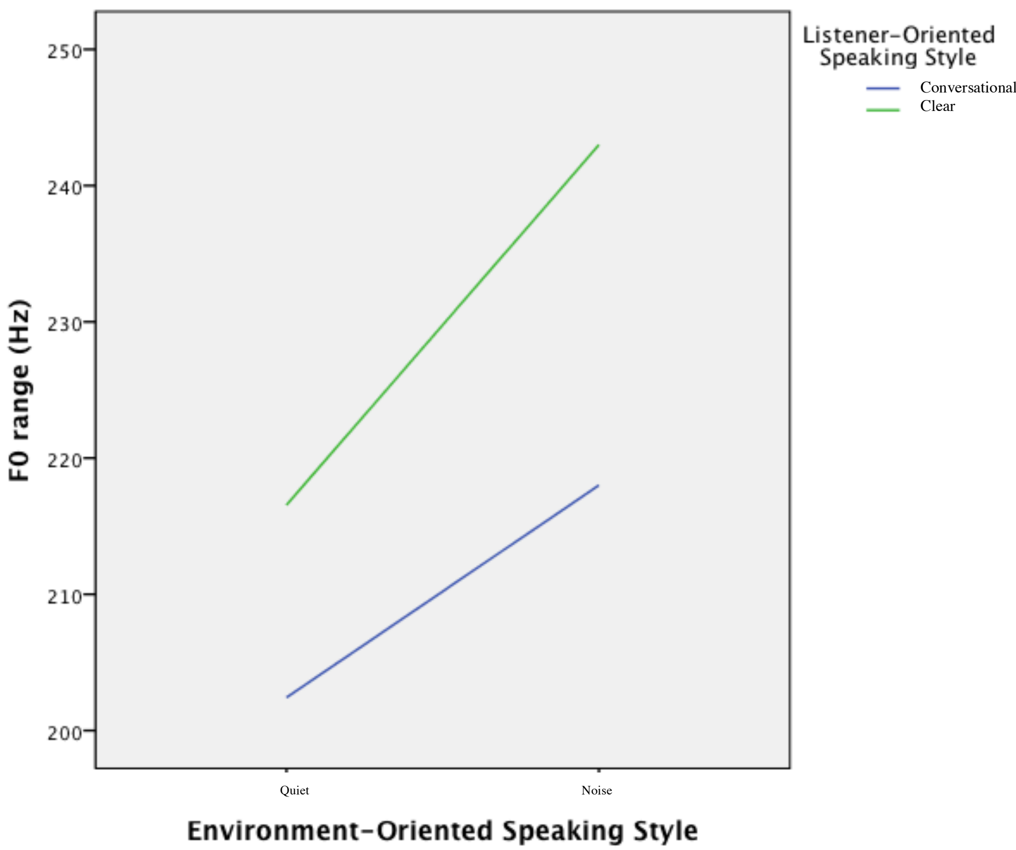
Figure 3: F0 range for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



See Figure 3 for an overview of results for the 3 talkers in the 4 speaking styles. There was no significant main effect of Talker Age Group on F0 range [F(2,27)=1.572, p=0.226]. However, there was a significant main effect of both Environment-Oriented Speaking Style [F(1,53.995)=6.444, p=0.014] and Listener-Oriented Speaking Style [F(1,53.995)=5.627, p=0.021], with F0 range showing a significant increase in noise-adapted speech and clear speech compared to baseline (speech produced in quiet and conversational speech, respectively) (see Tables 30 and 32).

Additionally, there was a significant 2-way interaction between Environment- and Listener-Oriented speaking style adaptations [F(1,7041.042)=5.060, p=0.025] (see Figure 4). The interaction revealed that the clear speech modifications to F0 range were only significant in noise-adapted speech. Likewise, the quiet-to-noise modifications were only significant in clear speech (see Tables 40 and 41). It appears that the main effect of Environment-Oriented speaking style on F0 range mainly arose from the F0 range increase that was present for clear speech (but nonsignificant for conversational speech). Likewise, the main effect of Listener-Oriented Speaking Style on F0 range mostly originated from the F0 range increase for noise-adapted speech (which was nonsignificant for speech produced in quiet).

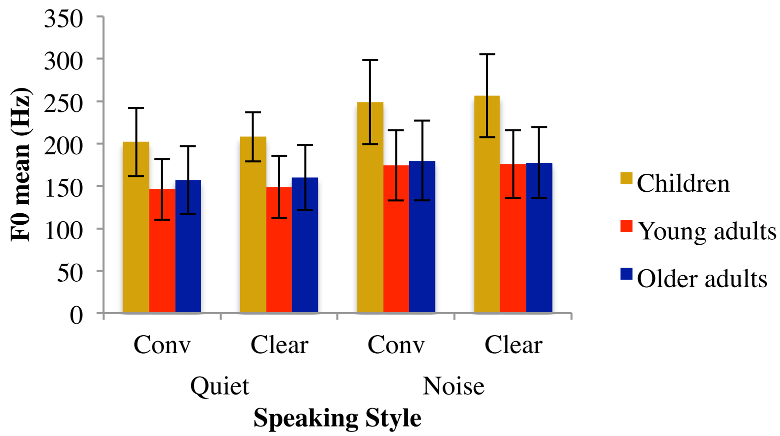
Figure 4: Interaction between Environment- and Listener-Oriented Speaking Styles for F0 range; x axis shows speech in quiet vs. in response to noise; lines represent conversational and clear speech



None of the other 3 interactions were significant: Environment-Oriented Speaking Style*Talker Age Group [F(2,53.995)=1.362, p=0.265], Listener-Oriented Speaking Style*Talker Age Group [F(2,53.995)=0.960, p=0.389], and Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [F(2,7041.042)=0.613, p=0.542].

F0 mean

Figure 5: F0 mean for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



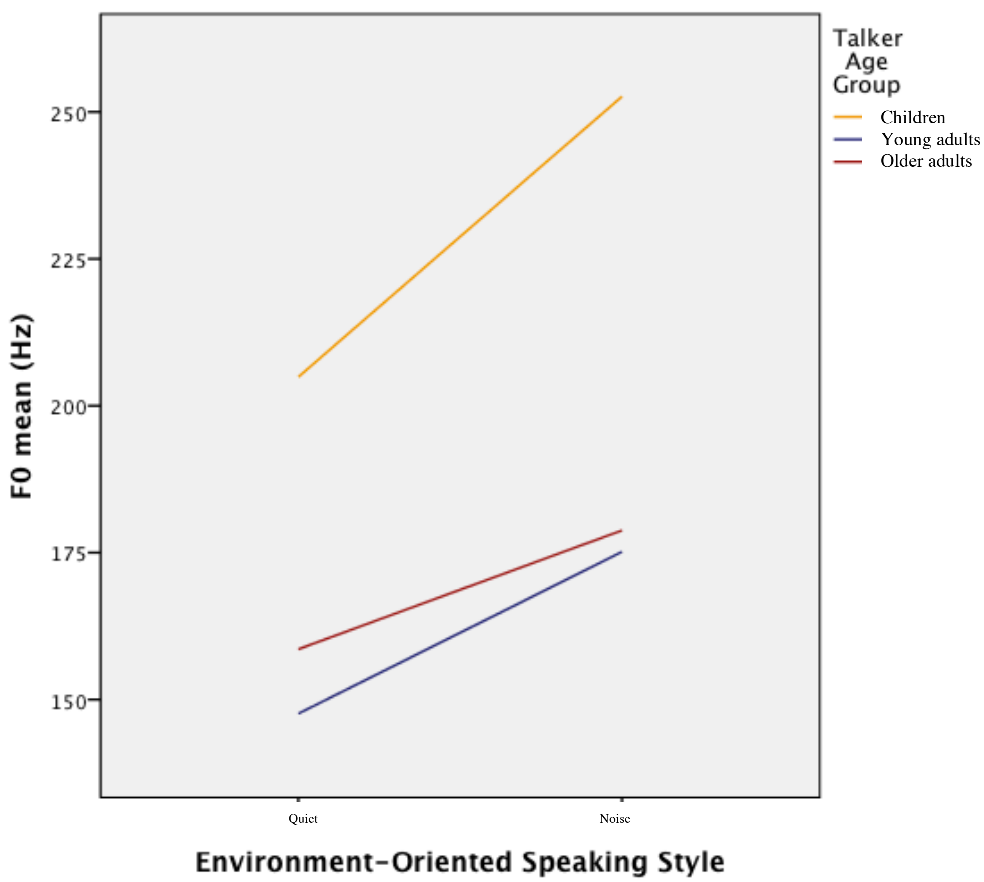
See Figure 5 for an overview of results for the 3 talkers in the 4 speaking styles. There was a significant main effect of Talker Age Group on F0 mean [F(2,27)=8.830, p=0.001]. Children showed a significantly higher mean F0 than both adult groups (see Table 28). A significant main effect of Environment-Oriented Speaking Style was also found, with mean F0 showing a significant increase in noise-adapted speech compared to speech produced in quiet [F(1,54.001)=122.514, p<0.001] (see Table 30). There was no significant main effect of Listener-Oriented Speaking Style [F(1,54.001)=1.130, p=0.292].

Several significant interactions for F0 mean also emerged. There was a significant 2-way interaction between Environment-Oriented Speaking Style and Talker Age Group [F(2,54.001)=8.184, p=0.001] (see Figure 6); pairwise comparisons showed that children, who although speaking with a significantly higher mean F0 than adults in quiet, made a significantly larger F0 increase than adults in response to noise (see Tables 34 and 35).

This shows that the main effect of Environment-Oriented Speaking Style on F0 mean is mainly due to children's F0 mean increase compared to the adults' increase.

Environment-Oriented Speaking Style also interacted with Listener-Oriented speaking style [$F(1,7041.013)=8.385$, $p=0.004$]. Pairwise comparisons revealed that, while noise-induced changes were significant both in conversational and clear speech, the increase was larger for quiet-to-noise changes in conversational speech than for those in clear speech (see Tables 40 and 41).

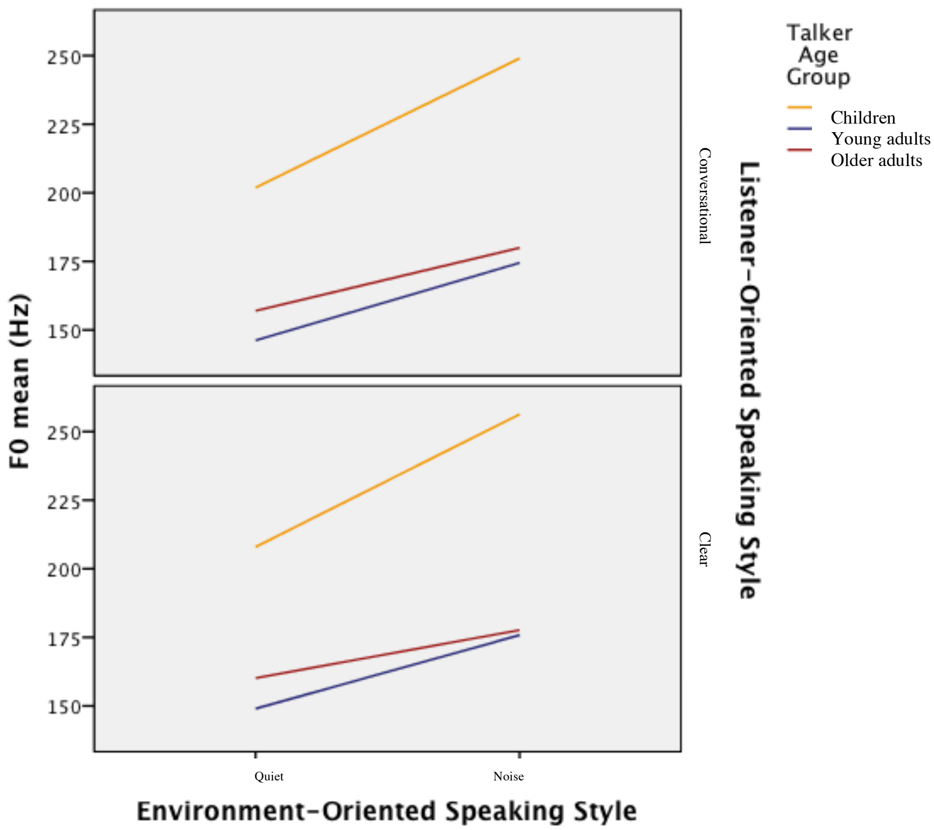
Figure 6: Interaction between Talker Age Group and Environment-Oriented Speaking Style for mean F0; x axis shows speech in quiet vs. in response to noise; lines represent children, young adult, and older adult talkers



Finally, a 3-way interaction was found between all 3 fixed effects: Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style [$F(2,7041.013)=9.311$, $p<0.001$] (see Figure 7 and Tables 43 to 45 for pairwise comparisons). While all talker groups showed significantly higher F0 in noise relative to quiet, the extent of the modification was modulated by Talker Age Group and Listener-Oriented Speaking Style. The quiet-to-noise F0 increase for older adults' clear speech was less than any of the other quiet-to-noise increases across talker groups as well as across listener-oriented speaking style conditions. That is to say, it was smaller than the older adults' quiet-to-noise increase in conversational speech, as well as the younger talker groups' quiet-to-noise increase in both conversational and clear speech. It appears that the interaction between Environment- and Listener-Oriented Speaking Styles (in which the quiet-to-noise changes were larger for conversational speech than for clear speech) arose from the older adult's significantly smaller quiet-to-noise increase in clear speech.

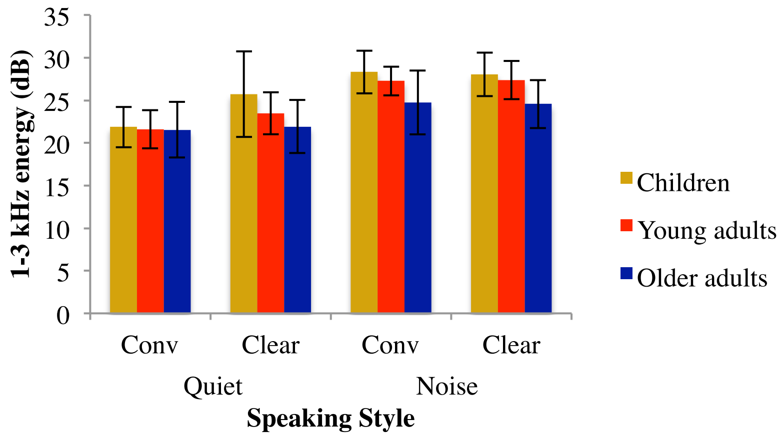
There was no significant interaction between Talker Age Group and Listener-Oriented Speaking Style [$F(2,54.001)=0.421$, $p=0.659$].

Figure 7: Interaction between Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style for mean F0; x axis shows speech in quiet vs. in response to noise; panels show conversational speech (at top) and clear speech (at bottom); lines represent children, young adult, and older adult talkers



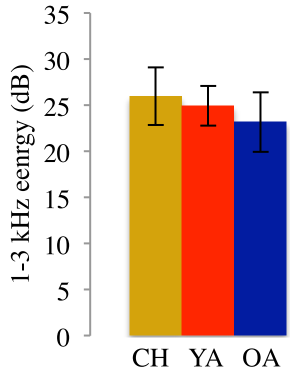
1-3 kHz energy

Figure 8: 1-3 kHz energy for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



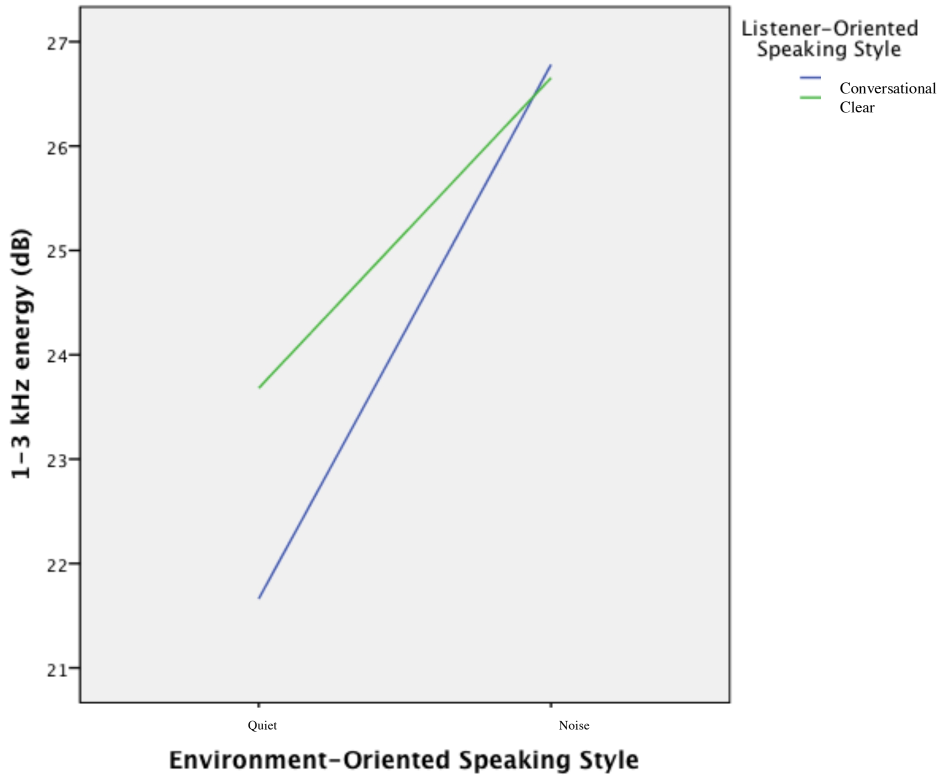
See Figure 8 for an overview of results for the 3 talkers in the 4 speaking styles. There was a significant main effect of Talker Age Group on 1-3 kHz energy [F(2,27)=5.444, p=0.010] (see Figure 9). Children (and marginally, young adults) spoke with significantly more energy in the 1-3 kHz region compared to older adults (see Table 28). There was also a significant main effect of Environment-Oriented Speaking Style [F(1,53.997)=58.312, p<0.001], with 1-3 kHz energy showing a significant increase in noise-adapted speech compared to speech produced in quiet (see Table 30). There was no significant main effect of Listener-Oriented Speaking Style [F(1,53.997)=3.205, p=0.079].

Figure 9: Main effect of Talker Age Group for 1-3 kHz energy; bars represent children (CH), young adult (YA), and older adult (OA) talkers. Error bars represent standard error,



Results revealed 2 significant interactions for 1-3 kHz energy. There was a significant 2-way interaction between Environment- and Listener-Oriented speaking style adaptations [F(1,7041.002)=416.141, p<0.001] (see Figure 10). The interaction revealed that the conversational-to-clear speech increase in 1-3 kHz energy was only significant for speech produced in quiet. And while noise-induced changes were significant for both conversational and clear speech, they were larger for conversational speech (see Tables 40 and 41).

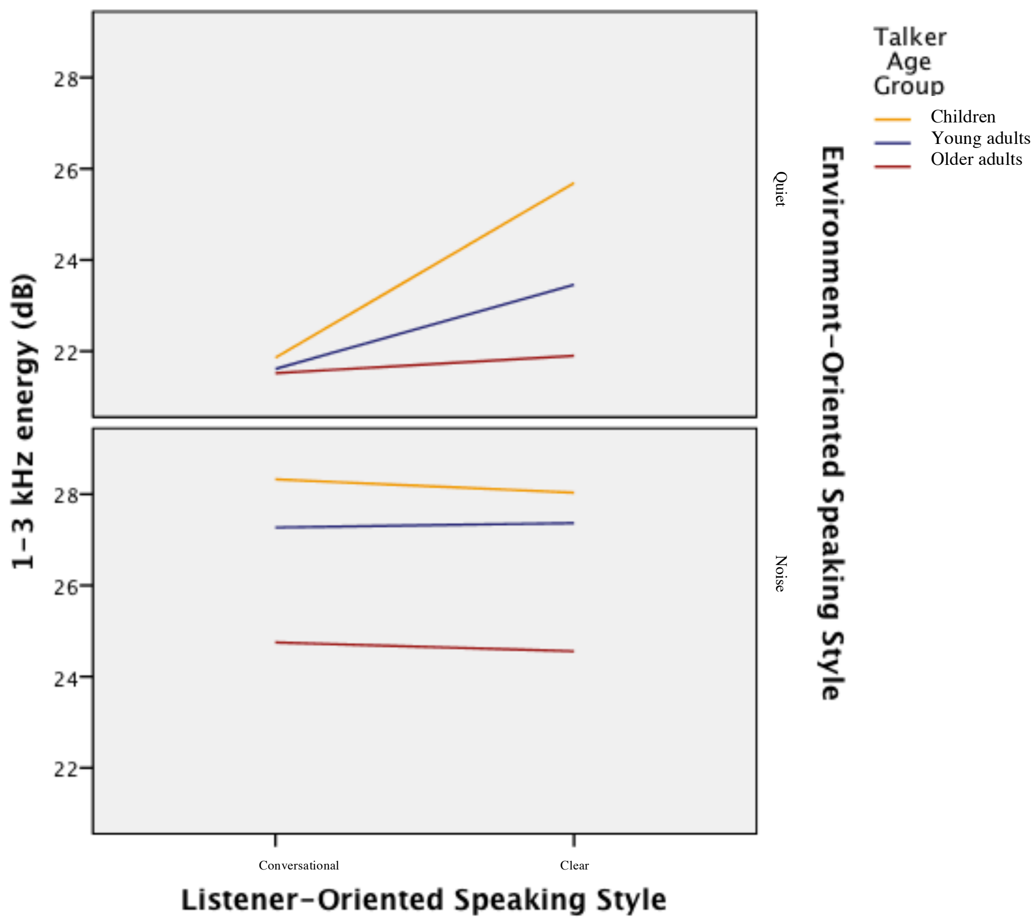
Figure 10: Interaction between Environment- and Listener-Oriented Speaking Styles for 1-3 kHz energy; x axis shows speech in quiet vs. in response to noise; lines represent conversational and clear speech



There was also a significant 3-way interaction between all 3 fixed effects: Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style [$F(2,7041.002)=9.311$, $p<0.001$] (see Figure 11 and Tables 43 to 45 for pairwise comparisons). The only significant conversational-to-clear speech increases in 1-3 kHz energy were made by children and young adult talkers speaking in quiet. The older adult talkers did not make significant conversational-to-clear speech modifications in quiet, and none of the talker groups made significant conversational-to-clear speech modifications in noise. Additionally, children exhibited significantly more 1-3 kHz energy than older adults in every speaking style except conversational speech produced in quiet.

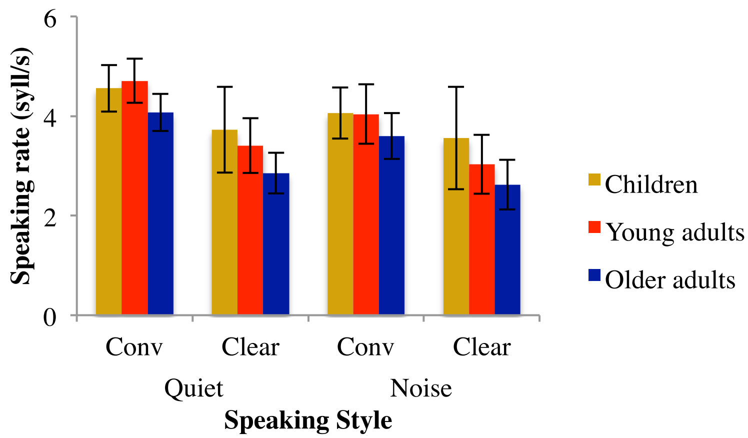
There were no other significant interactions: Environment-Oriented Speaking Style*Talker Age Group [F(2,53.997)=1.116, p=0.335] and Listener-Oriented Speaking Style*Talker Age Group [F(2,53.997)=0.844, p=0.436].

Figure 11: Interaction between Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style for 1-3 kHz energy; x axis shows conversational vs. clear speech; panels show speech in quiet (at top) and in response to noise (at bottom); lines represent children, young adult, and older adult talkers



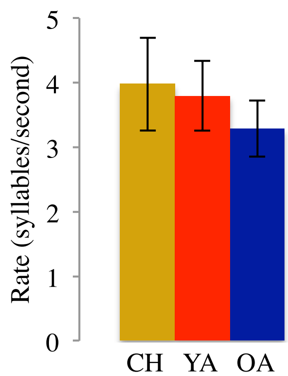
Speaking rate

Figure 12: Speaking rate for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



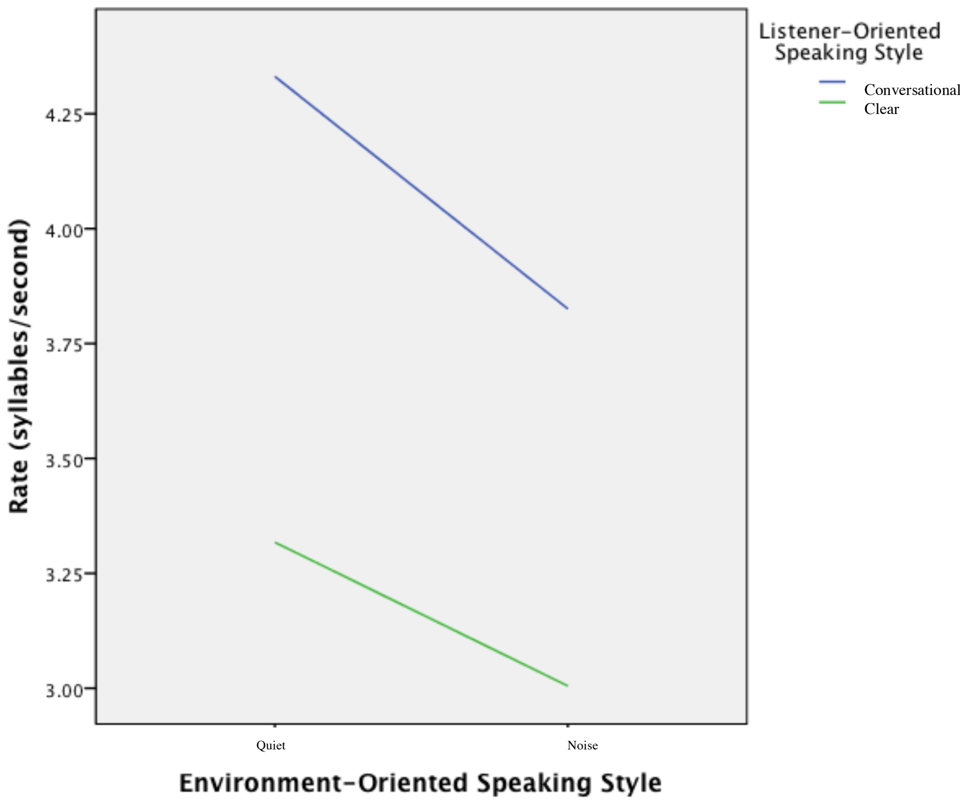
See Figure 12 for an overview of results for the 3 talkers in the 4 speaking styles. There was a significant main effect of Talker Age Group on speaking rate [$F(2,27.004)=8.526$, $p=0.001$] (see Figure 13). Both younger talker groups spoke significantly faster than the older adults (see Table 28). There were also significant main effects of both Environment-Oriented Speaking Style [$F(1,53.981)=26.801$, $p<0.001$] and Listener-Oriented Speaking Style [$F(1,53.981)=135.177$, $p<0.001$]; noise-adapted speech and clear speech were significantly slower compared to speech produced in quiet and conversational speech, respectively (see Tables 30 and 32).

Figure 13: Main effect of Talker Age Group for speaking rate; bars represent children (CH), young adult (YA), and older adult (OA) talkers. Error bars represent standard error.



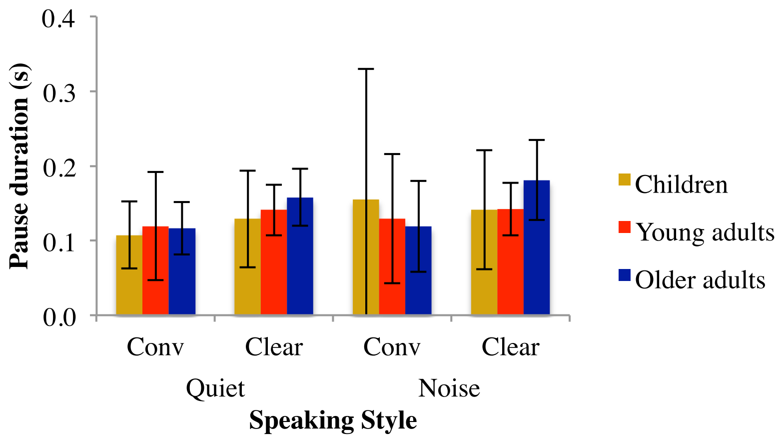
Only one significant interaction was found for speaking rate; Environment- and Listener-Oriented speaking style adaptations significantly interacted [F(1,7008.697)=416.141, p<0.001] (see Figure 14). While both noise-adapted and clear speech modifications were significant, the quiet-to-noise decrease was larger for conversational speech compared to clear speech. Likewise, the conversational-to-clear decrease was larger for speech produced in quiet than speech produced in response to noise (see Tables 40 and 41). Listener-Oriented Speaking Style marginally interacted with Talker Age Group [F(2,53.981)=3.048, p=0.056], with child-adult speaking rate differences being larger in clear speech due to children slowing down less than adults when producing clear speech (see Tables 37 and 38). No other significant interactions were found: Environment-Oriented Speaking Style*Talker Age Group [F(2,53.981)=0.949, p=0.394], and Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [F(2,7008.680)=1.069, p=0.343].

Figure 14: Interaction between Environment- and Listener-Oriented Speaking Styles for speaking rate; x axis shows speech in quiet vs. in response to noise; lines represent conversational and clear speech



Pause duration

Figure 15: Pause duration for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



See Figure 15 for an overview of results for the 3 talkers in the 4 speaking styles. There was no significant main effect of Talker Age Group [$F(2,27.016)=1.449$, $p=0.252$] nor Environment-Oriented Speaking Style [$F(1,54.064)=1.122$, $p=0.294$]. However, there was a significant main effect of Listener-Oriented Speaking Style [$F(1,54.064)=47.923$, $p<0.001$], with pause duration significantly increased in clear speech compared to conversational speech (see Table 32).

Results revealed several interactions for pause duration. There was a significant 2-way interaction between Listener-Oriented Speaking Style and Talker Age Group [$F(2,7000.855)=4.321$, $p=0.018$] (see Figure 16); while all talker groups made significant conversational-to-clear increases in pause duration, older adults produced significantly longer pauses than younger talkers in clear (but not in conversational) speech (see Tables 37 and 38). Listener-Oriented Speaking Style also interacted with Environment-Oriented Speaking Style [$F(1,7000.855)=17.408$, $p<0.001$] (see Figure 17). While the conversational-to-clear speech changes were significant in both quiet and in noise-adapted speech, they were larger in noise-adapted speech than in quiet (see Tables 40 and 41).

Figure 16: Interaction between Talker Age Group and Listener-Oriented Speaking Style for pause duration; x axis shows conversational and clear speech; lines represent children, young adult, and older adult talkers

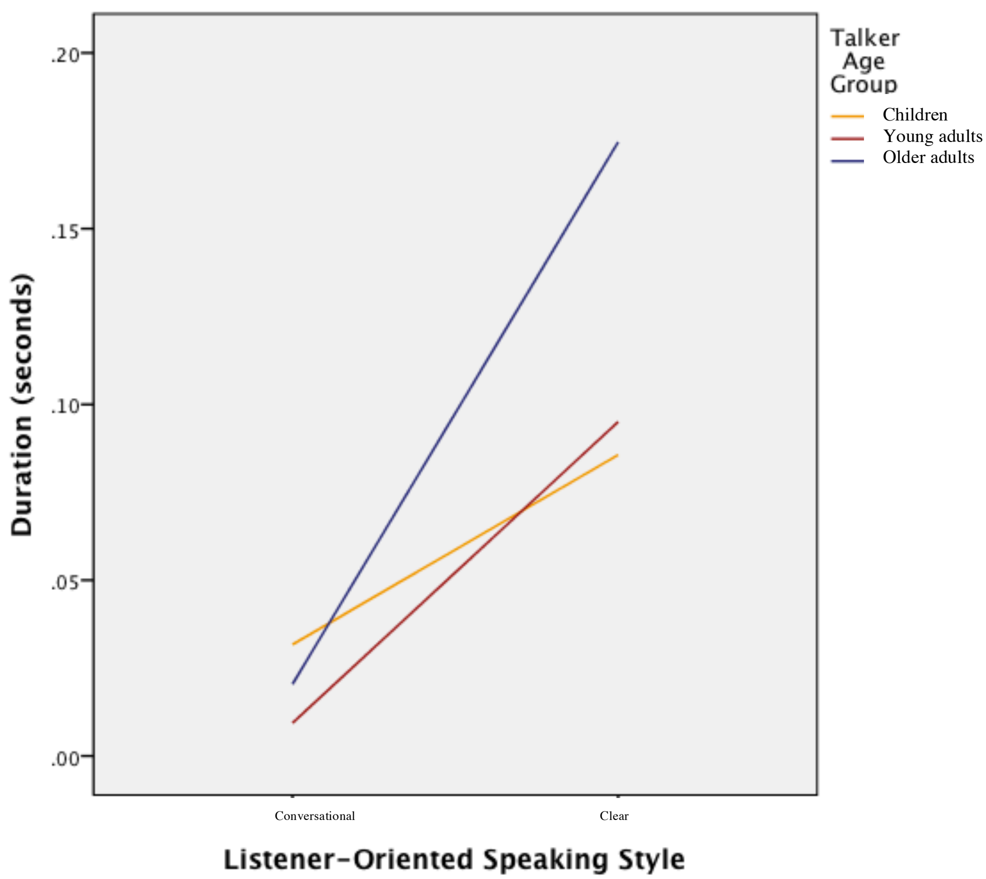
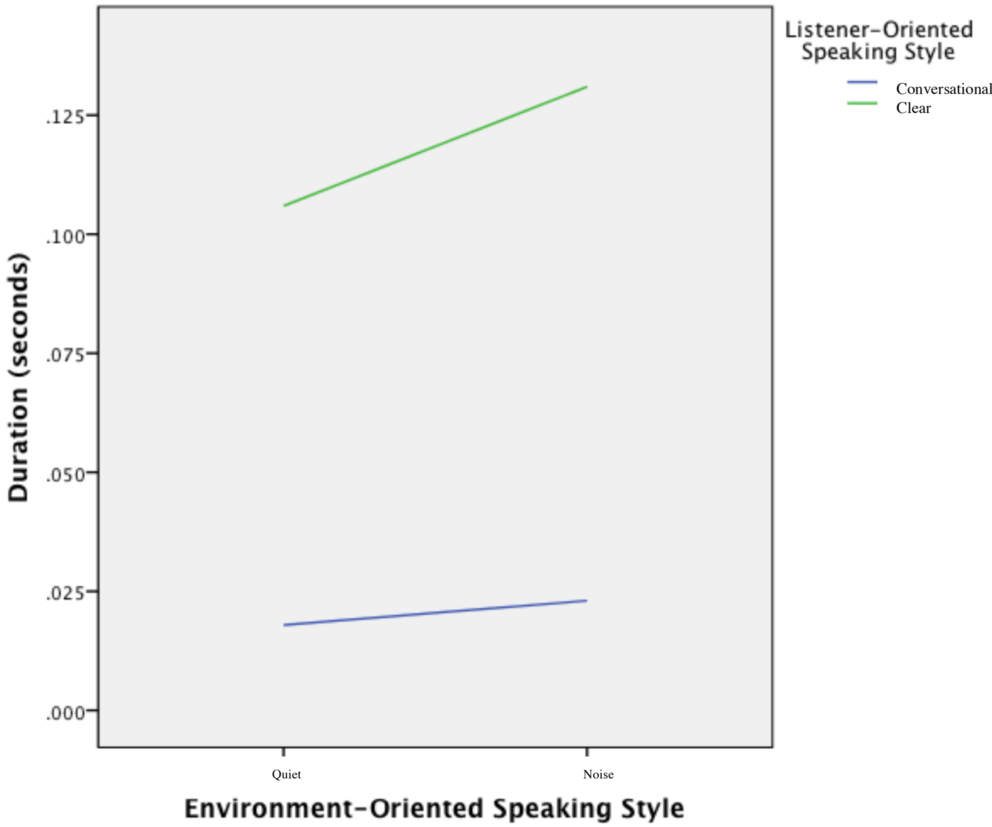


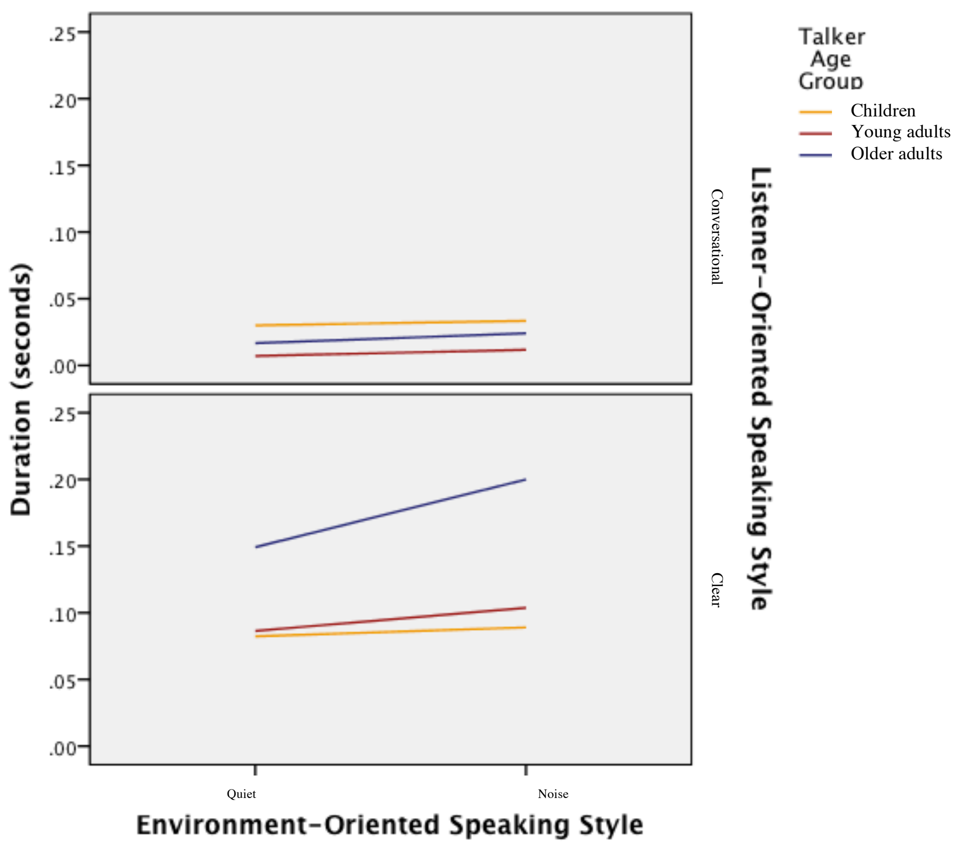
Figure 17: Interaction between Environment- and Listener-Oriented Speaking Styles for pause duration; x axis shows speech in quiet vs. in response to noise; lines represent conversational and clear speech



Finally, a 3-way interaction was found between all 3 fixed effects: Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style [F(2,7000.839)=6.502, p=0.002] (see Figure 18 and Tables 43 to 45 for pairwise comparisons). The only significant quiet-to-noise increase in pause duration occurred for older adults in clear speech. That is, only older adult talkers lengthened pause duration for noise-adapted clear speech compared to clear speech in quiet. In conversational speech, and for the younger talker groups in general, there were no significant quiet-to-noise increases in pause duration. As a result, older adults significantly differed from the younger talkers in noise-adapted clear speech. In all other styles, the talker groups were statistically similar.

There was no significant interaction between Talker Age Group and Environment-Oriented Speaking Style [F(2,54.064)=0.246, p=0.783].

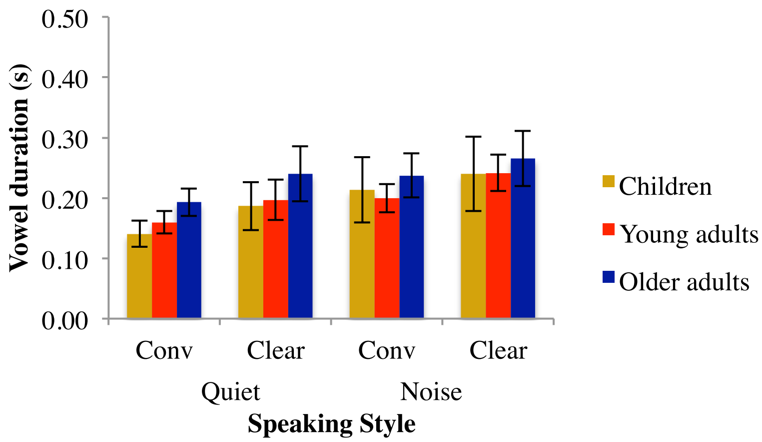
Figure 18: Interaction between Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style for pause duration; x axis shows speech in quiet vs. in response to noise; panels show conversational speech (at top) and clear speech (at bottom); lines represent children, young adult, and older adult talkers



Segmental

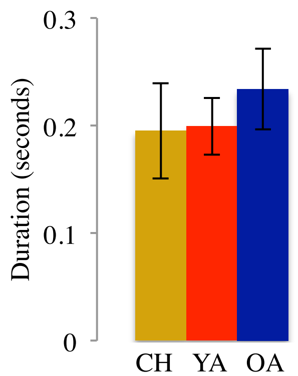
Vowel duration

Figure 19: Vowel duration for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



See Figure 19 for an overview of results for the 3 talkers in the 4 speaking styles. There was a significant main effect of Talker Age Group on vowel duration [F(2,27)=4.305, p=0.024] (see Figure 20), with older adults showing significantly longer vowels than both younger groups (see Table 28). There were significant main effects of both Environment-Oriented Speaking Style [F(1,54)=93.220, p<0.001] and Listener-Oriented Speaking Style [F(1,54)=61.175, p<0.001]; vowel duration was significantly lengthened in noise-adapted and clear speech relative to their baselines (speech in quiet and conversational speech, respectively) (see Tables 30 and 32).

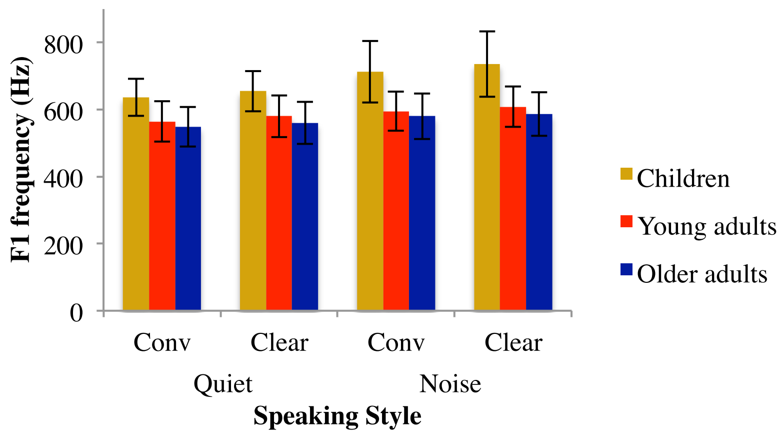
Figure 20: Main effect of Talker Age Group for vowel duration; bars represent children (CH), young adult (YA), and older adult (OA) talkers. Error bars represent standard error.



While there were no significant interactions found for vowel duration, 2 approached significance. Environment-Oriented Speaking Style marginally interacted with Talker Age Group [F(2,54)=3.028, p=0.057]. Although all talker groups significantly lengthened their vowel durations in response to noise, younger talker groups made larger increases. Thus, while older adults produced significantly longer vowels than the younger groups in quiet, there were no significant differences across the 3 talker groups in noise (see Tables 34 and 35). Environment-Oriented Speaking Style also marginally interacted with Listener-Oriented Speaking Style [F(1,860)=3.861, p=0.051]. Although both noise-adapted and clear speech modifications were significant, the quiet-to-noise lengthening was larger for conversational speech than for clear speech vowels, while the conversational-to-clear speech lengthening was larger for speech produced in quiet than speech produced in response to noise (see Tables 40 and 41). No other interactions were significant: Listener-Oriented Speaking Style*Talker Age Group [F(2,54)=0.048, p=0.953] and Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [F(2,860)=2.004, p=0.135].

F1

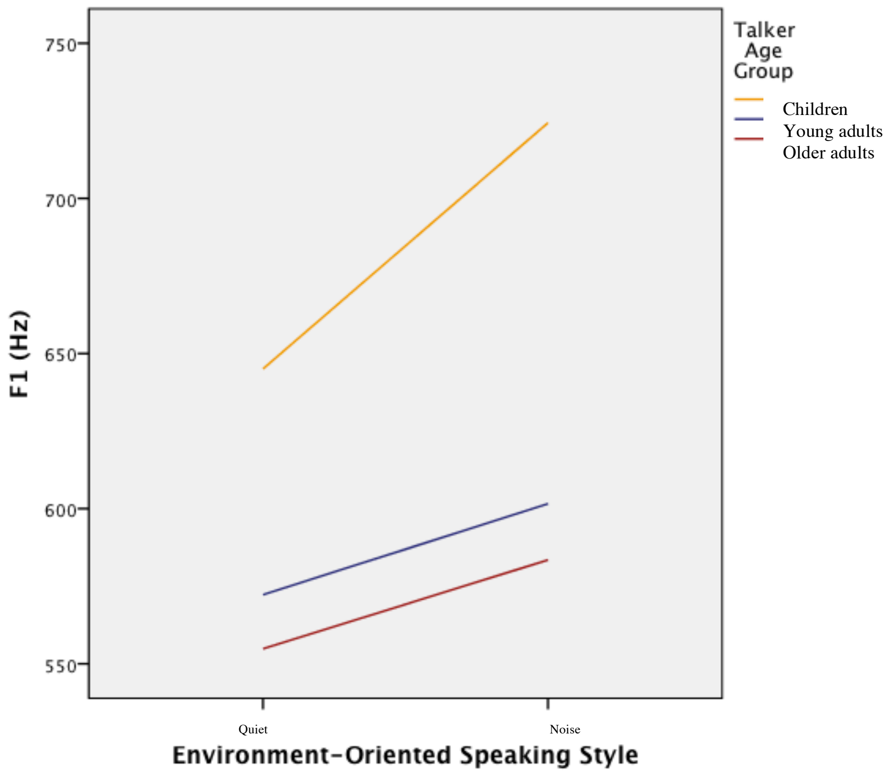
Figure 21: F1 for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



See Figure 21 for an overview of results for the 3 talkers in the 4 speaking styles. As expected, there was a significant main effect of Vowel, given that /ɑ, ae, i, u/ systematically differ in F1 [F(3,4)=171.994, p<0.001]. There was also a significant main effect of Talker Age Group on F1 frequency [F(2,27)=9.458, p=0.001]; children overall had higher F1s than adults, consistent with smaller vocal tracts (see Table 28). In terms of speaking styles, there were significant main effects of both Environment-Oriented Speaking Style [F(1,54)=66.935, p<0.001] and Listener-Oriented Speaking Style[F(1,54)=6.552, p=0.013], with F1 significantly raised in noise-adapted speech compared to speech produced in quiet, and in clear speech compared to conversational speech (see Tables 30 and 32).

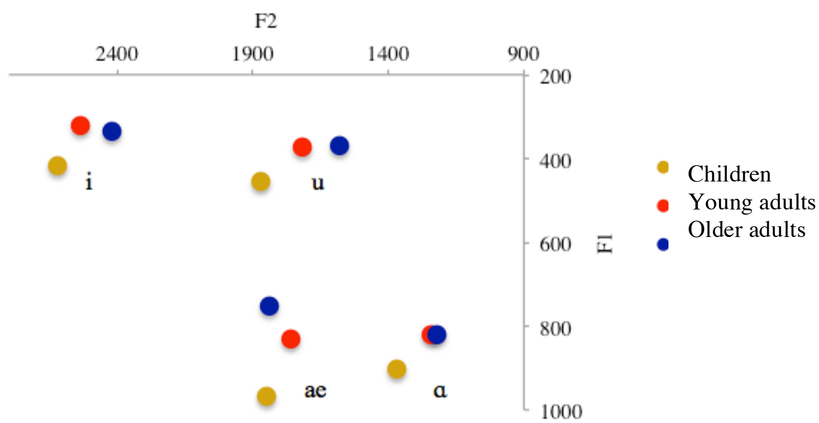
There was a significant interaction between Environment-Oriented Speaking Style and Talker Age Group [F(2,54)=8.989, p<0.001] (see Figure 22). All talker groups made significant quiet-to-noise increases in F1, but children raised F1 more than either adult group. As a result, the child-adult difference was even larger in noise-adapted speech compared to speech in quiet (albeit significant in both) (see Tables 34 and 35).

Figure 22: Interaction between Talker Age Group and Environment-Oriented Speaking Style for F1; x axis shows speech in quiet vs. in response to noise; lines represent children, young adult, and older adult talkers



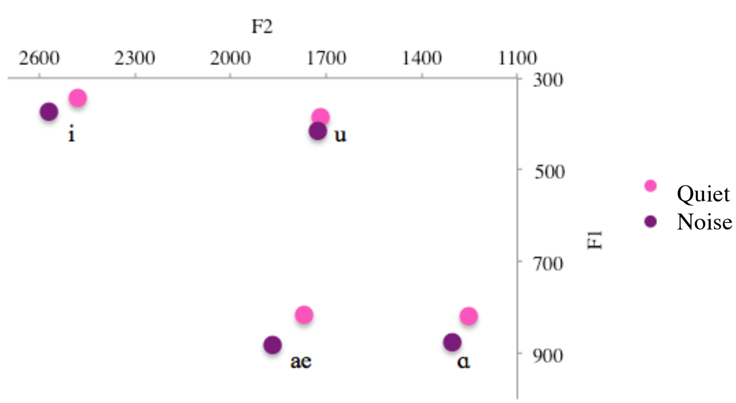
Finally, the 3 main effects were all significantly modulated by vowel type: Vowel*Talker Age Group [F(6,833)=23.605, p<0.001], Vowel*Environment-Oriented Speaking Style [F(3,833)=5.740, p=0.001], and Vowel*Listener-Oriented Speaking Style [F(3,833)=5.305, p=0.001] (see Tables 46 to 52). The effect of Talker Age Group depended on the vowel in that young and older adult talkers had similar F1s for all vowels except /ae/, for which young adults had a significantly higher frequency (i.e. a lower, more open /ae/) than older adults (see Figure 23).

Figure 23: Overall F1 and F2 for the 4 corner vowels for speech produced by children, young adults, and older adults



While Environment-Oriented Speaking Style had a significant effect on F1 for all vowels, it resulted in a larger F1 increase for the low vowels /ɑ, ae/ than for the high vowel /i, u/, making the low vowels more open (see Figure 24).

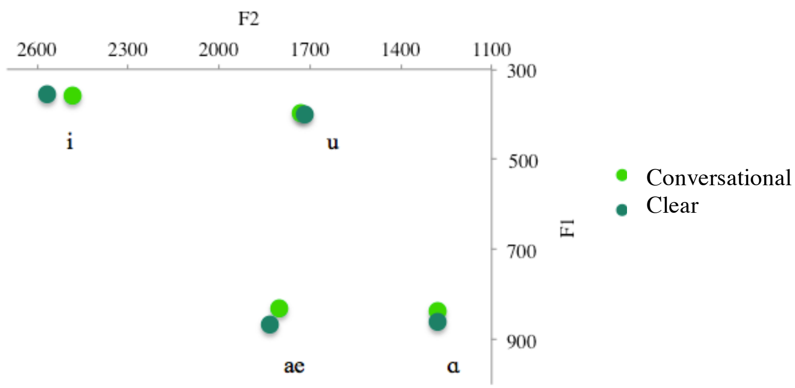
Figure 24: Overall F1 and F2 for the 4 corner vowels for speech produced in quiet and in response to noise



Similarly, Listener-Oriented Speaking Style had a significant effect on F1 frequency for the low vowels /ɑ, ae/ but not the high vowels /i, u/; thus the main effect of

Listener-Oriented Speaking Style arose more from the low vowels than from the high vowels (see Figure 25).

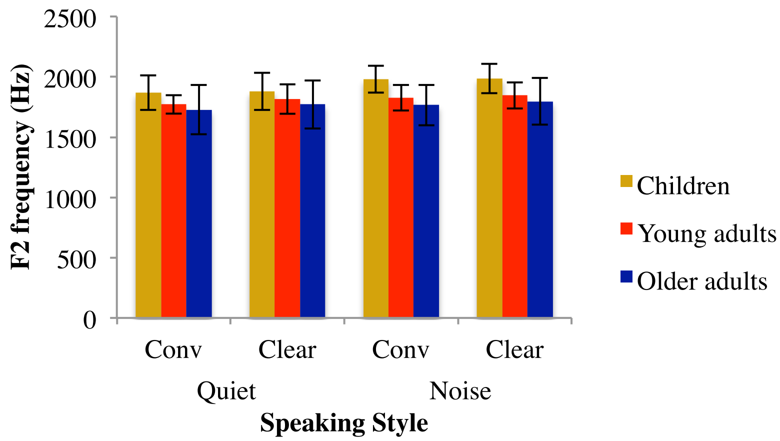
Figure 25: Overall F1 and F2 for the 4 corner vowels for conversational and clear speech



No other interactions were significant: Listener-Oriented Speaking Style*Talker Age Group [F(2,54)=3.313, p=0.733], Environment-Oriented Speaking Style*Listener-Oriented Speaking Style [F(1,860)=0.002, p=0.967], and Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [F(2,860)=0.160, p=0.852], Vowel*Talker Age Group*Environment-Oriented Speaking Style [F(6,833)=0.412, p=0.871], Vowel*Talker Age Group*Listener-Oriented Speaking Style [F(6,833)=0.706, p=0.645], and Vowel*Environment-Oriented Speaking Style*Listener-Oriented Speaking Style [F(3,833)=0.677, p=0.566].

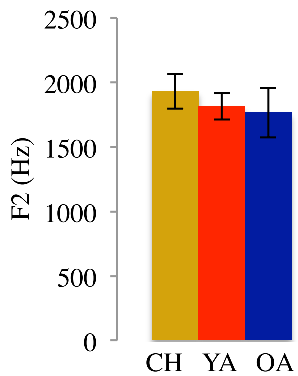
F2

Figure 26: F2 for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



See Figure 26 for an overview of results for the 3 talkers in the 4 speaking styles. As expected, there was a significant main effect of Vowel, given that /ɑ, ae, i, u/ systematically differ in F2 [F(3,4)=26.231, p=0.004]. There was a significant main effect of Talker Age Group [F(2,27)=3.938, p=0.032] (see Figure 27), with children showing significantly higher values than older adults, but only marginally higher values than young adults (see Table 28). There was also a significant main effect of Environment-Oriented Speaking Style [F(1,54)=15.827, p<0.001], with a significant increase in noise-adapted speech compared to speech produced in quiet (see Table 30). That is to say, vowels were fronted in noise-adapted speech compared to speech in quiet. There was no significant main effect of Listener-Oriented Speaking Style [F(1,54)=2.165, p=0.147].

Figure 27: Main effect of Talker Age Group for F2; bars represent children (CH), young adult (YA), and older adult (OA) talkers. Error bars represent standard error.



The main effect of Talker Age Group was also significantly modulated by vowel type [F(6,833)=6.111, p<0.001] (see Tables 46 to 47). The extent of age-related differences in F2 frequency differed by vowel type; there were no age-related differences in F2 for /ae/, while for /u/ children showed a significantly higher F2 than young adults, who showed a significantly higher F2 than older adults. For /ɑ/, children produced a significantly higher F2 than both adult talker groups, and for /i/, children produced a significantly higher F2 than older adults (young adults productions were in between and did not significantly differ from either group) (refer back to Figure 16).

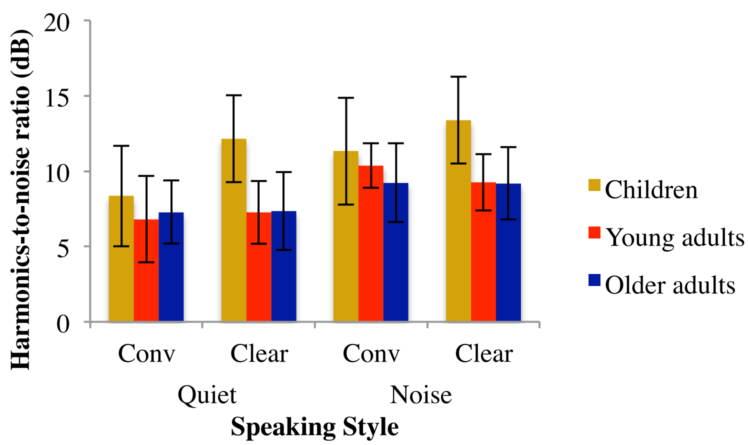
There were no other significant interactions: Environment-Oriented Speaking Style*Talker Age Group [F(2,54)=2.692, p=0.077], Listener-Oriented Speaking Style*Talker Age Group [F(2,54)=0.442, p=0.645], Environment-Oriented Speaking Style*Listener-Oriented Speaking Style [F(1,860)=0.126, p=0.723], Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [F(2,860)=0.072, p=0.931], Vowel*Environment-Oriented Speaking Style [F(3,833)=1.994, p=0.113], Vowel*Listener-Oriented Speaking Style [F(3,833)=1.983, p=0.115], Vowel*Talker Age Group*Environment-Oriented Speaking Style

[F(6,833)=0.232, p=0.966], Vowel*Talker Age Group*Listener-Oriented Speaking Style [F(6,833)=0.597, p=0.733], and Vowel*Environment-Oriented Speaking Style*Listener-Oriented Speaking Style [F(3,833)=0.114, p=0.952].

Voice

Harmonics-to-noise ratio (HNR)

Figure 28: HNR for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



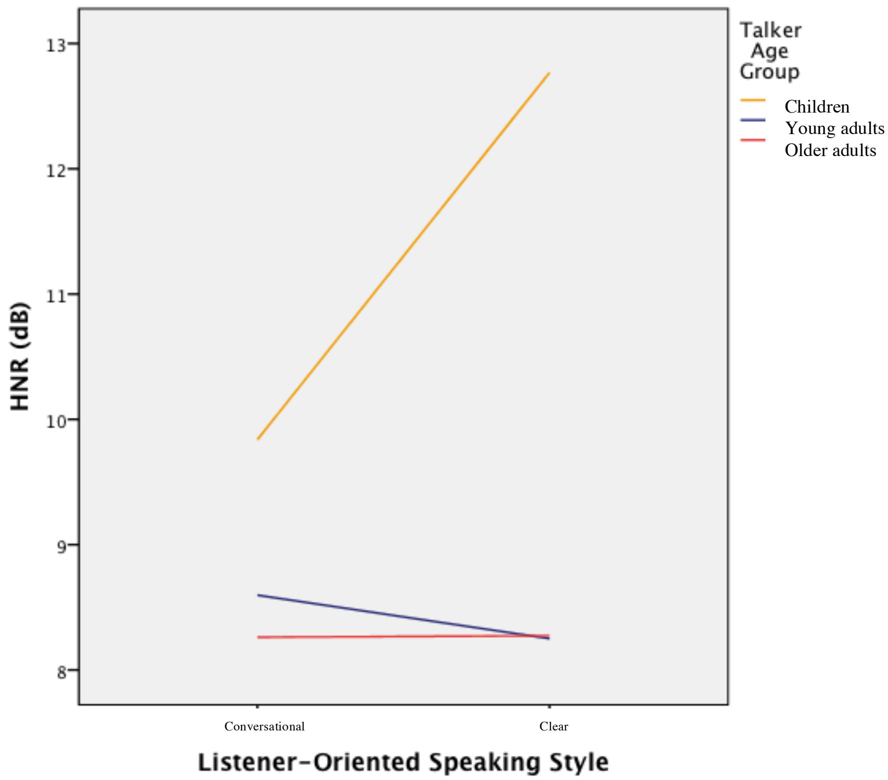
See Figure 28 for an overview of results for the 3 talkers in the 4 speaking styles. There was a significant main effect of Talker Age Group on HNR [F(2,27)=7.375, p=0.003], with children showing a significantly higher HNR than the adult talkers (see Table 28)[3]. There were also significant main effects of both Environment-Oriented Speaking Style [F(1,200)=37.978, p<0.001] and Listener-Oriented Speaking Style [F(1,200)=5.570, p=0.019], with a significant increase in HNR for noise-adapted speech compared to speech produced in quiet, and for clear speech relative to conversational speech (see Tables 30 and 32).

[3] It is important to note that an increase in intensity can alter HNR; however, intensity was not examined in this experiment

Only one interaction was significant: Listener-Oriented Speaking Style and Talker Age Group [$F(2,200)=8.016$, $p<0.001$] (see Figure 29). Pairwise comparisons revealed that only children implemented significant conversational-to-clear speech modifications by increasing HNR. As a result, the child-adult difference in HNR was significant in clear speech but not in conversational speech (see Tables 37 to 38). The main effect of Talker Age Group (with children significantly differing from both adult groups) thus stemmed from clear speech (in which children significantly differ from adults) more so than conversational speech (in which the age groups do not significantly differ).

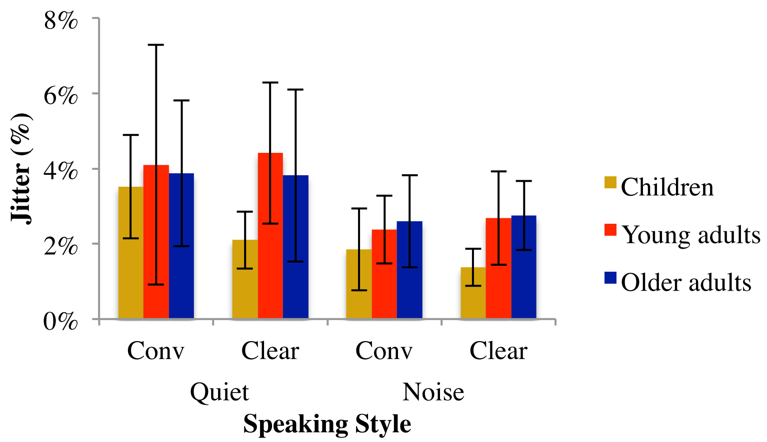
No other interactions were significant: Environment-Oriented Speaking Style*Talker Age Group [$F(2,200)=0.540$, $p=0.584$], Environment-Oriented Speaking Style*Listener-Oriented Speaking Style [$F(1,200)=2.440$, $p=0.120$], and Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [$F(2,200)=0.477$, $p=0.622$].

Figure 29: Interaction between Talker Age Group and Listener-Oriented Speaking Style for HNR; x axis shows conversational and clear speech; lines represent children, young adult, and older adult talkers



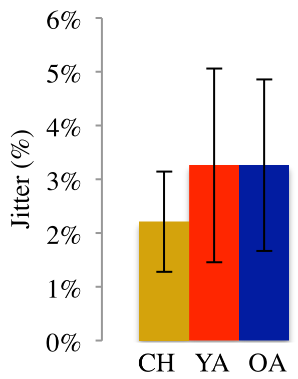
Jitter

Figure 30: Jitter for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



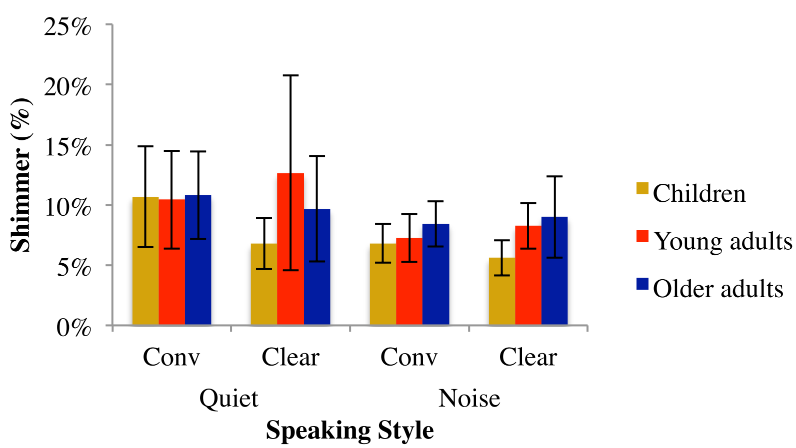
See Figure 30 for an overview of results for the 3 talkers in the 4 speaking styles. There was a significant main effect of Talker Age Group on jitter [F(2,27)=4.323, p=0.024] (see Figure 31); children showed significantly lower jitter than both the adult groups (see Table 28). There was also a significant main effect of Environment-Oriented Speaking Style [F(1,200)=31.591, p<0.001], with a significant jitter decrease in noise-adapted speech compared to speech produced in quiet (see Table 30). There was no significant main effect of Listener-Oriented Speaking Style [F(1,200)=1.225, p=0.270]. There were no significant interactions: Environment-Oriented Speaking Style*Talker Age Group [F(2,200)=0.996, p=0.371], Listener-Oriented Speaking Style*Talker Age Group [F(2,200)=1.660, p=0.193], Environment-Oriented Speaking Style*Listener-Oriented Speaking Style [F(1,200)=1.191, p=0.276], and Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [F(2,200)=0.153, p=0.858].

Figure 31: Main effect of Talker Age Group for jitter; bars represent children (CH), young adult (YA), and older adult (OA) talkers. Error bars represent standard error.



Shimmer

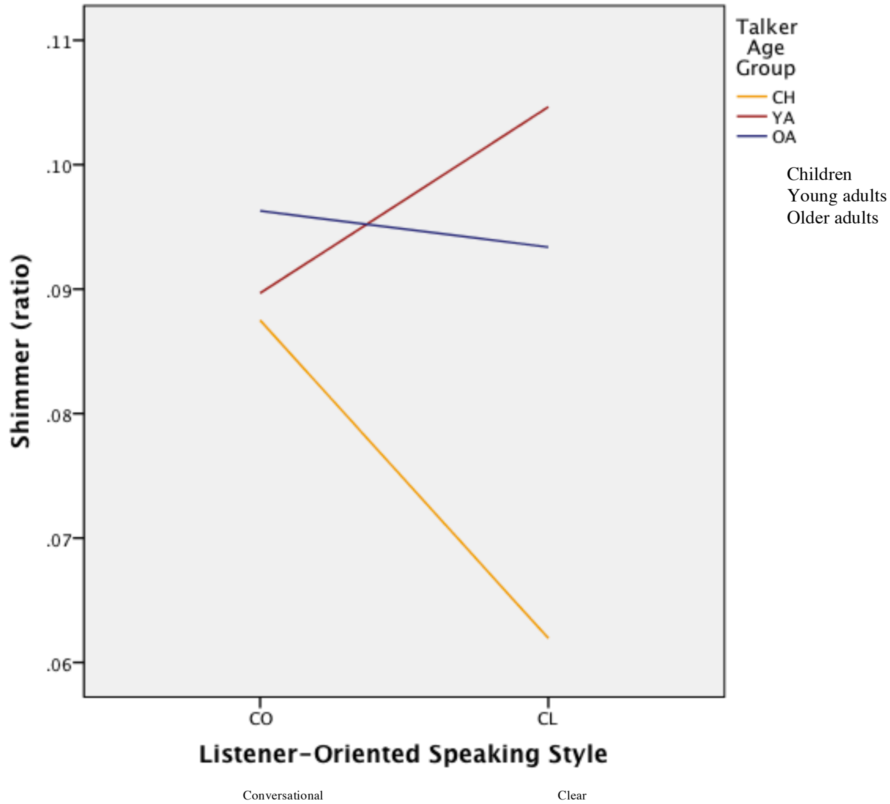
Figure 32: Shimmer for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults. Error bars represent standard error.



See Figure 32 for an overview of results for the 3 talkers in the 4 speaking styles. Results revealed a significant main effect of Environment-Oriented Speaking Style on shimmer [F(1,200)=19.676, p<0.001], with a significant decrease in noise-adapted speech compared to speech produced in quiet (see Table 30). There were no significant main effects of Talker Age Group [F(2,27)=2.727, p=0.083] nor Listener-Oriented Speaking Style [F(1,200)=0.714, p=0.399].

One 2-way interaction was found to be significant: Listener-Oriented Speaking Style and Talker Age Group [F(2,200)=3.425, p=0.035] (see Figure 33). Pairwise comparisons revealed that only children implemented significant clear speech modifications. As a result, the child-adult difference was only significant in clear speech (see Tables 37 to 38). No other interactions were significant: Environment-Oriented Speaking Style*Talker Age Group [F(2,200)=1.310, p=0.272], Environment-Oriented Speaking Style*Listener-Oriented Speaking Style [F(1,200)=1.095, p=0.297], and Environment-Oriented Speaking Style*Listener-Oriented Speaking Style*Talker Age Group [F(2,200)=0.695, p=0.500].

Figure 33: Interaction between Talker Age Group and Listener-Oriented Speaking Style for shimmer; x axis shows conversational and clear speech; lines represent children,



DISCUSSION

The overall goal of this experiment was to identify the acoustic-phonetic features that characterize the production of noise-adapted and clear speech speaking style adaptations in children, young adults, and older adults. 11 different acoustic features were examined: 5 global features (F0 range and mean, 1-3 kHz energy, speaking rate, and pause duration), 3 segmental features (vowel duration, F1, and F2), and 3 voice quality features (harmonics-to-noise ratio, jitter, and shimmer).

The results showed a large number of age-related differences in the implementation of noise-adapted and clear speaking style adaptations. A discussion of the various effects and interactions is presented below.

The production of noise-adapted speech across the lifespan

Several of the acoustic-phonetic changes that talker groups implemented when producing noise-adapted speech were consistent across age. For instance, all talker groups increased F0 range and decreased shimmer. Although children spoke with more energy in the 1-3 kHz region compared to older adults, all talker groups increased 1-3 kHz energy in response to noise. And while children and young adults spoke faster (with shorter vowels) than older adults, all talker groups slowed their speaking rate and lengthened their vowels in response to noise. Children showed a lower jitter and higher F2 than the adult talkers, but all talker groups lowered jitter and increased F2 when adapting their speech to noise. All talker groups increased HNR in response to noise as well. The results thus reveal that despite some general age-related differences in speech (in terms of 1-3 kHz energy, speaking rate, vowel duration, F2, jitter), speaking in response to noise elicited common strategies across talker groups.

However, differences between the 3 talker groups also emerged in the production of noise-adapted speech. Children, who although speaking with a significantly higher mean F0 and F1 than adults in quiet, made larger noise-induced F0 and F1 increases than adults in response to noise.

As these 3 talker groups have never been directly compared in terms of speaking style adaptations, there were no specific hypotheses regarding the interactions between the talker groups and the production of noise-adapted speech. All interactions highlighted child-adult differences; young and older adults made similar acoustic-phonetic modifications when speaking in response to noise. For instance, age-related differences in the production of F1 became more apparent in noise-adapted speech compared to speech in quiet, due to children's relatively larger response to noise.

That older adult speech was overall characterized by less energy in the 1-3 kHz range and a slower speaking rate (including longer vowels) was in accord with the hypotheses and previous research (Linville and Rens, 2001; Smiljanic, 2013). The overall child-adult differences in F0, F1, and F2 (higher mean F0 and higher formant frequencies for children compared to adults) were also expected. These differences were attributed to physiological differences (smaller vocal tracts and oral cavities in children compared to adults).

Results also revealed an absence of young adult-older adult differences in voice quality and formant frequencies. While children produced lower jitter and higher formant frequencies compared to the adult talkers, there were no significant differences between younger and older adults. As mentioned in the introduction, aging affects laryngeal and respiratory mechanisms (e.g. muscle atrophy, ossification of cartilage, reduction in breath support) that can result in age-related speech production differences (e.g. decreased HNR due to additive noise arising from inadequate closure of vocal folds/instability in vocal fold vibration) (Ferrand, 2002). The lack of younger-older adult differences found in the current study can likely be attributed to the good physiological condition of the older adult talkers in this study—these were all older adults active enough to volunteer for on-campus research, and several commented on their regular physical activities. Research has shown that age-related changes to speech are more due to physiological age than chronological (Ramig and Ringel, 1983; Ringel and Chodzko-Zajko, 1987).

For many of the acoustic features, there were no specific hypotheses given the paucity of previous findings. For instance, it was uncertain to what extent mean F0 would differ between younger and older adults; here mean F0 did not differ between the 2 groups (again, perhaps due to the older adults' excellent physical condition). It was also uncertain to what extent the durational measures (speaking rate, pause duration, vowel duration) would differ between children and young adults; here there were no differences. No specific hypotheses regarding the effects of age on F0 range were posited, and no effects of age on F0 range were found. The overall child-adult differences in jitter were novel, as this pattern has not been established in previous research. Although previous research has shown jitter does not change with age for children 7 to 15 years old (Linders et al., 1995), there is no work comparing jitter in 11-13 year olds and adults. This study contributes novel findings on how age affects several acoustic features that have been previously unexamined.

The production of clear speech across the lifespan

The 3 talker groups showed several similarities in the production of conversational-to-clear speech modifications. All talker groups produced clear speech with a wider F0 range and longer pause durations (although older adults lengthened pauses to a greater extent than the younger groups). Although the children and young adult groups spoke faster (with shorter vowels) than older adults, all talkers slowed their speaking rate and increased their vowel durations in clear speech. And while children showed a higher F1 than adults overall, all talker groups raised F1 when producing conversational-to-clear speech adaptations. This was specifically seen for vowels /ɑ, ae/ but not /i, u/, reflecting a lower jaw/tongue articulation for the low vowels.

The 3 groups also diverged in several of the acoustic-phonetic modifications used to produce clear speech. Children produced significant conversational-to-clear increases in HNR and decreases in shimmer, unlike the adult talkers. This led to children showing an overall higher HNR compared to the adult groups. Previous work has shown a lower HNR in 4 to 10 year-old children compared to adults (Ferrand, 2000). This higher HNR in 11 to 13 year-old children may be an acoustic parameter overshoot, as documented in Lee et al. (1999).

Similar to what was found for environment-oriented speaking style adaptations, the conversational-to-clear speech changes increased child-adult differences (specifically, for HNR and shimmer). The changes also increase differences between older adults and the younger talker groups (specifically, for pause duration).

The finding that older adults had an overall slower speaking rate is in line with Smiljanic (2013). However, unlike Smiljanic (2013) in which the older adult clear speech decrease to rate was smaller than that of young adults, young and older adults in the current study made similar decreases to speaking rate when producing clear speech. That both children and young adults decreased speaking rate in clear speech (although children did marginally less) is in line with Pettinato and Hazan (2013), although the children here did not show an overall slower speaking rate. This is likely due to the fact that the children in this study were 11-13 years old, while the children in Pettinato and Hazan (2013) were 9-10 years old (their 13-14 year old adolescents showed no difference from adults in speaking rate).

That children significantly modified HNR and shimmer when producing clear speaking style adaptations is a novel finding. This is similar to Glaze et al. (1990), who found that children increased HNR and decreased shimmer when asked to speak loudly. Although vocal levels were not included in this study, it appears that children's conversational-to-clear speech modifications reduce the amount of noise in their voice quality (harmonics, amplitude) to a greater extent than that of adults.

Noise-adapted vs. clear speech: similarities, differences, and interactions

This study is the first to directly compare both types of intelligibility-enhancing speaking style adaptations in multiple talkers. Results revealed a large overlapping set of cross-style modifications aimed to enhance intelligibility: a wider F0 range, increased F1 (which was shown to expand vowel space, in line with e.g. Smiljanic and Bradlow, 2005), a slower speaking rate along with longer vowels, as well as a higher HNR. But the different communicative intents resulted in slight acoustic-phonetic differences as well: for instance, the set of vowels for which F1 was raised differed between environment- and listener-oriented speaking styles. In noise-adapted speech (compared to speech in quiet), all 4 corner vowels showed raised F1, especially the low vowels; in clear speech (compared to conversational speech), only the low vowels showed raised F1. Likewise, while there was a significant main effect of both environment- and listener-oriented speaking style on HNR, the only clear speech increases to HNR were made by children (whereas noise-adapted increases were made by all 3 talker groups).

The two lines of speaking style adaptations differed in other acoustic parameters as well; increases to mean F0 and 1-3 kHz energy were more prevalent for noise-adapted speech than for clear speech adaptations (e.g. a significant conversational-to-clear increase in 1-3 kHz energy was found for speech in quiet, but not for speech in noise, whereas a significant quiet-to-noise increase in 1-3 kHz energy was found for both conversational and clear speech). Noise-adapted speech also showed decreases to jitter and shimmer not found for clear speech; while children decreased shimmer for clear speech modifications, there was no significant main effect of Listener-Oriented Speaking Style across all talker groups. On the other hand, longer pauses were found for clear speech modifications, but not for noise-adapted speech modifications.

As there is little work directly comparing clear speech and noise-adapted speech, it was unknown to what extent the adaptations would cumulate. The findings here suggest that the combined effect of clear and noise-adapted speech is dependent on the acoustic feature. This is likely due to the fact that each speaking style adaptation is comprised of a distinct set of acoustic-phonetic modifications; the extent of their interaction differs since the two lines of communicative intent enhance two different sets of cues (some of which overlap, some of which do not). For some features (F0 mean, 1-3 kHz energy, speaking rate), the speaking style increases each had a greater effect individually than in conjunction. In other words, the level at which a talker was able to enhance a contrast appeared to be capped. For other features (F0 range, pause duration), the speaking style adaptations were enhanced to an even greater extent when combined (i.e. clear speech adapted to noise). For many other features (HNR, jitter, shimmer, F1, F2), the two lines of speaking style adaptations did not interact; the magnitude of the clear speech adaptation was independent of the environment-oriented speaking style, and likewise, the magnitude of the noise-adapted speech adaptation was independent of the listener-oriented speaking style.

A few of the interactions mentioned above were also heavily modulated by talker age group (for mean F0, pause duration, and 1-3 kHz energy). For instance, while all talker groups increased mean F0 in response to noise, older adult talkers made less of a noise-adapted increase in clear speech than in conversational speech. There were no such differences for the younger talker groups. Likewise, only older adults' clear speech showed a noise-adapted lengthening of pause duration. There were no noise-adapted modifications to pause duration in conversational speech, nor for the younger talkers. Finally, only children's and young adults' speech in quiet exhibited clear speech modifications to 1-3 kHz energy. There were no clear speech modifications in noise-adapted speech, nor in the speech of older adult talkers. This indicates that the main effect of clear speech on 1-3 kHz energy which has been found in several previous studies (e.g. Krause and Braida, 1995) may not hold across all talker groups or environment-oriented speaking styles. In sum, the extent to which noise-adapted and clear speech adaptations are cumulative seems to not only depend on the acoustic feature, but also on the talker's age.

The greater number of significant acoustic-modifications elicited by quiet-to-noise speaking style adaptations compared to conversational-to-clear speaking style adaptations was likely due to the nature of the two elicitation methods. The elicitation method for noise-adapted speech involved a physically present communicative difficulty (noise presented over headphones, perhaps eliciting a more automatic adaptation), while that for clear speech asked talkers to imagine a communicative difficulty (directing speech towards an imaginary listener with low proficiency, perhaps a more intentional adaptation). While research has also shown this method of eliciting clear speech in young adults can result in even more exaggerated acoustic changes than that of more spontaneously-elicited clear speech, e.g. diapix tasks (Hazan and Baker, 2011), it is unknown to what extent children talkers respond in the same manner.

Many of the findings were in line with the hypotheses/previous work: noise-adapted speech exhibited a slower speaking rate (including longer vowels), increased 1-3 kHz energy and mean F0, along with a wider F0 range, while clear speech showed a slower speaking rate (including longer vowels and pauses), along with a wider F0 range (Pichora-Fuller et al., 2010). In addition to clear speech, noise-adapted speech also showed an expansion to vowel space area. The findings that mean F0, 1-3 kHz range, and speaking rate all increased in response to noise, along with the findings that F0 range increases and speaking rate decreases in clear speech, replicates previous work in our lab

(Gilbert et al., 2014). While we did not find (as hypothesized) increased pause duration in noise-adapted speech, nor mean F0/1-3 kHz energy increases in clear speech, the set of acoustic-phonetic modifications that characterizes noise-adapted and clear speech modifications tends to vary significantly from one study to another (Pichora-Fuller et al., 2010).

Additional new findings include the significant change of voice quality measures in response to speaking in noise (increased HNR, decreased jitter and shimmer) and clear speech (increased HNR). This was likely due to the increase in vocal effort required to compensate for talking over noise; previous research has shown that children speaking "as loud as they can" increased HNR and reduced jitter and shimmer (Glaze et al., 1990). However, this has not been looked at for these specific speaking style adaptations or in all 3 talker groups. That voice quality measures were affected by speaking style modifications is a novel finding in speech adaptation research.

In conclusion, the findings indicate that aging affected speech production in a number of ways. In particular, children and older adults implemented differential acoustic-phonetic modifications compared to young adults when producing noise-adapted and clear speaking style adaptations (e.g. in clear speech, older adults altered pause duration while children altered HNR). A direct comparison of the two speaking style adaptations revealed a large overlapping set of cross-style modifications aimed to enhance intelligibility, but differences remained between the two lines of communicative intent (e.g. changes to jitter were more prevalent in quiet-to-noise speech modifications; changes to pause duration were more typical in conversational-to-clear speech modifications). The two types of speaking style adaptations cumulated in an interactive manner for several acoustic features (e.g. greater than the sum of their parts for F0 range, less than the sum of their parts for speaking rate), but not for others (e.g. equal to the sum of their parts for jitter). These results contribute to a better understanding of how age-related peripheral and cognitive changes relate to speech production mechanisms across the lifespan.

Experiment 2: Perception

Experiment 2 was composed of two perception tasks: Experiment 2A assessed speech intelligibility via a word recognition in noise task, while Experiment 2B investigated the perceptual age of the talkers via an age estimate task. The two experiments are presented below.

EXPERIMENT 2A: WORD RECOGNITION IN NOISE

Research Aims

The goal of Experiment 2A was to assess the perceptual impact of the acoustic-articulatory adaptations reported in Experiment 1; specifically, to examine the extent to which acoustic changes that characterize noise-adapted and clear speaking style adaptations produced by different talker age groups improve word recognition in noise. The perceptual impacts were examined both as a function of age (children, young adults, and older adults) and communicative intent (i.e. environment- and listener-oriented speaking styles).

Methods

Materials

A subset of the stimuli from Experiment 1 were selected: the initial 40 of the 60 sentences as produced by all 30 talkers in all four speaking styles (4800 sentences total). This sentence subset was chosen so that the same listener could take part in both Experiments 2A and 2B (perceived age task) while only being exposed to each sentence once (i.e. a listener did Experiment 2A on sentences 1-40, and then Experiment 2B on sentences 41-60). All sentences were leveled for RMS amplitude and mixed with the speech-shaped noise that was used to elicit noise-adapted speech in Experiment 1[4]. Using the same noise in both the recording sessions and in the assessment of the intelligibility benefit of noise-adapted speech ensured a correlation between the target speech and its noise masker. Using SSN allowed a consistent level of masking across keywords (i.e. no glimpsing windows). Based on previous work and piloting, the signal-to-noise ratio (SNR) was set at -5 dB to avoid ceiling and floor results, as the stimuli were expected to range greatly in intelligibility (conversational speech in quiet vs. noise-adapted clear speech for a range of 30 talkers).

Listeners

A total of 61 native monolingual speakers of English (18-39 years old, mean 20.2 years) participated in Experiment 2A. They were all University of Texas at Austin students recruited from the Linguistics department subject pool. All listeners were screened for normal hearing (thresholds below 25 dB SPL at .5, 1, 2, and 4 kHz). Participants provided written informed consent and received course credit.

Procedure

Listeners took part in 1 of 3 conditions. Each condition consisted of 40 sentences from 10 talkers in 1 age group (children, young adult, or older adult talkers). Each talker's speech intelligibility was thus assessed by 20-21 listeners in total. The test began with 5 practice sentences to familiarize the listener with the task; practice sentences consisted of talkers and stimuli not included in the set of test stimuli. Test sentences were then presented to the listener, distributed as follows: 4 sentences from each of 10 speakers, 1 in each speaking style (QCO, QCL, NCO, NCL). Sentence presentation order

[4] RMS was leveled for the entire sentence and not just for the target word to provide a more naturalistic stimulus; additionally, given the target words were embedded in high context sentences, the test does not strictly rely on target word recognition in isolation

and talker-style-sentence pairing were randomized for each listener. The experiment was presented in MATLAB. Listeners were instructed to write what they hear, typing 1 sentence at a time on the keyboard after stimulus presentation. The test was self-paced. Each sentence contained one target word (the final word in the sentence) and was scored as correct (1) or incorrect (0).

Statistical Analysis

To examine the extent to which noise-adapted and clear speech produced by the 3 talker groups provided an intelligibility benefit, results were analyzed with mixed effects logistic regressions using the lme4 package in R. Keyword identification (i.e., correct or incorrect) was the dichotomous dependent variable. Talker, Sentence, and Listener were included in the model as random factors and Talker Age Group (children, young adults, or older adults), Environment-Oriented Speaking Style (produced in quiet or in response to noise), Listener-Oriented Speaking Style (conversational or clear), and their interactions were included as fixed effects. Random slopes were included in the model for both Environment- and Listener-Oriented Speaking Style at the level of Talker, since this level showed the greatest variance. This determined the impact of age and communicative intent on word recognition.

Hypotheses

With regard to overall intelligibility of the 3 talker groups, previous work found children and adults[5] to have equivalent inherent intelligibility (Hazan and Markham, 2004) while older adults were significantly less intelligible than young adults (Smiljanic, 2013). However, given this is only a relatively small number of studies that have

[5] mean age: 29.9 years, s.d. 7.10

examined speech intelligibility in children and older adults, and given the difference in the speech materials used here vs. in Smiljanic (2013) (meaningful vs. semantically anomalous sentences), I predicted that both children's and older adults' baseline speech (conversational and quiet speech) should be comparably intelligible to young adults.

With regard to the ability to produce intelligibility enhancing speaking style modifications, previous findings for older adults and children are inconclusive. Smiljanic (2013) found that the older adults provided a smaller intelligibility benefit via clear speech compared to young adults, while Schum (1996) found that older adults were able to produce an equivalent clear speech gain to that of young adults. Given the conflicting findings, and the lack of findings on the intelligibility of noise-adapted speech for these talkers, I had no specific predictions regarding older adult noise-adapted and clear speech intelligibility benefits. Given recent work suggesting that certain speech clarification strategies (e.g. vowel space expansion) still continue to develop even into late adolescence, I suspected that children will not be able to enhance their intelligibility to the same degree as the adult talker groups (Pettinato and Hazan, 2013).

I hypothesized that clear and noise-adapted speech would be more intelligible than conversational speech and quiet speech mixed with noise. I also hypothesized that the noise-adapted speech intelligibility benefit would be larger than that of clear speech, given that the noise-adapted speech was originally produced in response to the noise masker used in word recognition (and was thus specifically produced to overcome the masker). Based on a previous study that examined the 2 lines of speaking style adaptations in conjunction (Gilbert et al., 2014), I predicted that noise-adapted clear speech (NCL) might be of even greater intelligibility compared to the benefits of clear and noise-adapted speech separately.

Results

The ratio of words correctly transcribed by listeners for each talker group in each style ranged from 0.13 to 0.74, with a mean of 0.41. Figure 34 shows the breakdown of word recognition by talker group and speaking style. Random effects are summarized in Table 5, and fixed effects in Table 6.

Figure 34: Experiment 2A results for conversational and clear speech produced in quiet and in response to noise by children, young adults, and older adults

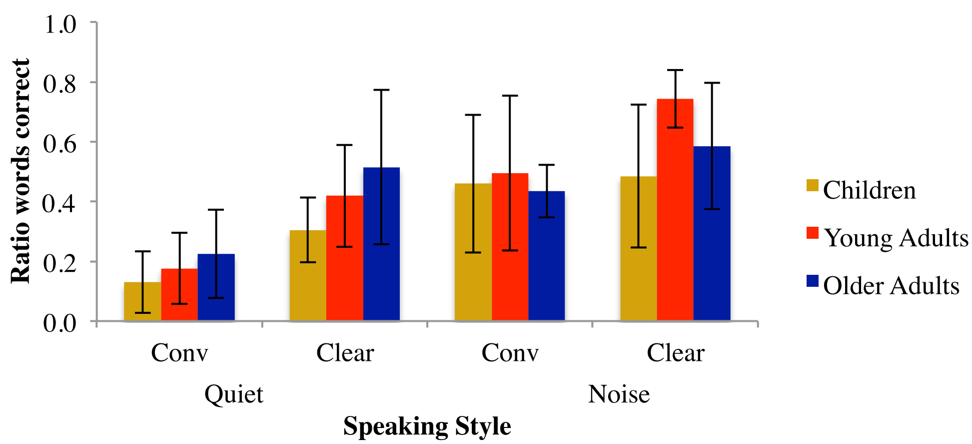


Table 5: Experiment 2A summary of random effects for the mixed model

Parameter		Variance	Std. Deviation
Sentence	Intercept	0.870	0.933
Listener	Intercept	0.020	0.143
Talker	Intercept	0.766	0.875
	QN	0.551	0.742
	COCL	0.394	0.628

Table 6: Experiment 2A summary of fixed effects for the mixed model. Significant effects are highlighted in yellow.

Source	Numerator df	Denominator df	F	Sig.
QN	1	2428	163.840	0.000
COCL	1	2428	101.838	0.000
CHYAOA	2	2428	13.964	0.000
QN*COCL	1	2428	6.201	0.013
QN*CHYAOA	2	2428	8.003	0.000
COCL*CHYAOA	2	2428	5.682	0.004
QN*COCL*CHYAOA	2	2428	1.804	0.165

The results of the mixed-effects logistic regression revealed that the probability of correct keyword identification was significantly affected by all 3 fixed effects: Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style. There were also significant 2-way interactions between all 3 fixed effects. The 3-way interaction was not significant [$F(2,2428)=1.804$, $p=0.165$].

With regard to the main effect of Talker Age Group [$F(2,2428)=13.964$, $p<0.001$], results showed that children were significantly less intelligible than young adults ($p=0.004$), but not older adults ($p=0.234$). Young and older adults did not significantly differ in intelligibility ($p=0.085$) (see Table 7 for pairwise comparisons, and Figure 35 for word recognition by talker group collapsed across speaking styles). Speaking in response to noise was significantly more intelligible than producing speech in quiet [$F(2,2428)=163.840$, $p<0.001$]. Likewise, clear speech was overall significantly more intelligible than conversational speech [$F(2,2428)=101.838$, $p<0.001$].

Figure 35: Main effect of Talker Age Group for word recognition in noise; boxes represent children, young adult, and older adult talkers

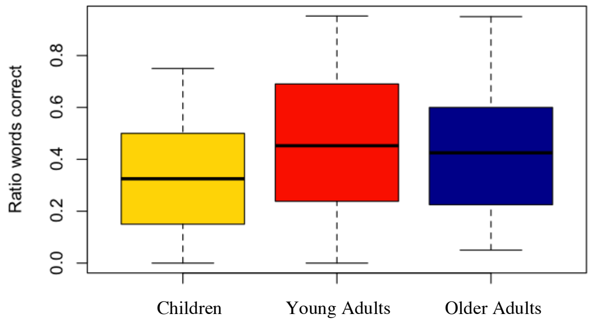


Table 7: Experiment 2A pairwise comparisons for the main effect of Talker Age Group. Significant effects are highlighted in yellow.

(I) CHYAOA	(J) CHYAOA	Estimate	Std. Error	z value	Sig.
CH	OA	0.546	0.458	1.190	0.234
	YA	1.342	0.463	2.899	0.004
OA	CH	0.546	0.458	1.190	0.234
	YA	0.796	0.462	1.723	0.085
YA	CH	1.342	0.463	2.899	0.004
	OA	0.796	0.462	1.723	0.085

The ability to enhance intelligibility via noise-adapted and clear speaking style adaptations was also examined in terms of proportional and net gain. Proportional gain was calculated as follows: proportional gain=(intelligibility in enhanced style – intelligibility at baseline)/intelligibility at baseline. Net gain was calculated as the intelligibility in the enhanced style – intelligibility at baseline[6]. Both these calculations were used to provide insight into the magnitude of the noise-adapted and clear speech intelligibility gains. Noise-adapted speech showed a proportional gain of 81% and a net gain of 24%. That is to say, noise-adapted speech was 81% more intelligible than speech produced in quiet, and improved speech intelligibility by 24%. Clear speech showed a

[6] Gain ratios were multiplied by 100 to convert into percentages

proportional gain of 59% and a net gain of 19%, i.e. clear speech was 59% more intelligible than conversational speech, and improved speech intelligibility by 19%.

The significant interaction between Talker Age Group and Environment-Oriented Speaking Style [$F(2,2428)=8.003$, $p<0.001$] revealed that all talker groups were significantly more intelligible when speaking in response to noise compared to speaking in quiet ($p<0.001$ for all groups; see Table 8). However, the extent to which the talker age groups differed in intelligibility varied by speaking style (see Figure 36 and Table 9). For speech produced in quiet, children were significantly less intelligible than older adults ($p=0.020$). Young adults did not significantly differ from either talker group. For noise-adapted speech, children were significantly less intelligible than young adults ($p=0.003$). Older adults did not significantly differ from either group. In sum, while children were significantly less intelligible than adults when speaking in quiet as well as in response to noise, the adult population they differed from depended on the communicative environment.

Figure 36: Interaction between Talker Age Group and Environment-Oriented Speaking Style for word recognition in noise; boxes represent children, young adult, and older adult talkers producing speech in quiet and in noise

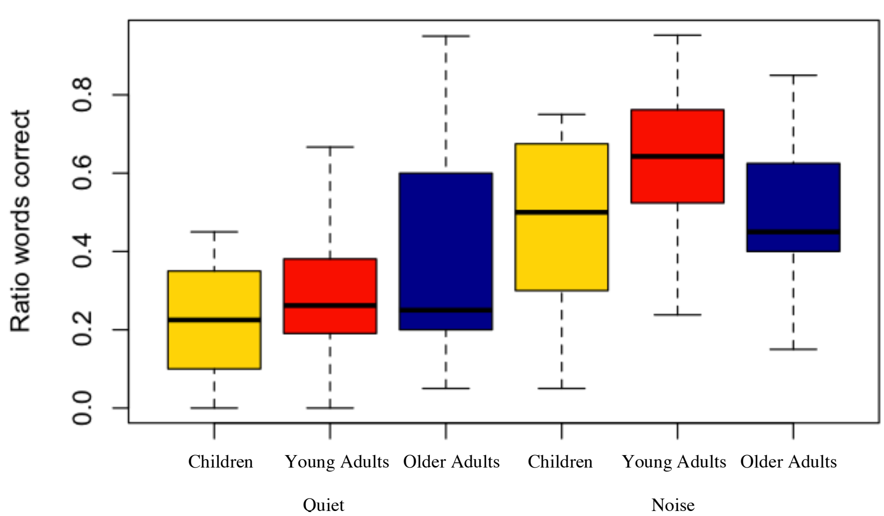


Table 8: Experiment 2A pairwise comparisons for the interaction between Talker Age Group and Environment-Oriented Speaking Style. Significant effects are highlighted in yellow.

CHYAOA	(I) QN	(J) QN	Df	F	Sig.
CH	N	Q	1	62.873	0.000
	Q	N	1	62.873	0.000
OA	N	Q	1	17.094	0.000
	Q	N	1	17.094	0.000
YA	N	Q	1	104.114	0.000
	Q	N	1	104.114	0.000

Table 9: Experiment 2A pairwise comparisons for the interaction between Talker Age Group and Environment-Oriented Speaking Style. Significant effects are highlighted in yellow.

QN	(I) CHYAOA	(J) CHYAOA	Estimate	Std. Error	z value	Sig.
N	CH	OA	0.510	0.430	1.186	0.236
		YA	1.289	0.435	2.966	0.003
	OA	CH	0.510	0.430	1.186	0.236
		YA	0.779	0.433	1.800	0.072
	YA	CH	1.289	0.435	2.966	0.003
		OA	0.779	0.433	1.800	0.072
Q	CH	OA	1.112	0.477	2.329	0.020
		YA	0.596	0.475	1.254	0.210
	OA	CH	-1.112	0.477	-2.329	0.020
		YA	-0.516	0.473	-1.092	0.275
	YA	CH	0.596	0.475	1.254	0.210
		OA	-0.516	0.473	-1.092	0.275

The significant interaction of Talker Age Group by Listener-Oriented Speaking Style revealed that all talker groups significantly enhanced their intelligibility through conversational-to-clear speaking style adaptations [$F(2,2428)=5.682$, $p=0.004$] (see Figure 37), although the clear speech enhancement for children was smaller compared to those for young and older adults (see Table 10). Results also revealed that the magnitude of talker age group differences depended on speaking style (Table 11). In conversational speech, there were no significant differences between talker groups. In clear speech, however, young adults were significantly more intelligible than children (p=0.004). Older adults did not significantly differ from either group. In sum, age-related differences in speech intelligibility only appeared in clear (but not in conversational) speech.

Figure 37: Interaction between Talker Age Group and Listener-Oriented Speaking Style for word recognition in noise; boxes represent children, young adult, and older adult talkers producing conversational and clear speech

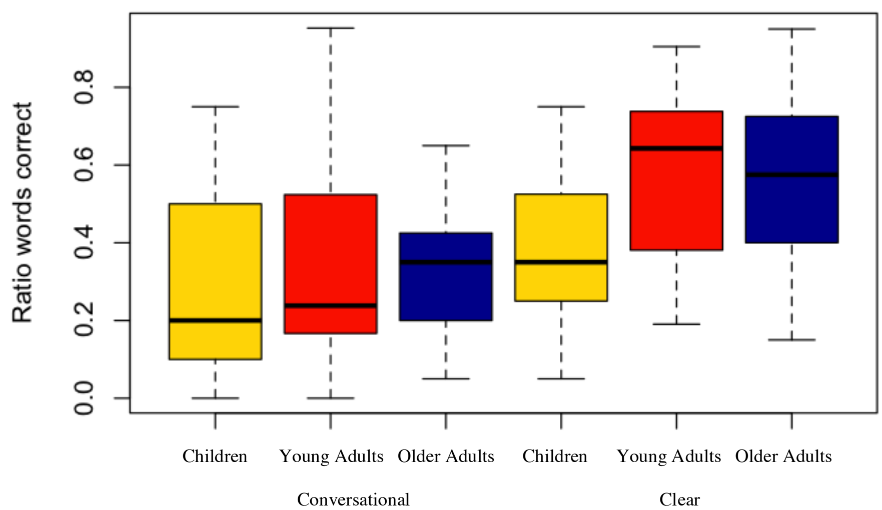


Table 10: Experiment 2A pairwise comparisons for the interaction between Talker Age Group and Listener-Oriented Speaking Style. Significant effects are highlighted in yellow.

CHYAOA	(I) COCL	(J) COCL	Df	F	Sig.
CH	CL	CO	1	9.669	0.002
	CO	CL	1	9.669	0.002
OA	CL	CO	1	42.211	0.000
	CO	CL	1	42.211	0.000
YA	CL	CO	1	60.606	0.000
	CO	CL	1	60.606	0.000

Table 11: Experiment 2A pairwise comparisons for the interaction between Talker Age Group and Listener-Oriented Speaking Style. Significant effects are highlighted in yellow.

COCL	(I) CHYAOA	(J) CHYAOA	Estimate	Std. Error	z value	Sig.
CL	CH	OA	0.521	0.448	1.163	0.245
		YA	1.318	0.452	2.915	0.004
	OA	CH	0.521	0.448	1.163	0.245
		YA	0.798	0.452	1.764	0.078
	YA	CH	1.318	0.452	2.915	0.004
		OA	0.798	0.452	1.764	0.078
CO	CH	OA	-0.100	0.480	-0.209	0.835
		YA	0.203	0.483	0.420	0.675
	OA	CH	-0.100	0.480	-0.209	0.835
		YA	0.303	0.480	0.632	0.528
	YA	CH	0.203	0.483	0.420	0.675
		OA	0.303	0.480	0.632	0.528

While both noise-adapted and clear speech significantly enhanced intelligibility ($p<0.001$), there was a significant interaction between Environment- and Listener-Oriented Speaking Style [$F(2,2428)=6.201$, $p=0.013$]. The magnitude of the intelligibility benefit depended on whether the enhancements (noise-adapted speech or clear speech) were produced individually or in conjunction (see Figure 38). Specifically, the conversational-to-clear speech intelligibility gain was larger for speech produced in quiet (proportional gain of 133%, net gain of 24%) compared to speech produced in noise (proportional gain of 30%, net gain of 14%). That is to say, clear speech in quiet was 133% more intelligible than conversational speech in quiet (improving intelligibility by 24%), while clear speech in noise was 30% more intelligible than conversational speech in noise (improving intelligibility by 14%).

Similarly, the quiet-to-noise intelligibility gain was larger for conversational speech (proportional gain of 162%, net gain of 29%) compared to clear speech (proportional gain of 46%, net gain of 19%). That is to say, noise-adapted conversational speech was 162% more intelligible than conversational speech in quiet (improving intelligibility by 29%), while noise-adapted clear speech was 46% more intelligible than clear speech in quiet (improving intelligibility by 19%). While noise-adapted clear speech was the most intelligible speaking style, it was less intelligible than if the individual noise-adapted and clear speech enhancements were summed together (see Tables 12 and 13 for pairwise comparisons).

Figure 38: Interaction between Environment- and Listener-Oriented Speaking Style for word recognition in noise; boxes represent conversational and clear speech produced in quiet and in noise

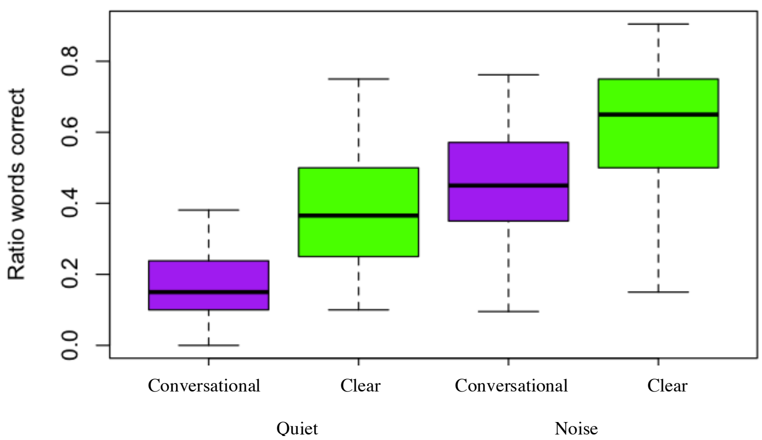


Table 12: Experiment 2A pairwise comparisons for the interaction between Environment- and Listener-Oriented Speaking Styles. Significant effects are highlighted in yellow.

QN	(I) COCL	(J) COCL	Df	F	Sig.
N	CL	CO	1	25.976	0.000
	CO	CL	1	25.976	0.000
Q	CL	CO	1	89.151	0.000
	CO	CL	1	89.151	0.000

Table 13: Experiment 2A pairwise comparisons for the interaction between Environment- and Listener-Oriented Speaking Styles. Significant effects are highlighted in yellow.

COCL	(I) QN	(J) QN	Df	F	Sig.
CL	N	Q	1	49.116	0.000
	Q	N	1	49.116	0.000
CO	N	Q	1	127.345	0.000
	Q	N	1	127.345	0.000

Discussion

The findings showed both age- and adaptation-related effects on intelligibility. Children were overall less intelligible than the young adult talkers, although interactions revealed that this originated from children being less intelligible than young adults in noise-adapted and clear speech only. For speech produced in quiet, children were less intelligible than older adults, and for conversational speech, there were no age-related differences. Young adults and older adults did not significantly differ in intelligibility in any speaking style. Noise-adapted and clear speech both significantly enhanced intelligibility. While noise-adapted clear speech was the most intelligible speaking style, the intelligibility gain was smaller than what would be expected if the individual noise-adapted and clear speech enhancements were additive (e.g. the clear speech increase was larger for speech produced in quiet than for speech produced in response to noise).

The lack of age-related differences in conversational speech is in line with the hypotheses, replicating the findings of Hazan and Markham (2004), who found that 13 year-old children's conversational speech was as intelligible as (young and middle-aged) adult men. Also in line with the hypotheses, all talker groups successfully increased their intelligibility in response to environment- and listener-oriented difficulty, but the extent of the intelligibility benefit varied for the 3 talker groups. Compared to children, young adult talkers produced the most intelligible noise-adapted and clear speech, while older adults produced the most intelligible quiet speech.

The intelligibility of older adults did not significantly differ from that of young adults in any speaking style. This finding differs from Smiljanic (2013), in which older adults were overall less intelligible than young adults and exhibited a smaller clear speech intelligibility gain, but is in accord with Schum (1996), who found older adults able to produce a clear speech intelligibility benefit comparable to that of young adults (though no baseline levels were reported). The difference between the findings here and those of Smiljanic (2013) can be attributed in part to the stimuli. Both this study and Schum (1996) used relatively short, meaningful sentences (e.g for Schum, 1996, *Their room was clean*). In contrast, the materials used in Smiljanic (2013) consisted of more difficult, semantically anomalous sentences (e.g. *A cabbage would sink his tired Tuesday)*. This could have elicited age-related production difficulties not found here. Additionally, the set of talkers in Smiljanic (2013) was relatively smaller, with only 5 older adults included in the study versus 10 here and in Schum (1996). It is also important to consider the differing age ranges across the studies; talkers in Smiljanic (2013) ranged from 65 to 78 years old, with a mean of 71.4 years while Schum used younger talkers (range: 62-70 years old, no mean given). However, since the age range in the current study more closely resembles that of Smiljanic's (range: 60-84 years old, mean: 70.2 years), the difference in findings does not appear to be due to the talkers' chronological age.

The findings that both noise-adapted and clear speaking style adaptations enhanced intelligibility were in line with previous work. That noise-adapted and clear speech increases to intelligibility were larger individually than when in conjunction suggests a limitation in the extent to which talkers are able to modify their speech with the goal of increasing intelligibility. This interaction differs from Gilbert et al. (2014) who found noise-adapted clear speech provided a larger intelligibility benefit than the sum of the two adaptations. The difference in findings likely originates from the large talker-to-talker variability; while the findings here are based on the productions of 30 different talkers, Gilbert et al. (2014) only examined a single young adult talker.

A notable finding is that the magnitude of the intelligibility benefit for noise-adapted vs. clear speech varied as a function of age (QCL columned compared to NCO columns in Figure 4). For both children and young adults, conversational quiet-to-noise speech modifications enhanced intelligibility more than quiet conversational-to-clear speech modifications. That is to say, the adaptation to environment-oriented difficulties increased intelligibility to a larger extent than the adaptation to listener-oriented difficulties. One explanation could be that the noise-adapted speech was produced in response to the actual masker used in the word recognition task, which made it more resistant to the masking impact of the noise. Another explanation could be that, independent of the noise-adapted speech's correlation with the noise masker, the set of acoustic-articulatory modifications that characterized the noise-adapted speech productions gave rise to a larger intelligibility benefit compared to the set of modifications that characterized the clear speech productions.

The opposite pattern was observed for older adults: older adults increased their intelligibility through conversational-to-clear speech modifications more than through quiet-to-noise speech modifications. This may have originated from the elevated hearing thresholds that naturally occur with aging. Research has shown the magnitude of the noise adaptation to depend on the intensity of the noise (Lu and Cooke, 2008), suggesting that talkers with diminished acoustic sensitivity might adapt their speech to noise to a lesser degree.

Finally, it is important to acknowledge that the listeners in this study were only young adults. While it is possible that young adult listeners would find young adult talkers more intelligible, the results indicate otherwise. Older adult were found to be as intelligible as young adults. Furthermore, the relative talker intelligibility results (see Figure 32 in the next chapter) show that the individual talkers found to be most intelligible comprised all 3 age groups. This finding lends support for the notion that talker intelligibility is to a large extent dependent on the acoustic-phonetic characteristics of the talker's speech (Bradlow et al., 1996; Hazan and Markham, 2004; Smiljanic and Bradlow, 2007). However, the finding that children were less intelligible than young adults in noise-adapted and clear speech could be due to the young adult listeners' lack of exposure to children's speech relative to young and older adults' speech. In order to differentiate between the effects of acoustic-articulatory modifications vs. effects of talker-listener interactions, it is important to include children and older adult listeners in future work.

These results extend our understanding of speaking style adaptations across age, yielding novel findings about the intelligibility of children and older adults producing noise-adapted speech, and the intelligibility of clear speech in all 3 talker groups at once.

EXPERIMENT 2B: PERCEIVED AGE

Research Aims

In order to accurately discuss the acoustic and perceptual impacts of talker age, it is important to confirm that the talkers in this study did indeed exhibit typical vocal aging and were perceived to belong to 3 distinct age groups. The goal of Experiment 2B was thus to examine the relationship between talkers' chronological and perceived ages, and confirm that the 3 talker groups were indeed representative of their age groups.

Methods

Stimuli

A subset of the stimuli from Experiment 1 was used: the 20 sentences not used in Experiment 2A, as produced by all 30 talkers in all 4 speaking styles (2400 sentences total).

Listeners

48 listeners from Experiment 2A also provided age assessments in this experiment (range 18-35 years old, mean 20.0 years).

Procedure

Sentences were presented in quiet, beginning with the same 5 practice sentences used in Experiment 2A to familiarize listeners with the task. Each listener participated in 1 of 12 conditions, in which they were exposed to 20 talkers composing 2 of the 3 age groups (Experiment 2A was run on the talkers in the 3rd age group) (see table below for distribution). This design ensured no overlap between talkers/materials across the 2 perception experiments. Within a session, the stimuli were all presented in one speaking style. Sessions alternated between all 4 speaking styles, so as to obtain perceived age data on the range of speaking styles presented in Experiment 2A. Every listener heard each of the 20 talkers producing 1 sentence from the pool of 20 sentences (no sentence repetition within a listening condition). Each talker's age was thus estimated by 32 listeners. Presentation order and talker-sentence pairing were randomized for each listener. The experiment was presented in MATLAB. Listeners were instructed to guess how old each talker was, typing one estimate at a time on the keyboard after stimulus presentation. Responses were open-ended. The test was self-paced.

Table 14: Conditions for Experiment 2B

Condition		**Description**	
1	A	Age estimates for young and older adults (following child condition for Exp. 2A)	in QCO
	B		in QCL
	C		in NCO
	D		in NCL
2	A	Age estimates for children and older adults (following young adult condition for Exp. 2A)	in QCO
	B		in QCL
	C		in NCO
	D		in NCL
3	A	Age estimates for children and young adults (following the older adult condition for Exp. 2A)	in QCO
	B		in QCL
	C		in NCO
	D		in NCL

Statistical Analysis

In order to investigate how accurate the age estimates were in relation to the talkers' chronological ages, results were submitted to a mixed effects linear regression in SPSS. Listener accuracy (the difference between each estimate and the talker's actual age) served as the continuous dependent variable, Talker, Sentence, and Listener were included in the model as random factors and Talker Age Group (children, young adults, or older adults), Environment-Oriented Speaking Style (produced in quiet or in response to noise), Listener-Oriented Speaking Style (conversational or clear), and their interactions were included as fixed effects. Random slopes were included in the model for both Listener- and Environment- Oriented Speaking Style at the level of Talker, since this level showed the greatest variance.

Hypotheses

I expected the 3 chronological age groups to form similar perceived age groups, although the older adult perceived ages could be lower than their chronological ages due to the self-selection of the subject pool (i.e. these were older talkers still active enough to travel to campus, and perceived age is strongly affected by physiological condition) (Ramig and Ringel, 1983). I expected the ages of children and young adults to be accurately estimated, since inaccurate estimation (in both directions) tends to occur with single word stimuli and here listeners were exposed to the entire sentence (Amir et al., 2012; Assman et al., 2013).

Results

Random effects from the regression are summarized in Table 15, and fixed effects in Table 16. The results revealed that the accuracy of estimating a talker's age was significantly affected by Talker Age Group [$F(2,27.849)=46.691$, $p<0.001$]. Neither speaking style had a significant effect, nor were there any significant interactions. Actual ages vs. average perceived age estimates for the 30 talkers are shown in Figure 39.

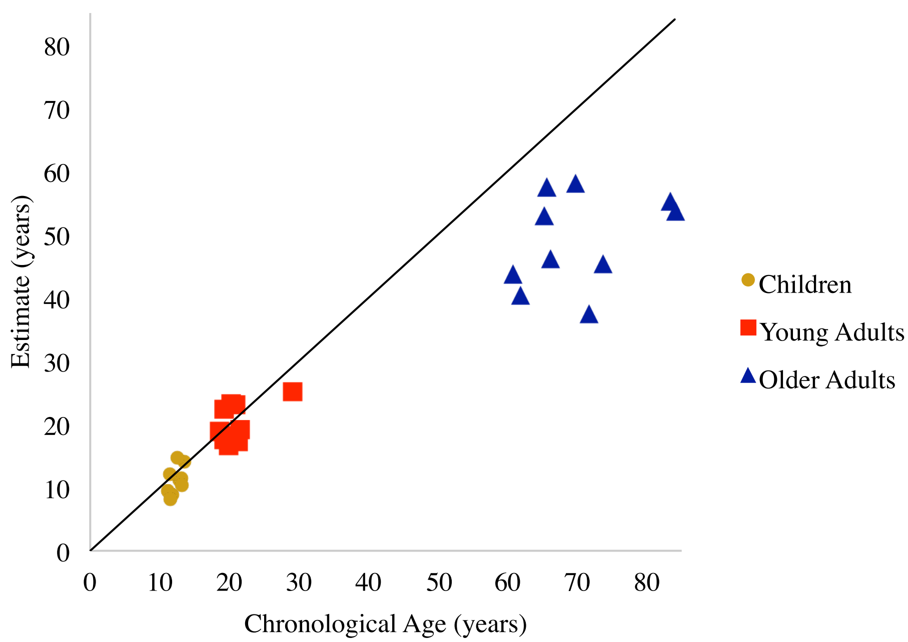
Table 15: Experiment 2B summary of random effects for the mixed model

<table>
<tr><td colspan="2">Parameter</td><td>Estimate</td><td>Std. Error</td></tr>
<tr><td colspan="2">Residual</td><td>49.962867</td><td>2.547228</td></tr>
<tr><td>Intercept [subject = Sentence]</td><td>Variance</td><td>1.733959</td><td>0.93216</td></tr>
<tr><td>Intercept [subject = Listener]</td><td>Variance</td><td>8.144578</td><td>2.375773</td></tr>
<tr><td rowspan="2">Intercept + QN + COCL [subject = Talker]</td><td>CS diagonal offset</td><td>0.917898</td><td>0.830105</td></tr>
<tr><td>CS covariance</td><td>2.939513</td><td>0.918828</td></tr>
</table>

Table 16: Experiment 2B summary of fixed effects for the mixed model. Significant effects are highlighted in yellow.

Source	Numerator df	Denominator df	F	Sig.
QN	1	43.697	0.008	0.929
COCL	1	43.708	0.908	0.346
CHYAOA	2	27.849	46.691	0.000
QN*COCL	1	41.128	0.428	0.517
QN*CHYAOA	2	76.412	2.595	0.081
COCL*CHYAOA	2	76.355	2.338	0.103
QN*COCL*CHYAOA	2	816.275	2.388	0.092

Figure 39: Accuracy of perceived age for children, young adult, and older adult talkers



The results showed that the 3 talker groups formed 3 separate age groups as perceived by listeners. Listener accuracy for estimating the age of the older adult talkers significantly differed from that for the younger talker groups (estimates and pairwise comparisons can be found in Tables 17 and 18). That is to say, listener accuracy for older adult talkers' ages was significantly different than listener accuracy for children and young adult talkers' ages. The means in Table 17 indicate that this stemmed from an underestimation of older adult talkers' ages compared to the younger talkers' ages. On average, listeners estimated the children talkers to be 0.43 years (about 5 months) younger than their actual age, i.e. an average perceptual age of 11.9 years vs. their average chronological age of 12.3 years. For young adult talkers, listeners underestimated age by 1.45 years (about 17 months), i.e. an average perceptual age of 19.6 years vs. their average chronological age of 21.0 years. Older adult talkers, on the other hand, were underestimated by 21.56 years on average, i.e. an average perceptual age of 48.6 years vs. their average chronological age of 70.2 years. These trends are evident in Figure 31.

Table 17: Experiment 2B estimates for the main effect of Talker Age Group

CHYAOA	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
CH	-0.428	1.813	-4.119	3.263
OA	-21.529	1.810	-25.215	-17.843
YA	-1.453	1.811	-5.141	2.235

Table 18: Experiment 2B pairwise comparisons for the main effect of Talker Age Group. Significant effects are highlighted in yellow.

(I) CHYAOA	(J) CHYAOA	Mean Difference (I-J)	Std. Error	Sig.[c]	95% Confidence Interval for Difference[c]	
					Lower Bound	Upper Bound
CH	OA	21.101*	2.464	0.000	16.053	26.149
	YA	1.025	2.466	0.681	-4.027	6.077
OA	CH	-21.101*	2.464	0.000	-26.149	-16.053
	YA	-20.076*	2.463	0.000	-25.123	-15.029
YA	CH	-1.025	2.466	0.681	-6.077	4.027
	OA	20.076*	2.463	0.000	15.029	25.123

Discussion

The results of Experiment 2B showed that listeners accurately estimated children and young adult talkers' ages (underestimations of 5 and 17 months, respectively), but significantly underestimated the ages of the older adult talkers (over 21 years). In spite of this underestimation, older adult talkers still formed a 3rd perceptually-distinct group (see Figure 31; the older adult talkers in blue still form a separate group along the y axis).

These results support the findings from Experiment 1. In line with the set of acoustic-phonetic differences between the children and adult talkers found in Experiment 1 (e.g. increased F0 in children relative to young adults), these results showed that the children talkers were perceived as younger than the adult talkers (listener accuracy was high for the younger groups). That older adult talkers were judged to be significantly younger than their actual age is in line with the finding that many of the acoustic-phonetic features that typically accompany aging (e.g. decreased HNR, increased shimmer) were not found for these talkers in Experiment 1; as previous work has shown, age-related changes to speech production are not only influenced by chronological age, but also by physiological condition (Ramig and Ringel, 1983). However, although they sounded (on average) 22 years younger than their actual age, and exhibited voice quality characteristics (HNR, jitter, shimmer) comparable to those of young adults, they did still differ from younger talkers in certain acoustic-phonetic aspects such as less energy in the 1-3 kHz range, and a slower speaking rate with longer vowels (and formed a separate perceptual group).

The lack of differences between young and older adults in Experiment 2A intelligibility rates thus cannot be attributed to these talkers being perceived as belonging to the same age group. Rather, it reflects the ability of both age groups to enhance intelligibility to a similar degree. The underestimation of the older adult talkers' ages may also explain differences between the results of Experiment 2A and the findings from Smiljanic (2013), in which older adults were less intelligible than young adults, unlike here.

In conclusion, the talkers selected for this study appear to be appropriate for a discussion of age-related changes in speech, although future studies will need to include older adults with greater physiological aging.

Relative talker intelligibility

RESEARCH AIMS

The additional analyses reported here address the extent to which the relative intelligibility of individual talkers is consistent across speaking styles. Previous research has shown relative intelligibility for a number of talkers and talker groups to be consistent across multiple listener groups (such as native and nonnative listeners, or normal-hearing and cochlear implant listeners) and listening environments (such as vocoded speech and speech masked by multi-talker babble) (Hazan and Markham, 2004; Green et al., 2007; Bent et al., 2009; van Dommelen and Hazan, 2012).

Findings regarding relative talker intelligibility across speaking styles are less consistent; studies have shown enormous between-talker variability in the ability to enhance intelligibility via speaking style adaptations (Gagne et al., 2002; Bradlow et al., 2003; Ferguson, 2004; Ferguson and Kewley-Port, 2007). However, Ferguson (2004) found a significant correlation between the intelligibility of 41 adult talkers' conversational and clear speech vowels, indicating that the ranking of least to most intelligible talkers was similar in both speaking styles. It remains to be seen if relative talker intelligibility is consistent across a more diverse group of talkers producing multiple types of speaking style adaptations. Here I examine the extent to which young adult listeners find the same talkers (across children, young adults, and older adults) to be intelligible across multiple speaking styles (conversational and clear speech produced in quiet and in response to noise).

METHODS

The materials, listeners, and procedure were those from Experiment 2A.

Statistical Analysis

In order to investigate the extent to which talkers were consistent in their intelligibility across speaking styles, Experiment 2A scores were aggregated for each talker in each speaking style. Four correlation analyses (n=30 for each correlation) were carried out using the cor.test() function in R. The analyses comparing cross-style relative talker intelligibility were as follows: 1) speech produced in quiet (QCO vs. QCL), 2) noise-adapted speech (NCO vs. NCL), 3) conversational speech (QCO vs. NCO) and 4) clear speech (QCL vs. NCL).

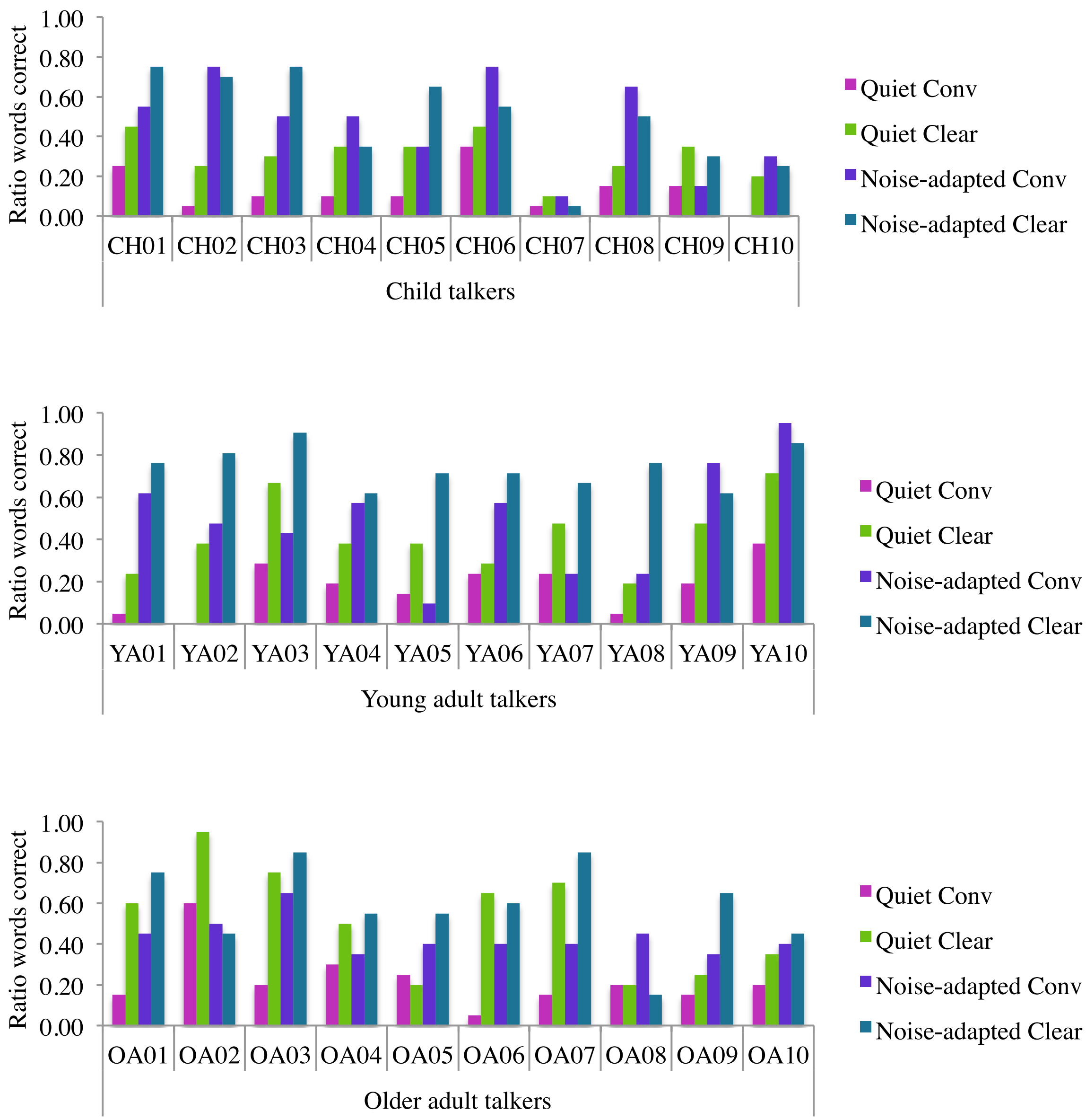
Hypotheses

While relative talker intelligibility is consistent across different listener populations and listening conditions, findings across speaking styles have been significantly more mixed. Given that Experiment 1 found environment- and listener-oriented speaking style adaptations to be characterized by several different acoustic-phonetic modifications depending on the talker age group, it was suspected that relative talker intelligibility would vary significantly across speaking styles, reducing the correlational significance.

RESULTS

As shown in Figure 40, the ratio of words correctly transcribed by listeners for each talker in each speaking style ranged from 0.00 to 0.95, with a mean of 0.41—a large enough range of intelligibility to be used as the basis for an investigation of relative talker intelligibility. For each talker, a ranking of intelligibility in the 4 speaking styles can be found in Table 20, and a list of proportional gains [proportional gain=(intelligibility in enhanced style – intelligibility at baseline)/intelligibility at baseline] for clear and noise-adapted speaking style adaptations can be found in Table 21[7].

Figure 40: Individual talker intelligibility for conversational and clear speech produced in quiet and in response to noise. Top: children, middle: young adults, bottom: older adults.



[7] Gain ratios were multiplied by 100 to convert into percentages

The correlation analyses examining relative talker intelligibility in each speaking style are shown in Table 19. Three of the 4 correlations were significant; that is to say, the rankings of relative intelligibility among the 30 talkers were significantly correlated in all but one of the comparisons. Most strongly correlated were intelligibility scores for talkers across speaking styles in quiet (i.e. conversational and clear speech produced in quiet) ($r=0.620$, $p<0.001$) (see Figure 41). Significantly correlated, but more weakly, were intelligibility scores for talkers in clear speech (i.e. clear speech both produced in quiet and in response to noise) ($r=0.432$, $p=0.017$), and noise-adapted speech (i.e. conversational and clear speech produced in response to noise) ($r=0.374$, $p=0.042$) (see Figures 42 and 43). Intelligibility scores for talkers in conversational speech (i.e. conversational speech in quiet and in response to noise) were not significantly correlated ($r=0.307$, $p=0.099$); that is, the intelligibility of talkers in quiet conversational speech did not significantly correlate with their intelligibility in noise-adapted conversational speech (Figure 44).

Table 19: Relative talker intelligibility correlations between speaking style adaptions. Significant effects are highlighted in yellow.

Analysis	Datasets	p	r
Quiet	QCO, QCL	0.000	0.620
Noise	NCO, NCL	0.042	0.374
Conversational	QCO, NCO	0.099	0.307
Clear	QCL, NCL	0.017	0.432

Figure 41: Relative talker intelligibility correlation between talkers producing conversational and clear speech in quiet (QCO and QCL, respectively)

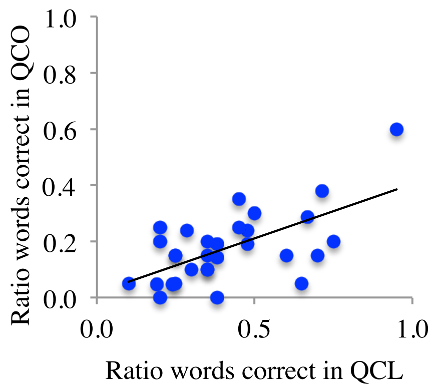


Figure 42: Relative talker intelligibility correlation between talkers producing clear speech in quiet and in response to noise (QCL and NCL, respectively)

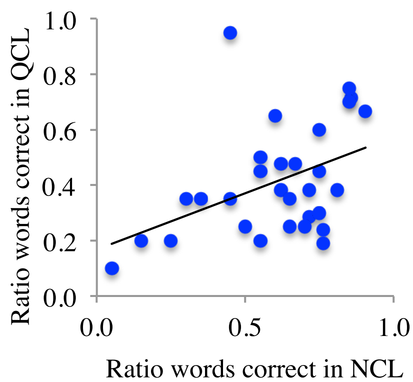


Figure 43: Relative talker intelligibility correlation between talkers producing conversational and clear speech in noise (NCO and NCL, respectively)

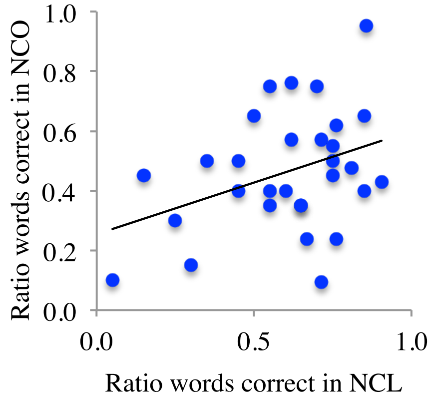


Figure 44: Relative talker intelligibility correlation between talkers producing conversational speech in quiet and in response to noise (QCO and NCO, respectively)

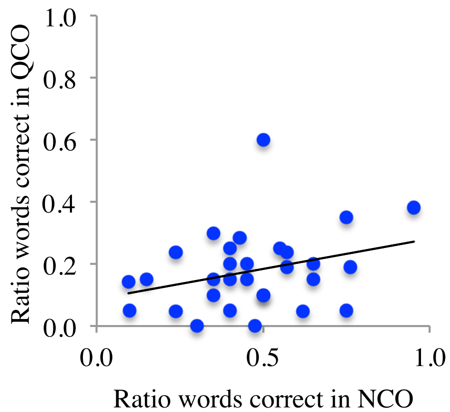


Table 20: Ranking of relative talker intelligibility by speaking style

Talker	Rank from most to least intelligible			
	QCO	QCL	NCO	NCL
CH01	7	11	10	8
CH02	25	22	3	13
CH03	21	20	12	9
CH04	22	16	13	26
CH05	23	17	24	16
CH06	3	12	4	20
CH07	24	30	29	30
CH08	16	24	6	23
CH09	17	18	28	27
CH10	29	26	25	28
YA01	28	25	7	6
YA02	30	13	14	5
YA03	5	5	17	1
YA04	13	15	8	18
YA05	20	14	30	12
YA06	8	21	9	11
YA07	9	10	26	14
YA08	27	29	27	7
YA09	14	9	2	17
YA10	2	3	1	2
OA01	18	7	15	10
OA02	1	1	11	24
OA03	12	2	5	3
OA04	4	8	22	22
OA05	6	28	18	21
OA06	26	6	19	19
OA07	19	4	20	4
OA08	10	27	16	29
OA09	15	23	23	15
OA10	11	19	21	25

Table 21: Proportional gain for clear (CL) and noise-adapted (N) speaking style enhancements

Talker	QCO (baseline)	Proportional gain		
		QCL	NCO	NCL
CH01	0.25	80%	120%	200%
CH02	0.05	400%	1400%	1300%
CH03	0.10	200%	400%	650%
CH04	0.10	250%	400%	250%
CH05	0.10	250%	250%	550%
CH06	0.35	29%	114%	57%
CH07	0.05	100%	100%	0%
CH08	0.15	67%	333%	233%
CH09	0.15	133%	0%	100%
CH10	0.00[8]	n/a	n/a	n/a
YA01	0.05	400%	1200%	1500%
YA02	0.00	n/a	n/a	n/a
YA03	0.29	133%	50%	217%
YA04	0.19	100%	200%	225%
YA05	0.14	167%	-33%	400%
YA06	0.24	20%	140%	200%
YA07	0.24	100%	0%	180%
YA08	0.05	300%	400%	1500%
YA09	0.19	150%	300%	225%
YA10	0.38	88%	150%	125%
OA01	0.15	300%	200%	400%
OA02	0.60	58%	-17%	-25%
OA03	0.20	275%	225%	325%
OA04	0.30	67%	17%	83%
OA05	0.25	-20%	60%	120%
OA06	0.05	1200%	700%	1100%
OA07	0.15	367%	167%	467%
OA08	0.20	0%	125%	-25%
OA09	0.15	67%	133%	333%
OA10	0.20	75%	100%	125%

[8] Note that proportional gain cannot be calculated for talkers who exhibit a baseline score of 0

DISCUSSION

The results of the correlational analyses showed that the consistency of relative talker intelligibility depended on the speaking style adaptations examined. For example, relative talker intelligibility showed a strong correlation between conversational and clear speech in quiet. Significant but weak correlations between clear speech produced in quiet and in response to noise, and between conversational and clear speech in noise were also found. No significant correlation was found for relative talker intelligibility in conversational speech produced in quiet vs. in response to noise.

The present finding that talkers are consistently more/less intelligible across conversational and clear speech in quiet replicates the results from Ferguson (2004), who found that intelligibility for 41 talkers was consistent across these 2 speaking styles for vowels. Here, 30 talkers across 3 age groups also showed similar intelligibility rankings across conversational and clear speech styles for monosyllabic target words embedded in high context sentences. This finding suggests that talkers who are inherently clear have available strategies to increase their intelligibility further. It is not the case that a relatively intelligible talker's baseline intelligibility level precludes them from further implementing acoustic-articulatory changes that enhance their intelligibility. Conversely, talkers whose baseline intelligibility levels are low may not have the ability to modify their speech patterns in a way that increases their intelligibility to the levels of highly intelligible talkers.

Adapting speech to noise induced a larger amount of variability in relative talker intelligibility, even for talkers that showed comparable intelligibility levels in quiet. That is, producing speech in noise induced variable adaptations that resulted in less consistent intelligibility rankings across talkers. For example, Talker YA10 and Talker OA02 were 2 of the most intelligible talkers in quiet (refer back to Table 20). YA10 maintained this high level of intelligibility in noise-adapted speech, remaining one of the most intelligible talkers (proportional gains of 150% in NCO and 125% in NCL). OA02, on the other hand, dropped to the 11th most intelligible talker in noise-adapted conversational speech, and to the 24th most intelligible talker in noise-adapted clear speech (proportional gains of -17% and -25%, respectively).

While Talker YA10 and a few other talkers (e.g. Talker CH10) exhibited a consistent level of relative intelligibility across speaking styles, more striking was the large relative talker variability in intelligibility from style to style. The variation in proportional gain reflects the variation in relative talker intelligibility; the conversational-to-clear speech proportional intelligibility gain in quiet (QCL compared to QCO) ranged from 0% to 1200% (mean=191%), the quiet-to-noise-adapted conversational speech proportional intelligibility gain (NCO compared to QCO) ranged from -17% to 1400% (mean=258%), and the proportional intelligibility gain for the 2 enhancements in conjunction (NCL compared to QCO) ranged from -25% to 1500% (mean=386%) (refer back to Table 21).

Some talkers were able to enhance their intelligibility much better than others; for instance, Talker YA02 was the least intelligible talker when producing conversational speech in quiet (30th most intelligible), but became relatively much more intelligible when enhancing intelligibility via noise-adapted and clear speech adaptations (5th most intelligible talker in noise-adapted clear speech). Other talkers seemed to respond better to one type of enhancement than the other; for example, Talker YA01 was relatively unintelligible in quiet, regardless of the listener-oriented speaking style (28th most intelligible in QCO and 25th most intelligible in QCL) but enhanced intelligibility more successfully than other talkers in response to noise (7th most intelligible in NCO and 6th most intelligible in NCL, with proportional gains of 1200% and 1500% respectively). In contrast, Talker OA04 was a relatively intelligible talker in quiet conversational speech (4th most intelligible), but dropped in relative intelligibility with each adaptation (8th most intelligible in QCL, 11th most intelligible in NCO, and 24th most intelligible in NCL).

Although relative talker intelligibility scores were significantly correlated across speaking styles in quiet, in noise, and in clear speech, they were not between speaking styles in conversational speech. Additionally, when all 4 speaking styles were examined in conjunction, the ranking of least/most intelligible talkers showed considerable variability. Unlike the relative talker intelligibility findings shown across different listener populations and environments, the relative talker intelligibility across different speaking styles appears to be less consistent. This could be due in part to the fact that the cross-listener/environment studies were directed towards the relative ranking of talkers' intrinsic clarity as a function of listener group/environmental degradation. These cross-style studies, on the other hand, highlight the relative ranking of talkers' ability to implement intelligibility-enhancing speech modifications (which, as evidenced by Table 21, are largely variable).

As shown in Experiment 1, talker groups implemented a diverse range of modifications when enhancing intelligibility, resulting in varying degrees of intelligibility benefit (as shown in Experiment 2A). The large cross-style inconsistency in relative talker intelligibility found here may in part reflect the variation in intelligibility gain across the 3 talker groups, e.g. older adults showed a smaller noise-adapted gain than the younger talkers, which would reduce correlational significance between speech in quiet and speech in noise[9]. This variation in relative talker intelligibility across speaking styles mirrors the large individual variability in the production and perception of speaking styles

[9] Correlation analyses examining relative talker intelligibility for different age groups would require more than 10 talkers per age group, thus age-related differences in relative talker intelligibility are not directly examined here

found in previous research (Tartter et al., 1993; Krause and Braida, 2004; Ferguson, 2004; Ferguson and Kewley-Port, 2007). The cross-style variability illustrated in these analyses suggests that who listeners perceive as an intrinsically “good” vs. “bad” talker may change depending on the communicative context.

Production & Perception

Research Aims

The goal of this set of analyses was to examine more closely the relationship between production and production; that is, to identify the specific acoustic-phonetic features associated with improved speech intelligibility[10]. The production study (Experiment 1) examined how the acoustic measures vary as a function of style. Here, the question is how variation in the acoustic measures affects intelligibility. For example, does a speaking rate decrease (characteristic of environment- and listener-oriented speaking style modifications) contributes to increased intelligibility? Despite a large number of studies examining the acoustic-phonetic characteristics of more/less intelligible speech, the impact of individual acoustic-phonetic features on intelligibility is unclear. Studies have traditionally examined either the acoustic-phonetic characteristics of 1) ''intrinsically clear'' talkers, i.e. talkers who are relatively more intelligible than other talkers, or 2) "deliberately clear" speaking style adaptations, i.e. speaking styles that are relatively more intelligible than other styles.

Some studies of "intrinsically clear" talkers have shown F0 range and vowel measures (range in F1, vowel space dispersion, F2-F1 distance for /i/ and F2-F1 distance for /a/) to be more correlated with intelligibility compared to speaking rate and mean F0 (Bradlow et al., 1996). Others have found longer word and vowel durations, differentiated vowel space, maximal cues for consonantal contrasts, and low variation in stressed vowel amplitude to characterize intrinsically more intelligible speakers (Bond and Moore, 1994). Additional studies found 1-3 kHz energy and word duration to significantly correlate with intelligibility, more so than the long-term average spectrum

[10] Given that the stimuli sets differ (e.g. acoustic measures done on stimuli in quiet vs. word recognition in noise), this is an exploratory analysis

slope, F0 measures, CV ratios, and vowel formant measures (Hazan and Markham, 2004; Green et al., 2007; van Dommelen and Hazan, 2012).

Studies of "deliberately clear" speech have mainly been limited to a discussion of the set of acoustic-phonetic changes that characterize the speaking style adaptations, i.e. the typical acoustic-phonetic modifications made in quiet-to-noise and conversational-to-clear speaking style adaptations (slower speaking rate, wider F0 range, expanded vowel space, etc.). Few of these studies have examined the individual impacts of different acoustic-phonetic features on intelligibility (i.e. modifications not in conjunction). Some studies have manipulated duration cues of conversational and clear speech, showing that longer durations and slower speaking rates do not necessarily lead to greater intelligibility (Picheny et al. 1986; Picheny et al., 1989; Uchanski et al., 1996; Krause and Braida, 2002; Summers et al., 1988; Pittman and Wiley, 2001; Goy et al., 2007; Cooke, Mayo, Villegas, 2014). Both Krause and Braida (2004) and Cooke, Mayo, and Villegas (2014) found that speaking style adaptation increases in spectral energy were associated with improved intelligibility. Ferguson and Kewley-Port (2002, 2007) found that significantly longer vowel duration and larger vowel space expansion characterized talkers who produced a large clear speech benefit compared to talkers who produced small clear speech benefit.

Although some acoustic correlates of intelligibility have been identified, there is significant variability in the characteristics of intelligible speech. There is a need to examine talker intelligibility by including: 1) a set of talkers differing in intrinsic clarity, 2) a set of speaking styles differing in communicative intent and clarity, and 3) a comprehensive list of acoustic-phonetic features, including additional measures that have previously been excluded from such analyses (e.g. voice quality measures, which Experiment 1 showed to be significant features of intelligibility-enhancing speaking style adaptations). Given that this study revealed that both acoustic-articulatory adjustments (Experiment 1) and word recognition in noise (Experiment 2A) vary as a function of speaking style and age, it is important to consider the production-perception link more closely. To this end, I examined the extent to which each of the 11 acoustic features from in Experiment 1 predicted the word recognition results from Experiment 2A. This analysis took into account a set of speaking styles aimed at deliberately enhancing clarity, produced by a range of 30 talkers differing in their baseline intrinsic clarity.

METHODS

The materials, listeners, and procedure were those from Experiment 1 and Experiment 2A.

Statistical Analysis

In order to examine the relationship between production and perception (Experiments 1 and 2A), mixed effects logistic regressions were carried out using the lme4 package in R. Eleven regressions were run in total, with keyword identification (i.e. correct or incorrect) as the dichotomous dependent variable and, for each regression, one of the 11 acoustic features as the independent variable. Talker, Sentence, and Listener were included in the model as random factors, except in the voice analyses (where there was not enough Sentence variation to include Sentence as a third random factor). Given that the itemized datasets from Experiments 1 and 2A differed in number (global measures: n=7200, segmental measures: n=960, voice measures: n=240, word recognition measures: n=2440), analyses were run on the set of overlapping items (e.g.

for segmental, on just the intelligibility results that had a corresponding segmental measure)[11]. This determined the impact of acoustic variation on word recognition

Hypotheses

Several studies have examined acoustic-phonetic change in relation to the change in intelligibility, with largely differing results. Both Bradlow et al. (1996) and Krause and Braida (2002) did not find durational measures such as speaking rate to correlate with intelligibility, while other studies found durational measures such as word duration to significantly correlate with intelligibility (longer durations with improved intelligibility) (Hazan and Markham, 2004; Green et al., 2007; van Dommelen and Hazan, 2012). Likewise, Bradlow et al. (1996) found F0 range to strongly correlate with intelligibility (increased range with improved intelligibility), while the opposite was found in Hazan and Markham (2004).

Given these differing findings, I held no predictions regarding which of the typical acoustic-phonetic correlates of intelligibility (durational measures, F0 measures, 1-3 kHz energy) measured in Experiment 1 would predict speech intelligibility scores from Experiment 2A. It was also unclear whether additional features in Experiment 1 that have never been examined in relation to intelligibility (voice quality measures, vowel duration, pause duration) would significantly predict intelligibility.

---

[11] It is important to note that the sets of intelligibility measure and acoustic measures do not align 1:1. That is, this analysis relates acoustic measures obtained from the entire sentences (e.g., F0) or segments that are both within and without the target keyword used to assess intelligibility. This approach allowed an examination of the acoustic-phonetic properties of speaking style adaptations for connected speech, which presents a more realistic speech sample. Furthermore, the overall acoustic-phonetic characteristics of the target sentences contributed to the recognition of the final keywords in listeners' responses, and are thus important to include in an acoustic and perceptual examination.

RESULTS

Given that the overall ratio of words correctly transcribed by listeners for each talker in each style ranged from 0.00 to 0.95, with a mean 0.41 and standard deviation of 0.24, the amount of variance was substantial enough to serve as the basis for an investigation of the effects of acoustic-phonetic talker characteristics on overall speech intelligibility (refer back to Figure 40).

Fixed effects are summarized in Table 22, and random effects in the appendix (Table 52). The results of the mixed-effects logistic regressions revealed that the probability of correct keyword identification was significantly affected by 4 acoustic cues: 1-3 kHz energy [$F(1,2438)=51.250$, $p<0.001$], speaking rate [$F(1,2435)=13.420$, $p<0.001$], pause duration [$F(1,2431)=30.950$, $p<0.001$], and vowel duration [$F(1,364)=4.252$, $p=0.040$]. With regard to the main effect of 1-3 kHz energy, results showed that speech with more 1-3 kHz energy was significantly more intelligible (see Figure 45). A slower speaking rate with longer vowels and longer pauses also predicted more intelligible speech (see Figures 46-48).

Table 22: Summary of fixed effects for the mixed models. Significant effects are highlighted in yellow.

Regression No.	Feature	Numerator df	Denominator df	F	Sig.
1	F0 range	1	2438	3.062	0.080
2	F0 mean	1	2438	1.164	0.281
3	1-3 kHz energy	1	2438	51.250	0.000
4	Speaking rate	1	2435	13.420	0.000
5	Pause duration	1	2431	30.950	0.000
6	Vowel duration	1	364	4.252	0.040
7	F1	1	364	0.874	0.350
8	F2	1	364	2.828	0.094
9	HNR	1	59	0.970	0.329
10	Jitter	1	59	0.462	0.500
11	Shimmer	1	59	0.288	0.594

Figure 45: Histogram, curve, and scatterplot for the relationship between 1-3 kHz energy and word recognition in noise

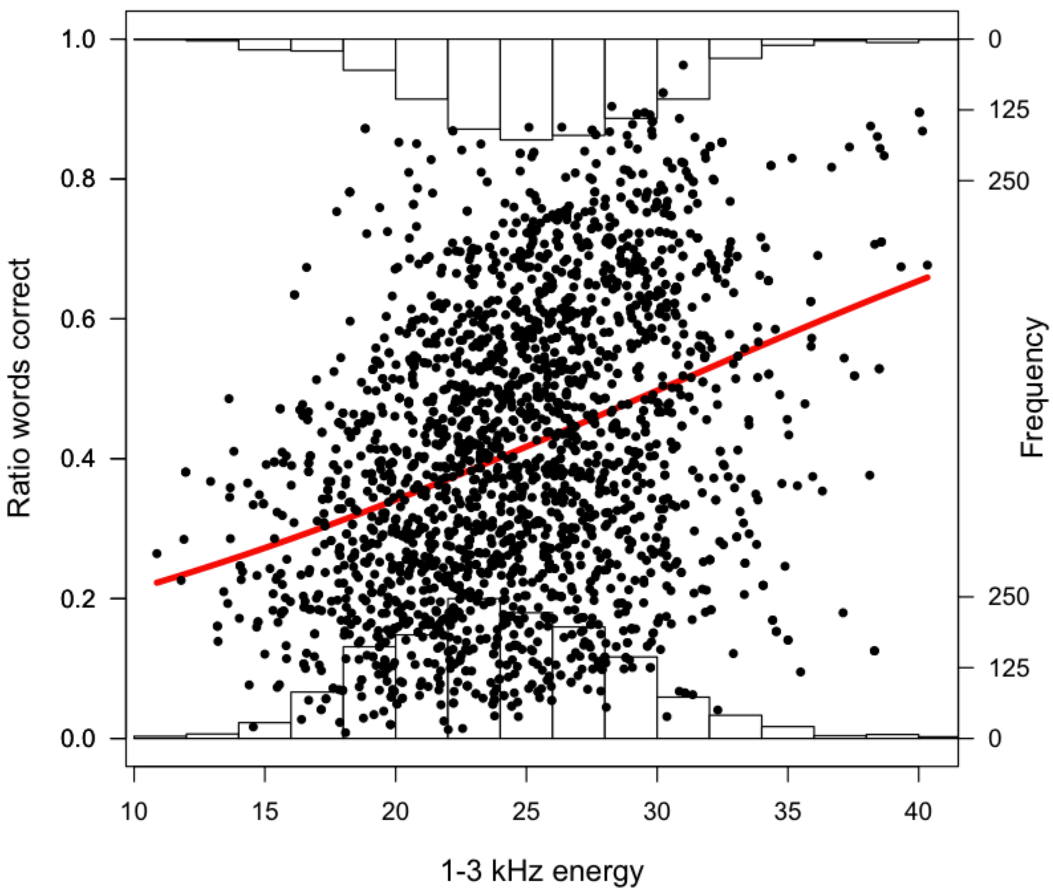


Figure 46: Histogram, curve, and scatterplot for the relationship between speaking rate and word recognition in noise

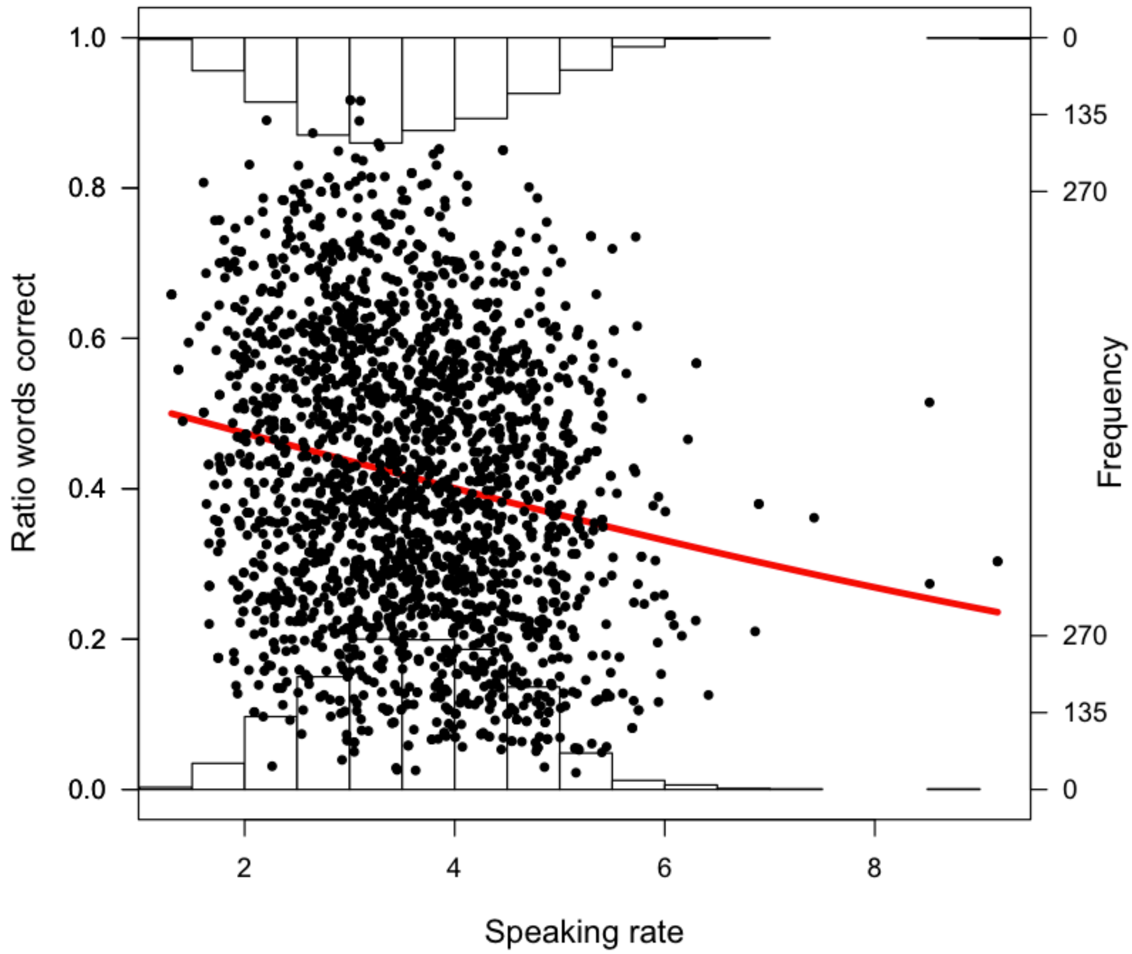


Figure 47: Histogram, curve, and scatterplot for the relationship between pause duration and word recognition in noise

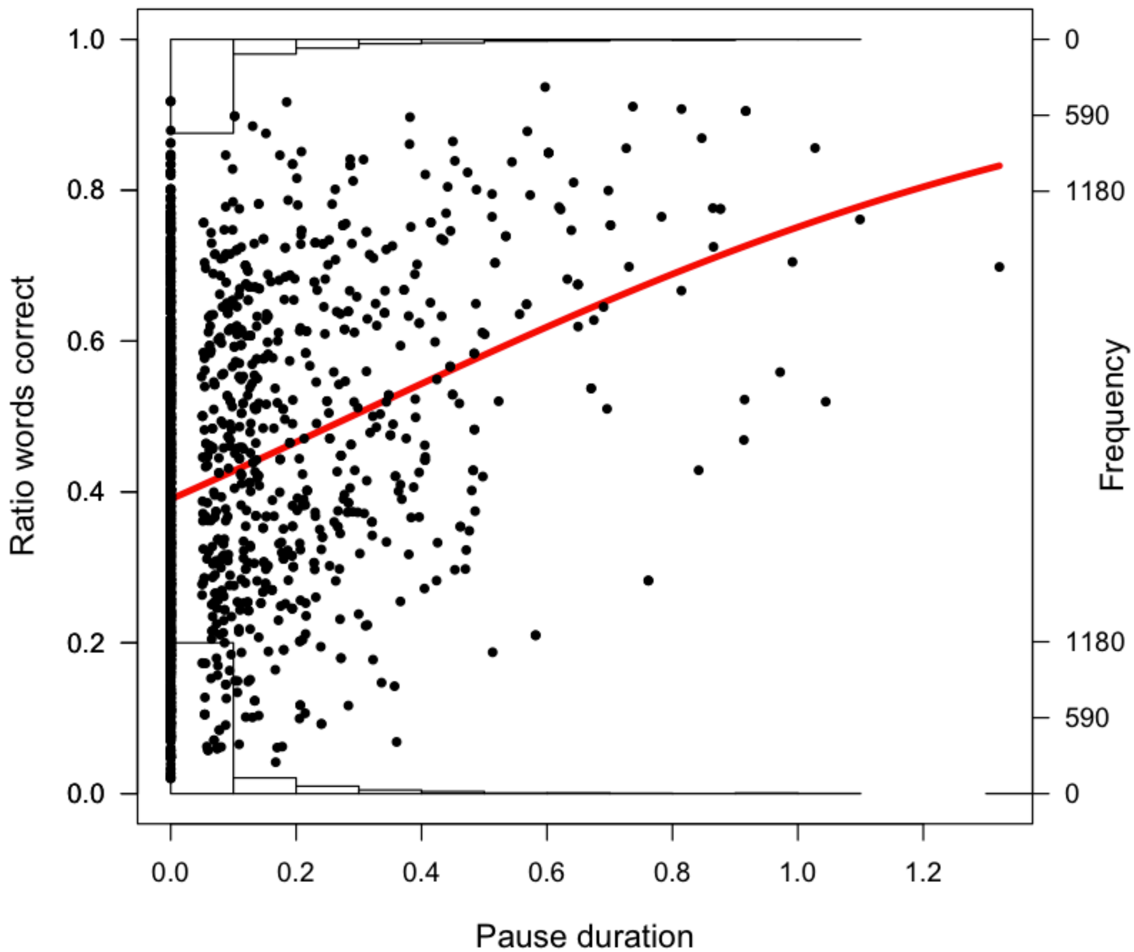
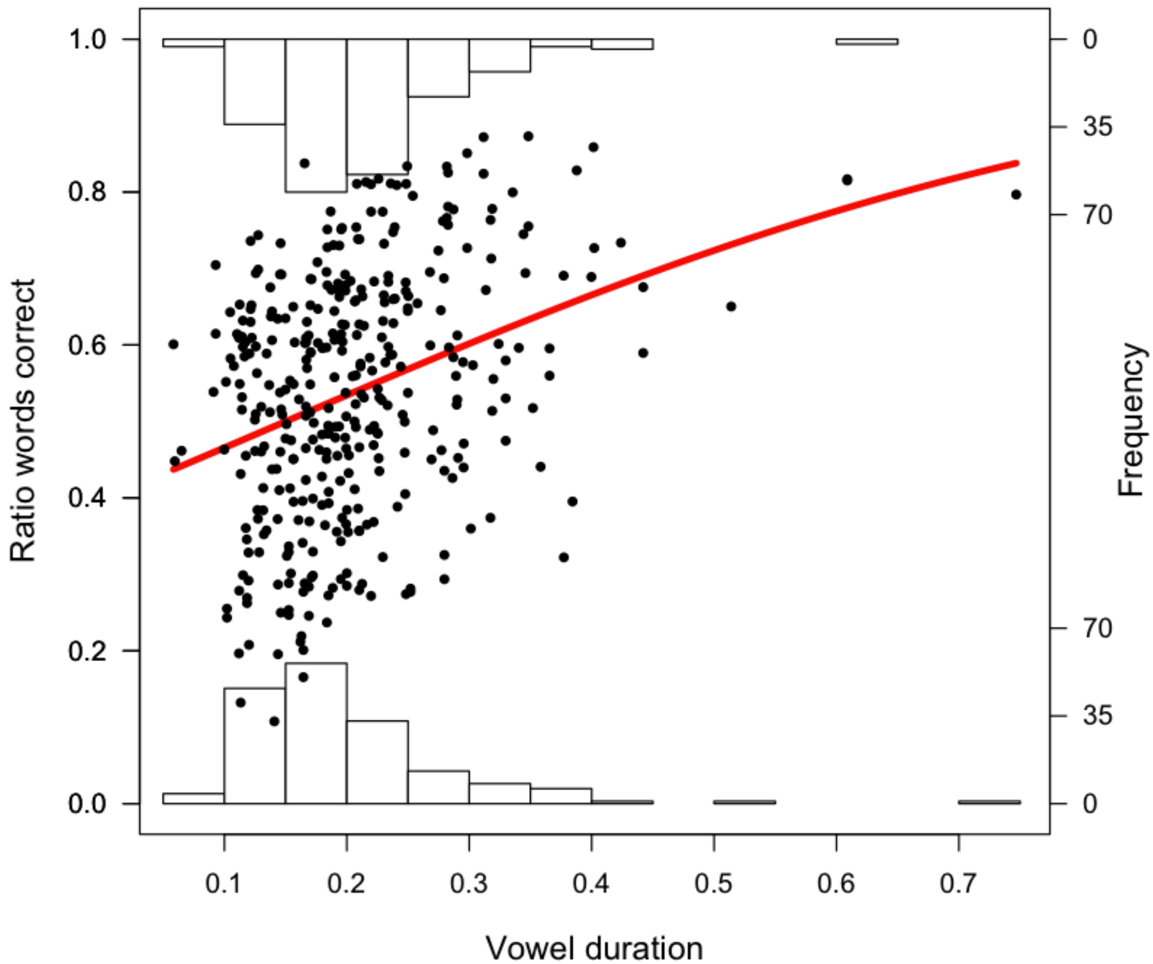


Figure 48: Histogram, curve, and scatterplot for the relationship between vowel duration and word recognition in noise



DISCUSSION

Regressions analyzing acoustic features (Experiment 1) as predictors of word recognition scores (Experiment 2A) revealed that spectral and durational cues most significantly impacted the variance in intelligibility. That is to say, the acoustic features that most predicted increased word recognition were increased 1-3 kHz energy, a slower speaking rate, increased pause duration, and increased vowel duration.

One acoustic cue that emerged as a strong predictor of intelligibility was 1-3 kHz energy. This is in line with previous studies which found 1-3 kHz energy to significantly correlate with word recognition scores (Krause and Braida, 2004; Hazan and Markham, 2004; Green et al., 2007; van Dommelen and Hazan, 2012). Similarly, Cooke et al. (2014) found that the noise-adapted speech intelligibility benefit is largely due to spectral modifications. This aligns with the Experiment 1 results which showed that talkers significantly increased 1-3 kHz energy in noise-adapted speech compared to speech produced in quiet (they also made increases in clear speech compared to conversational speech, but only in quiet).

Durational cues (speaking rate, pause duration, and vowel duration) also emerged as significant factors affecting intelligibility. Similar findings of durational measures (slower speaking rate, longer vowels and pauses) being associated with improved intelligibility were reported in Hazan and Markham (2004), Ferguson and Kewley-Port (2002; 2007), Bond and Moore (1994), and van Dommelen and Hazan (2012). Generally speaking, the beneficial effects of a slower speaking rate on intelligibility may arise from allowing the listener more processing time, and thus improving their word recognition in noise. Slower speaking rates may also result in the production of more salient acoustic-phonetic cues (e.g. more stop burst releases) and greater articulatory precision (longer and more peripheral vowels).

The difficulty in assessing the contribution of an acoustic cue to improved word recognition is shown in Bradlow et al. (1996), who did not find a correlation between speaking rate and intelligibility for conversational speech in quiet. Along the same line, Krause and Braida (2002) reported that trained talkers were able to enhance intelligibility via clear speech independent of speaking rate. While Cooke et al. (2014) found a large intelligibility benefit for spectrally modifying speech, they did not find any benefit from durational modifications. This suggests that speaking slowly is not entirely responsible for enhancing intelligibility, but rather that there may be several different strategies for enhancing intelligibility, including strategies independent of speaking rate. As shown here, these strategies may furthermore be dependent on age; e.g. Bradlow et al. (1996) and Krause and Braida (2002) did not examine children or older adult talkers (although the exact ages of their adult talkers were not given).

In the current study, measures of F0 mean and range and formant frequencies were not found to significantly predict intelligibility, contrary to findings from previous studies (Bradlow et al., 1996; Bond and Moore, 1994; Hazan and Markham, 2004; van Dommelen and Hazan, 2012). Although these measures were shown to be significantly modified in the production of the intelligibility-enhancing speaking adaptations (e.g. both noise-adapted and clear speech showed higher mean F0 and wider F0 ranges relative to baseline in Experiment 1), these measures were not significant predictors of the intelligibility results. Increased intelligibility may be a combination of different possible strategies that vary across individual talkers.

One possible reason for the lack of significant effects for some acoustic measures is the large amount of variation included within each set of acoustic measures. Here the acoustic variation simultaneously included variation across speaking styles and across age. Results from previous studies examining one age group or one speaking style could thus differ from these results due to the acoustic-phonetic changes across age and speaking style. For example, children showed significantly higher overall F1s than adults. This is likely due to the shorter vocal tracts and did not lead to increased intelligibility. The age-related variation in baseline F1 could have thus masked what would otherwise be a significant contribution of F1 increase (related to speaking style adaptations) to intelligibility scores. This could lead to F1 not being identified as a significant predictor of intelligibility, unlike in other studies that have examined intelligbility variation within a talker group.

Also of note is that here, mixed-effects logistic regression models were used in place of the more traditional correlational analyses, thus accounting for random variance traditionally excluded from analysis. Given that several of the acoustic-articulatory cues found to characterize intelligibility-enhancing speaking style adaptations did not directly contribute to more intelligible speech, it appears that the process of enhancing intelligibility relies on a set of multiple, covarying acoustic-phonetic changes that may change across talkers of different ages, or across speech with varying communicative intent.

General discussion and conclusions

The goal of this dissertation was to increase our understanding of how speech intelligibility is shaped by talker-related factors such as age and communicative intent. Although talkers differ in intrinsic clarity, they can improve the ease to which they are understood via systematic alterations in their speech patterns (e.g. adapting their speech in response to noise or listener difficulty). These intelligibility-enhancing speaking style adaptations have been well researched for healthy young adult talkers, but our knowledge of how they interact and develop across the lifespan is limited. Some of the questions addressed here were: what are the acoustic-phonetic characteristics of quiet-to-noise and conversational-to-clear speech modifications produced by children vs. young adult vs. older adult talkers? How do these acoustic modifications impact perceptual tasks like word recognition? Does relative talker intelligibility vary across speaking styles? Given a diverse range of talkers and speaking styles, what are the acoustic-phonetic predictors of intelligibility?

Understanding the effects of age and communicative intent on speech production (and how this variation shapes speech intelligibility) is critical given the prevalence of communicative difficulties in daily interactions such talking to hearing-impaired listeners or communicating in noisy classrooms. This dissertation enhances our understanding of variation in speech intelligibility by providing a bigger-picture account of within-talker variability in response to adverse communicative situations for different talker populations, and its effects on speech perception. This was the first study to investigate the extent to which clear speech and noise-adapted speech benefits interact with each other across multiple talker groups. It was also one of the first to examine how relative talker intelligibility varies as a function of speaking style.

The major goals of this dissertation were as follows: 1) examine age- and adaptation-related variation in speech production, 2) examine the extent to which age- and adaptation-related changes shape speech intelligibility for young adult listeners (and confirm the age-related changes were accurately perceived), 3) examine the extent to which relative talker intelligibility varies across speaking styles, and 4) identify the acoustic-phonetic predictors of speech intelligibility. To this end, noise-adapted and clear speaking style adaptations were examined in children (11-13 years old), young adults (18-29 years old), and older adults (60-84 years old). Three experiments were run: a production study (Experiment 1) and 2 perception studies assessing word recognition in noise and perceived age (Experiments 2A and 2B). Additional analyses examined relative talker intelligibility across speaking styles and the production-perception link.

The goal of Experiment 1 was to perform a comprehensive acoustic comparison of noise-adapted and clear speech production in children, young adults, and older adults. Eleven acoustic cues spanning global (F0 mean and range, 1-3 kHz energy, speaking rate, pause duration), segmental (vowel duration, F1, F2), and voice (HNR, shimmer, jitter) characteristics were examined.

The findings showed several quiet-to-noise adaptations independent of age. All 3 talker age groups produced quiet-to-noise speech modifications with increased F0 range, increased HNR, and decreased shimmer. While children spoke with more energy in the 1-3 kHz range compared to older adults, all talker groups increased 1-3 kHz energy in response to noise. Overall, older adults exhibited a slower speaking rate (with longer vowels) than the younger talkers, but all talker groups slowed their speaking rate and lengthened their vowels. Similarly, although children spoke with less jitter and higher F2s than the adult talkers, all groups lowered jitter and raised F2 when adapting their speech to noise.

Findings showed several cross-age clear speech strategies as well: all talker groups produced clear speech with a wider F0 range relative to conversational speech. All talker groups also lengthened pauses, although older adults did so to a greater extent. Overall, older adults exhibited a slower speaking rate (with longer vowels) than the younger talkers, but all talker groups slowed their speaking rate and lengthened their vowels when producing clear speech. While children showed higher F1s than adults, all talker groups raised F1 when producing conversational-to-clear speaking style adaptations.

However, several age-related differences emerged in the production of intelligibility-enhancing strategies. As expected, children had higher F0 and formant frequencies due to smaller vocal tracts/oral cavities and shorter/thinner vocal folds. They also modified these F0 and F1 cues more than adults when producing noise-adapted speech, thus revealing age-related differences in response to environmentally-oriented communicative issues. In response to listener-oriented communicative issues, children increased HNR and decreased shimmer, unlike the adult talkers. These changes implemented by children but not adults may be due to the overshoot in acoustic parameter values documented by Lee and colleagues (1999). Future work should investigate the origin of these child-adult differences (e.g. if they are peripheral vs. central in nature).

The direct comparison of the 2 speaking style adaptations (to the environment vs. to the listener) revealed a number of commonalities as well as several acoustic-phonetic differences. For example, voice quality changes were more prevalent in noise-adapted speech while pause duration changes were more salient in clear speech. Several interactions were found as well, e.g. when noise-adapted and clear speaking style adaptations were produced in conjunction, pause duration changes were larger than expected while those for speaking rate were smaller. That environment- and listener-oriented speaking style adaptations were characterized by overlapping but distinct sets of acoustic-phonetic modifications is natural given their communicative intents. While both lines of adaptations aim to enhance intelligibility, listener-oriented speech is focused on enhancing the clarity of phonetic cues to helping the listener to retrieve and decode information. In contrast, speech modifications induced by environmental factors are primarily focused on preserving audibility (Cooke, King, Garnier, Aubanel, 2014). These results contribute to a better understanding of how age- related peripheral and cognitive changes relate to speech production mechanisms across the lifespan.

Experiment 2A examined the impact of these acoustic-phonetic changes on intelligibility. The goal was to examine whether the 3 talker groups differed in their intrinsic intelligibility, and the extent to which they could all implement intelligibility-enhancing modifications. The results revealed that noise-adapted and clear speech both significantly enhanced intelligibility for young adult listeners. Results also showed several age-related differences in intelligibility. Children were overall less intelligible than the young adult talkers, although interactions revealed that this originated from children being less intelligible than young adults in noise-adapted and clear speech only. For speech produced in quiet, children were less intelligible than older adults, and for conversational speech, there were no age-related differences. Young adults and older adults did not significantly differ in intelligibility in any speaking style. It appears that the acoustic-phonetic modifications implemented by children but not adult talkers when producing speaking style adaptations (e.g. children showed greater changes to F0, F1, HNR, shimmer than adults) did not lead to increased intelligibility in children's speech relative to adults'. In fact, children were relatively less successful in enhancing intelligibility for young adult listeners. However, this could be due to the young adult listeners' lack of exposure to children's speech relative to young and older adults' speech. In order to differentiate between the effects of acoustic-articulatory modifications vs. effects of talker-listener interactions, future work should include children and older adult listeners. Finally, the noise-adapted clear speech intelligibility gain was less than what would be expected given the individual gains of noise-adapted speech and clear speech, suggesting a limit to the extent one can enhance intelligibility. Thus it appears that the additional acoustic-phonetic modifications implemented by children in Experiment 1 (e.g. HNR increases in clear speech) were not as successful in enhancing intelligibility compared to the set of modifications made by adults.

Experiment 2B sought to examine whether the 3 sets of talkers (children, young adults, older adults) formed perceptually-distinct groups. Results confirmed that the talkers formed 3 perceptual groups (in line with speech production differences such as a slower speaking rate in the older adult talkers, a higher F0 in children talkers, etc.). Listeners were highly accurate in identifying the ages of the younger talkers, but strongly underestimated the ages of the older adult talkers. This is in line with the lack of age-related voice quality differences found between younger and older adults in Experiment 1, and likely due to the fact that the older adult talkers in this study were in excellent physical condition. Future work should include older adult talkers with greater physiological aging.

Although studies have shown talker intelligibility to be relatively consistent across different listener populations and listening environments, little is known about the extent to which this holds across speaking style. An additional goal of this dissertation was to examine relative talker intelligibility across speaking styles using the results from Experiment 2A. Findings revealed that the extent of relative talker intelligibility largely depended on the speaking style conditions. Strong correlations for relative talker intelligibility were found between speaking styles in quiet (conversational and clear speech in quiet). Weaker, although still significant, correlations were found between speaking styles in noise (conversational and clear speech in noise), and in clear speech (quiet and noise-adapted clear speech). Relative talker intelligibility was not significantly correlated for speaking styles in conversational speech. Adapting speech to noise induced a larger amount of variability in relative talker intelligibility, even for talkers that showed comparable intelligibility levels in quiet. Additionally, the ranking of least/most intelligible talkers showed considerable variability when all 4 speaking styles were examined in conjunction. The cross-style variability illustrated in these analyses suggests that who listeners perceive as an intrinsically "good" vs. "bad" talker may change depending on the communicative context. The cross-style variability in relative talker intelligibility further suggests that talkers may not have consistent intelligibility-enhancing strategies for different adverse communicative situations.

Given the limited knowledge regarding which specific acoustic cues contribute to improved intelligibility, results from Experiment 1 were analyzed in conjunction with results from Experiment 2A to provide insight as to which of the acoustic-phonetic changes most impacted speech intelligibility. Results revealed that spectral and durational cues best predicted the variance in intelligibility. Improved word recognition was associated with increased 1-3 kHz energy, along with a slower speaking rate (including lengthened vowels and pauses). Several of the acoustic-articulatory cues found to characterize intelligibility-enhancing speaking style adaptations (e.g. voice quality) did not necessarily contribute to more intelligible speech. It appears that the process of enhancing intelligibility relies on a set of multiple, covarying acoustic-phonetic changes.

The results of this dissertation expand our knowledge of how children, young adults, and older adults enhance their intelligibility via noise-adapted and clear speaking style adaptations. The results also further our understanding of how intelligibility-enhancing speech adaptations differ based on communicative intent, how relative talker intelligibility varies across different communicative settings, and which acoustic-phonetic changes underlie improved speech intelligibility. The results hold practical implications for the classroom, the clinic, and speech technology. For example, audiologic standards and rehabilitation strategies have typically assumed a large degree of speaker heterogeneity. Expanding our knowledge of talker variability in production and its effects on perception has the potential to contribute to a wide range of clinical applications such as fitting hearing aids and defining rehabilitative criteria. For example, an understanding of how speech intelligibility varies across age and speaking style could improve the design of rehabilitative strategies (e.g. tailoring instruction to the spouse vs. child of an adult with severe hearing loss). Additionally, speech recognition systems could be improved by taking into account these natural age- and adaptation-related variations in human speech production. The discovery of which acoustic properties of speech result in greater intelligibility could also be extremely useful in developing new signal processing algorithms in nonlinear hearing aids.

Appendix

Table 23: List of all sentences (Fallon et al., 2002)

No.	**Target word**	**Sentence**
1	bag	Mom packed my lunch in a bag.
2	ball	We played catch with the ball.
3	barn	Farm animals stay in a barn
4	bed	I fell asleep on my bed.
5	bee	I got stung by a bee.
6	belt	I bought Dad a leather belt
7	book	I like to read a book.
8	boots	When it snows, I put on my boots.
9	bread	Sandwiches are made with bread.
10	broom	I cleaned the floor with a broom.
11	brush	To untangle my hair, I use a brush.
12	bus	Dad rides to work on the bus.
13	cage	I put the bird back in its cage.
14	cake	For dessert, we ate cake.
15	car	We drove to the store in our car.
16	cat	The dog chased the cat.
17	chair	I sat down on the chair.
18	cheese	Mice like to eat cheese.
19	clock	I knew the time when I looked at the clock.
20	cloud	Rain poured from the cloud.
21	clown	We laughed at the funny clown.
22	corn	Farmers plant rows of corn.
23	cow	At the farm, I saw a cow.
24	crown	The king wore a gold crown.
25	cup	I drink juice out of a cup.
26	deer	In the forest, I saw a deer.
27	doll	The girl played with her doll.
28	door	Mom asked me to open the door.
29	dress	She wore a pretty dress.
30	drum	Mike banged on a drum.

31	duck	At the pond, I fed a duck.
32	fan	I was hot, so I turned on the fan.
33	fish	I went to the pond and caught a fish.
34	flag	At the soccer game, I waved my flag.
35	fork	I eat spaghetti with a fork.
36	horse	I learned how to ride a horse.
37	hose	To water the lawn, Dad used the hose.
38	house	Ann's family lives in a house.
39	key	To open the door, Dad used a key.
40	kite	I like to fly my kite.
41	net	Nick catches bugs with a net.
42	nose	The bully punched my nose.
43	pail	We carried the water in a pail.
44	pants	I fell and ripped my pants.
45	phone	I answered the phone.
46	pig	The farmer fed the pig.
47	pin	To hold cloth together, we use a pin.
48	pot	Mom cooks dinner in a pot.
49	rose	I gave my mom a pretty rose
50	shell	At the beach, I found a shell.
51	shoe	I know how to tie a shoe.
52	skunk	An animal that smells bad is a skunk.
53	snake	I got bitten by a snake.
54	snow	I like to play in the snow.
55	soap	We wash our hands with soap.
56	sock	We put the shoe on after the sock
57	star	In the sky, I saw a bright star.
58	tie	When Dad gets dressed up, he wears a tie.
59	tree	A bird built its nest in our tree.
60	wheel	My wagon has a broken wheel.

Table 24: List of tokens for Experiment 1 vowel analyses

No.	**Vowel**	**Sentence**

2	ae	We played *catch* with the ball.
8	u	When it snows, I put on my *boots*.
16	ae	The dog chased the *cat*.
18	i	Mice like to eat *cheese*.
25	u	I drink *juice* out of a cup.
32	ɑ	I was *hot*, so I turned on the fan.
50	i	At the *beach*, I found a shell.
56	ɑ	We put the shoe on after the *sock*.

Table 25: Experiment 1 summary of random effects for the 11 mixed models

No.	Acoustic feature	Parameter		Estimate	Std. Error
1	F0 range	Residual		10452.051	176.157
		Intercept [subject = Sentence]	Variance	474.516	103.433
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	938.038	197.319
			CS covariance	68.921	88.442
2	F0 mean	Residual		180.616	3.044
		Intercept [subject = Sentence]	Variance	16.449	3.306
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	122.394	23.845
			CS covariance	145.087	47.211
3	1-3 kHz energy	Residual		5.020	0.085
		Intercept [subject = Sentence]	Variance	1.784	0.336
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	4.160	0.809
			CS covariance	-0.519	0.211
4	Speaking rate	Residual		0.749	0.013
		Intercept [subject = Sentence]	Variance	0.156	0.030

		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	0.087	0.018
			CS covariance	0.003	0.007
5	Pause duration	Residual		0.011	0.000
		Intercept [subject = Sentence]	Variance	0.002	0.000
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	0.003	0.001
			CS covariance	0.000	0.000
6	Vowel duration	Residual		0.002	0.000
		Intercept [subject = Sentence]	Variance	0.002	0.001
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	0.000	0.000
			CS covariance	0.000	0.000
7	F1	Residual		3993.807	195.695
		Intercept [subject = Sentence]	Variance	826.548	607.994
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	219.874	91.177
			CS covariance	392.039	125.407
8	F2	Residual		53751.970	2633.825
		Intercept [subject = Sentence]	Variance	19858.888	14359.108
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	433.980	748.384
			CS covariance	1648.427	550.837
9	HNR	Residual		8.055	0.805
		Intercept [subject = Sentence]	Variance	9.849	14.023
		Intercept + QN + COCL	CS diagonal offset	0.000	0.000

		[subject = Talker]	CS covariance	0.328	0.120
10	Jitter	Residual		0.000	0.000
		Intercept [subject = Sentence]	Variance	0.000	0.000
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	0.000	0.000
			CS covariance	0.000	0.000
11	Shimmer	Residual		0.002	0.000
		Intercept [subject = Sentence]	Variance	0.000	0.001
		Intercept + QN + COCL [subject = Talker]	CS diagonal offset	0.000	0.000
			CS covariance	0.000	0.000

Table 26: Experiment 1 summary of fixed effects for the 11 mixed models

No.	Acoustic feature	Source	Numerator df	Denominator df	F	Sig.
1	F0 range	QN	1	53.995	6.444	0.014
		COCL	1	53.995	5.627	0.021
		CHYAOA	2	27.000	1.572	0.226
		QN*COCL	1	7041.042	5.060	0.025
		QN*CHYAOA	2	53.995	1.362	0.265
		COCL*CHYAOA	2	53.995	0.960	0.389
		QN*COCL*CHYAOA	2	7041.041	0.613	0.542
2	F0 mean	QN	1	54.001	122.514	0.000
		COCL	1	54.001	1.130	0.292
		CHYAOA	2	27.000	8.830	0.001
		QN*COCL	1	7041.013	8.385	0.004
		QN*CHYAOA	2	54.001	8.184	0.001
		COCL*CHYAOA	2	54.001	0.421	0.659
		QN*COCL*CHYAOA	2	7041.013	9.311	0.000
3	1-3 kHz energy	QN	1	53.997	58.312	0.000
		COCL	1	53.997	3.205	0.079
		CHYAOA	2	27.000	5.444	0.010
		QN*COCL	1	7041.002	416.141	0.000

		QN*CHYAOA	2	53.997	1.116	0.335
		COCL*CHYAOA	2	53.997	0.844	0.436
		QN*COCL*CHYAOA	2	7041.002	98.321	0.000
4	Speaking rate	QN	1	53.980	26.801	0.000
		COCL	1	53.981	135.177	0.000
		CHYAOA	2	27.004	8.526	0.001
		QN*COCL	1	7008.697	21.434	0.000
		QN*CHYAOA	2	53.981	0.949	0.394
		COCL*CHYAOA	2	53.981	3.048	0.056
		QN*COCL*CHYAOA	2	7008.680	1.069	0.343
5	Pause duration	QN	1	54.064	1.122	0.294
		COCL	1	54.064	47.923	0.000
		CHYAOA	2	27.016	1.449	0.252
		QN*COCL	1	7000.855	17.408	0.000
		QN*CHYAOA	2	54.064	0.246	0.783
		COCL*CHYAOA	2	54.064	4.321	0.018
		QN*COCL*CHYAOA	2	7000.839	6.502	0.002
6	Vowel duration	QN	1	54.000	93.220	0.000
		COCL	1	54.000	61.175	0.000
		CHYAOA	2	27.000	4.305	0.024
		QN*COCL	1	860.000	3.861	0.051
		QN*CHYAOA	2	54.000	3.028	0.057
		COCL*CHYAOA	2	54.000	0.048	0.953
		QN*COCL*CHYAOA	2	860.000	2.004	0.135
7	F1	QN	1	54.000	66.935	0.000
		COCL	1	54.000	6.552	0.013
		CHYAOA	2	27.000	9.458	0.001
		QN*COCL	1	860.000	0.002	0.967
		QN*CHYAOA	2	54.000	8.989	0.000
		COCL*CHYAOA	2	54.000	3.313	0.733
		QN*COCL*CHYAOA	2	860.000	0.160	0.852
8	F2	QN	1	54.000	15.827	0.000
		COCL	1	54.000	2.165	0.147
		CHYAOA	2	27.000	3.938	0.032
		QN*COCL	1	860.000	0.126	0.723
		QN*CHYAOA	2	54.000	2.692	0.077
		COCL*CHYAOA	2	54.000	0.442	0.645
		QN*COCL*CHYAOA	2	860.000	0.072	0.931

9	HNR	QN	1	200.000	37.978	0.000
		COCL	1	200.000	5.570	0.019
		CHYAOA	2	27.000	7.375	0.003
		QN*COCL	1	200.000	2.440	0.120
		QN*CHYAOA	2	200.000	0.540	0.584
		COCL*CHYAOA	2	200.000	8.016	0.000
		QN*COCL*CHYAOA	2	200.000	0.477	0.622
10	Jitter	QN	1	200.000	31.591	0.000
		COCL	1	200.000	1.225	0.270
		CHYAOA	2	27.000	4.323	0.024
		QN*COCL	1	200.000	1.191	0.276
		QN*CHYAOA	2	200.000	0.996	0.371
		COCL*CHYAOA	2	200.000	1.660	0.193
		QN*COCL*CHYAOA	2	200.000	0.153	0.858
11	Shimmer	QN	1	200.000	19.676	0.000
		COCL	1	200.000	0.714	0.399
		CHYAOA	2	27.000	2.727	0.083
		QN*COCL	1	200.000	1.095	0.297
		QN*CHYAOA	2	200.000	1.310	0.272
		COCL*CHYAOA	2	200.000	3.425	0.035
		QN*COCL*CHYAOA	2	200.000	0.695	0.500

Table 27: Experiment 1 estimates for the main effects of Talker Age Group

	CHYAOA	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
F0 mean	CH	228.794	12.466	203.219	254.368
	OA	168.704	12.466	143.129	194.278
	YA	161.317	12.466	135.742	186.891
1-3 kHz energy	CH	25.980	0.630	24.696	27.264
	OA	23.181	0.630	21.897	24.465
	YA	24.927	0.630	23.643	26.211
Speaking rate	CH	3.961	0.151	3.654	4.267
	OA	3.158	0.151	2.852	3.464
	YA	3.737	0.151	3.430	4.043
Vowel duration	CH	0.195	0.019	0.155	0.235
	OA	0.234	0.019	0.194	0.274

	YA	0.199	0.019	0.159	0.239
F1	CH	684.704	91.587	472.171	897.237
	OA	569.156	91.587	356.623	781.688
	YA	586.929	91.587	374.396	799.461
F2	CH	1926.736	177.951	1514.541	2338.931
	OA	1765.361	177.951	1353.166	2177.556
	YA	1814.313	177.951	1402.118	2226.508
HNR	CH	11.303	2.307	-10.298	32.905
	OA	8.267	2.307	-13.334	29.869
	YA	8.423	2.307	-13.178	30.025
Jitter	CH	0.022	0.010	-0.062	0.106
	OA	0.033	0.010	-0.051	0.116
	YA	0.035	0.010	-0.048	0.118

Table 28: Experiment 1 pairwise comparisons for the main effects of Talker Age Group

	(I) CHYAOA	(J) CHYAOA	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
F0 mean	CH	OA	60.090*	17.614	0.002	23.948	96.232
		YA	67.477*	17.614	0.001	31.335	103.619
	OA	CH	-60.090*	17.614	0.002	-96.232	-23.948
		YA	7.387	17.614	0.678	-28.755	43.529
	YA	CH	-67.477*	17.614	0.001	-103.619	-31.335
		OA	-7.387	17.614	0.678	-43.529	28.755
1-3 kHz energy	CH	OA	2.799*	0.857	0.003	1.041	4.557
		YA	1.053	0.857	0.230	-0.706	2.811
	OA	CH	-2.799*	0.857	0.003	-4.557	-1.041
		YA	-1.746	0.857	0.051	-3.505	0.012
	YA	CH	-1.053	0.857	0.230	-2.811	0.706
		OA	1.746	0.857	0.051	-0.012	3.505
Speaking rate	CH	OA	.803*	0.201	0.000	0.391	1.214
		YA	0.224	0.201	0.274	-0.188	0.635
	OA	CH	-.803*	0.201	0.000	-1.214	-0.391
		YA	-.579*	0.201	0.008	-0.990	-0.167
	YA	CH	-0.224	0.201	0.274	-0.635	0.188
		OA	.579*	0.201	0.008	0.167	0.990
Vowel	CH	OA	-.039*	0.014	0.013	-0.068	-0.009

duration		YA	-0.004	0.014	0.774	-0.034	0.026
	OA	CH	.039*	0.014	0.013	0.009	0.068
		YA	.035*	0.014	0.024	0.005	0.064
	YA	CH	0.004	0.014	0.774	-0.026	0.034
		OA	-.035*	0.014	0.024	-0.064	-0.005
F1	CH	OA	115.548*	28.611	0.000	56.844	174.253
		YA	97.775*	28.611	0.002	39.071	156.480
	OA	CH	-115.548*	28.611	0.000	-174.253	-56.844
		YA	-17.773	28.611	0.540	-76.478	40.932
	YA	CH	-97.775*	28.611	0.002	-156.480	-39.071
		OA	17.773	28.611	0.540	-40.932	76.478
		OA	3.399	21.566	0.876	-40.850	47.649
F2	CH	OA	161.375*	58.964	0.011	40.391	282.358
		YA	112.423	58.964	0.067	-8.560	233.406
	OA	CH	-161.375*	58.964	0.011	-282.358	-40.391
		YA	-48.952	58.964	0.414	-169.935	72.032
	YA	CH	-112.423	58.964	0.067	-233.406	8.560
		OA	48.952	58.964	0.414	-72.032	169.935
HNR	CH	OA	3.036*	0.890	0.002	1.209	4.863
		YA	2.880*	0.890	0.003	1.053	4.707
	OA	CH	-3.036*	0.890	0.002	-4.863	-1.209
		YA	-0.156	0.890	0.862	-1.983	1.670
	YA	CH	-2.880*	0.890	0.003	-4.707	-1.053
		OA	0.156	0.890	0.862	-1.670	1.983
Jitter	CH	OA	-.011*	0.005	0.033	-0.020	-0.001
		YA	-.013*	0.005	0.010	-0.023	-0.003
	OA	CH	.011*	0.005	0.033	0.001	0.020
		YA	-0.002	0.005	0.617	-0.012	0.007
	YA	CH	.013*	0.005	0.010	0.003	0.023
		OA	0.002	0.005	0.617	-0.007	0.012

Table 29: Experiment 1 estimates for the main effects of Environment-Oriented Speaking Style

	QN	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
F0 range	N	230.509	10.472	209.406	251.612

	Q	209.523	10.472	188.420	230.626
F0 mean	N	202.177	7.352	187.151	217.203
	Q	170.366	7.352	155.340	185.391
1-3 kHz energy	N	26.717	0.471	25.778	27.656
	Q	22.675	0.471	21.736	23.614
Speaking rate	N	3.414	0.104	3.206	3.623
	Q	3.822	0.104	3.614	4.030
Vowel duration	N	0.233	0.017	0.195	0.270
	Q	0.186	0.017	0.148	0.224
F1	N	636.482	90.129	424.819	848.145
	Q	590.710	90.129	379.047	802.374
F2	N	1867.103	174.845	1456.808	2277.398
	Q	1803.836	174.845	1393.541	2214.131
HNR	N	10.460	2.256	-14.960	35.881
	Q	8.202	2.256	-17.218	33.623
Jitter	N	0.023	0.010	-0.080	0.125
	Q	0.037	0.010	-0.065	0.139
Shimmer	N	0.076	0.016	-0.075	0.227
	Q	0.103	0.016	-0.047	0.253

Table 30: Experiment 1 pairwise comparisons for the main effects of Environment-Oriented Speaking Style

	(I) QN	(J) QN	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
F0 range	N	Q	20.986*	8.267	0.014	4.411	37.561
	Q	N	-20.986*	8.267	0.014	-37.561	-4.411
F0 mean	N	Q	31.811*	2.874	0.000	26.049	37.574
	Q	N	-31.811*	2.874	0.000	-37.574	-26.049
1-3 kHz energy	N	Q	4.042*	0.529	0.000	2.980	5.103
	Q	N	-4.042*	0.529	0.000	-5.103	-2.980
Speaking rate	N	Q	-.408*	0.079	0.000	-0.566	-0.250
	Q	N	.408*	0.079	0.000	0.250	0.566
Vowel duration	N	Q	.047*	0.005	0.000	0.037	0.056
	Q	N	-.047*	0.005	0.000	-0.056	-0.037
F1	N	Q	45.771*	5.595	0.000	34.555	56.988
	Q	N	-45.771*	5.595	0.000	-56.988	-34.555

F2	N	Q	63.267*	15.903	0.000	31.384	95.150
	Q	N	-63.267*	15.903	0.000	-95.150	-31.384
HNR	N	Q	2.258*	0.366	0.000	1.535	2.980
	Q	N	-2.258*	0.366	0.000	-2.980	-1.535
Jitter	N	Q	-.015*	0.003	0.000	-0.020	-0.009
	Q	N	.015*	0.003	0.000	0.009	0.020
Shimmer	N	Q	-.027*	0.006	0.000	-0.039	-0.015
	Q	N	.027*	0.006	0.000	0.015	0.039

Table 31: Experiment 1 estimates for the main effects of Listener-Oriented Speaking Style

	COCL	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
F0 range	CL	229.821	10.472	208.719	250.924
	CO	210.211	10.472	189.108	231.313
Speaking rate	CL	3.160	0.104	2.952	3.369
	CO	4.076	0.104	3.868	4.284
Pause duration	CL	0.119	0.015	0.089	0.148
	CO	0.021	0.015	-0.009	0.050
Vowel duration	CL	0.228	0.017	0.191	0.266
	CO	0.191	0.017	0.153	0.228
F1	CL	620.756	90.129	409.093	832.419
	CO	606.436	90.129	394.773	818.099
HNR	CL	9.764	2.256	-15.657	35.184
	CO	8.899	2.256	-16.521	34.319

Table 32: Experiment 1 pairwise comparisons for the main effects of Listener-Oriented Speaking Style

	(I) COCL	(J) COCL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
F0 range	CL	CO	19.611*	8.267	0.021	3.036	36.186
	CO	CL	-19.611*	8.267	0.021	-36.186	-3.036
Speaking rate	CL	CO	-.916*	0.079	0.000	-1.074	-0.758
	CO	CL	.916*	0.079	0.000	0.758	1.074

Pause duration	CL	CO	.098*	0.014	0.000	0.070	0.126
	CO	CL	-.098*	0.014	0.000	-0.126	-0.070
Vowel duration	CL	CO	.038*	0.005	0.000	0.028	0.048
	CO	CL	-.038*	0.005	0.000	-0.048	-0.028
F1	CL	CO	14.321*	5.595	0.013	3.104	25.537
	CO	CL	-14.321*	5.595	0.013	-25.537	-3.104
HNR	CL	CO	.865*	0.366	0.019	0.142	1.587
	CO	CL	-.865*	0.366	0.019	-1.587	-0.142

Table 33: Experiment 1 estimates for the interactions between Talker Age Group and Environment- Oriented Speaking Style

	QN	CHYAOA	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
F0 mean	N	CH	252.652	12.712	226.663	278.641
		OA	178.825	12.712	152.836	204.814
		YA	175.054	12.712	149.065	201.043
	Q	CH	204.936	12.712	178.947	230.925
		OA	158.582	12.712	132.593	184.571
		YA	147.579	12.712	121.590	173.568
Vowel duration	N	CH	0.227	0.019	0.186	0.267
		OA	0.251	0.019	0.211	0.292
		YA	0.22	0.019	0.180	0.261
	Q	CH	0.164	0.019	0.123	0.204
		OA	0.216	0.019	0.176	0.257
		YA	0.178	0.019	0.138	0.219
F1	N	CH	724.362	91.715	511.747	936.977
		OA	583.485	91.715	370.870	796.100
		YA	601.598	91.715	388.983	814.213
	Q	CH	645.046	91.715	432.431	857.661
		OA	554.826	91.715	342.211	767.441
		YA	572.259	91.715	359.644	784.874

Table 34: Experiment 1 pairwise comparisons for the interactions between Talker Age Group and Environment-Oriented Speaking Style

	QN	(I) CHYAOA	(J) CHYAOA	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
F0 mean	N	CH	OA	73.827*	17.963	0.000	37.099	110.555
			YA	77.597*	17.963	0.000	40.869	114.325
		OA	CH	-73.827*	17.963	0.000	-110.555	-37.099
			YA	3.771	17.963	0.835	-32.957	40.499
		YA	CH	-77.597*	17.963	0.000	-114.325	-40.869
			OA	-3.771	17.963	0.835	-40.499	32.957
	Q	CH	OA	46.354*	17.963	0.015	9.626	83.082
			YA	57.357*	17.963	0.003	20.629	94.085
		OA	CH	-46.354*	17.963	0.015	-83.082	-9.626
			YA	11.003	17.963	0.545	-25.725	47.731
		YA	CH	-57.357*	17.963	0.003	-94.085	-20.629
			OA	-11.003	17.963	0.545	-47.731	25.725
Vowel duration	N	CH	OA	-0.025	0.016	0.123	-0.056	0.007
			YA	0.006	0.016	0.693	-0.025	0.038
		OA	CH	0.025	0.016	0.123	-0.007	0.056
			YA	0.031	0.016	0.056	-0.001	0.063
		YA	CH	-0.006	0.016	0.693	-0.038	0.025
			OA	-0.031	0.016	0.056	-0.063	0.001
	Q	CH	OA	-.053*	0.016	0.002	-0.085	-0.021
			YA	-0.015	0.016	0.356	-0.046	0.017
		OA	CH	.053*	0.016	0.002	0.021	0.085
			YA	.038*	0.016	0.020	0.006	0.070
		YA	CH	0.015	0.016	0.356	-0.017	0.046
			OA	-.038*	0.016	0.020	-0.070	-0.006
F1	N	CH	OA	140.877*	29.42	0.000	80.805	200.949
			YA	122.763*	29.42	0.000	62.691	182.835
		OA	CH	-140.877*	29.42	0.000	-200.949	-80.805
			YA	-18.113	29.42	0.543	-78.185	41.959
		YA	CH	-122.763*	29.42	0.000	-182.835	-62.691

			OA	18.113	29.42	0.543	-41.959	78.185
	Q	CH	OA	90.220*	29.42	0.005	30.148	150.292
			YA	72.788*	29.42	0.019	12.715	132.860
		OA	CH	-90.220*	29.42	0.005	-150.292	-30.148
			YA	-17.432	29.42	0.558	-77.505	42.640
		YA	CH	-72.788*	29.42	0.019	-132.860	-12.715
			OA	17.432	29.42	0.558	-42.640	77.505

Table 35: Experiment 1 pairwise comparisons for the interactions between Talker Age Group and Environment-Oriented Speaking Style

	CHYAOA	(I) QN	(J) QN	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
F0 mean	CH	N	Q	47.716*	4.978	0.000	37.736	57.696
		Q	N	-47.716*	4.978	0.000	-57.696	-37.736
	OA	N	Q	20.243*	4.978	0.000	10.263	30.223
		Q	N	-20.243*	4.978	0.000	-30.223	-10.263
	YA	N	Q	27.476*	4.978	0.000	17.495	37.456
		Q	N	-27.476*	4.978	0.000	-37.456	-17.495
Vowel duration	CH	N	Q	.063*	0.008	0.000	0.046	0.080
		Q	N	-.063*	0.008	0.000	-0.080	-0.046
	OA	N	Q	.035*	0.008	0.000	0.018	0.052
		Q	N	-.035*	0.008	0.000	-0.052	-0.018
	YA	N	Q	.042*	0.008	0.000	0.025	0.059
		Q	N	-.042*	0.008	0.000	-0.059	-0.025
F1	CH	N	Q	79.316*	9.69	0.000	59.888	98.743
		Q	N	-79.316*	9.69	0.000	-98.743	-59.888
	OA	N	Q	28.659*	9.69	0.005	9.231	48.086
		Q	N	-28.659*	9.69	0.005	-48.086	-9.231
	YA	N	Q	29.340*	9.69	0.004	9.912	48.767
		Q	N	-29.340*	9.69	0.004	-48.767	-9.912

Table 36: Experiment 1 estimates for the interactions between Talker Age Group and Listener-Oriented Speaking Style

	COCL	CHYAOA	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Speaking rate	CL	CH	3.639	0.165	3.306	3.972
		OA	2.647	0.165	2.314	2.979
		YA	3.196	0.165	2.863	3.528
	CO	CH	4.282	0.165	3.949	4.615
		OA	3.669	0.165	3.336	4.002
		YA	4.278	0.165	3.945	4.610
Pause duration	CL	CH	0.086	0.024	0.037	0.135
		OA	0.174	0.024	0.126	0.223
		YA	0.095	0.024	0.047	0.144
	CO	CH	0.032	0.024	-0.017	0.081
		OA	0.021	0.024	-0.028	0.070
		YA	0.009	0.024	-0.040	0.058
HNR	CL	CH	12.768	2.328	-7.516	33.052
		OA	8.273	2.328	-12.011	28.557
		YA	8.25	2.328	-12.034	28.534
	CO	CH	9.839	2.328	-10.445	30.123
		OA	8.262	2.328	-12.022	28.545
		YA	8.597	2.328	-11.687	28.881
Shimmer	CL	CH	0.062	0.018	-0.030	0.154
		OA	0.093	0.018	0.001	0.185
		YA	0.105	0.018	0.013	0.197
	CO	CH	0.088	0.018	-0.004	0.179
		OA	0.097	0.018	0.006	0.188
		YA	0.091	0.018	0.001	0.181

Table 37: Experiment 1 pairwise comparisons for the interactions between Talker Age Group and Listener-Oriented Speaking Style

	COCL	(I) CHYAOA	(J) CHYAOA	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
Speaking rate	CL	CH	OA	.992*	0.223	0.000	0.542	1.442
			YA	0.443	0.223	0.053	-0.006	0.893
		OA	CH	-.992*	0.223	0.000	-1.442	-0.542

			YA	-.549*	0.223	0.018	-0.999	-0.099
		YA	CH	-0.443	0.223	0.053	-0.893	0.006
			OA	.549*	0.223	0.018	0.099	0.999
	CO	CH	OA	.613*	0.223	0.009	0.163	1.063
			YA	0.004	0.223	0.985	-0.446	0.454
		OA	CH	-.613*	0.223	0.009	-1.063	-0.163
			YA	-.609*	0.223	0.009	-1.059	-0.159
		YA	CH	-0.004	0.223	0.985	-0.454	0.446
			OA	.609*	0.223	0.009	0.159	1.059
Pause duration	CL	CH	OA	-.089*	0.034	0.011	-0.156	-0.021
			YA	-0.01	0.034	0.774	-0.077	0.058
		OA	CH	.089*	0.034	0.011	0.021	0.156
			YA	.079*	0.034	0.023	0.012	0.147
		YA	CH	0.01	0.034	0.774	-0.058	0.077
			OA	-.079*	0.034	0.023	-0.147	-0.012
	CO	CH	OA	0.011	0.034	0.745	-0.057	0.079
			YA	0.022	0.034	0.508	-0.045	0.090
		OA	CH	-0.011	0.034	0.745	-0.079	0.057
			YA	0.011	0.034	0.736	-0.056	0.079
		YA	CH	-0.022	0.034	0.508	-0.090	0.045
			OA	-0.011	0.034	0.736	-0.079	0.056
HNR	CL	CH	OA	4.495*	0.997	0.000	2.483	6.507
			YA	4.518*	0.997	0.000	2.506	6.530
		OA	CH	-4.495*	0.997	0.000	-6.507	-2.483
			YA	0.023	0.997	0.982	-1.989	2.034
		YA	CH	-4.518*	0.997	0.000	-6.530	-2.506
			OA	-0.023	0.997	0.982	-2.034	1.989
	CO	CH	OA	1.577	0.997	0.121	-0.435	3.589
			YA	1.242	0.997	0.220	-0.770	3.254
		OA	CH	-1.577	0.997	0.121	-3.589	0.435
			YA	-0.335	0.997	0.738	-2.347	1.677
		YA	CH	-1.242	0.997	0.220	-3.254	0.770
			OA	0.335	0.997	0.738	-1.677	2.347
Shimmer	CL	CH	OA	-.031*	0.013	0.020	-0.058	-0.005
			YA	-.043*	0.013	0.002	-0.069	-0.016
		OA	CH	.031*	0.013	0.020	0.005	0.058
			YA	-0.011	0.013	0.395	-0.038	0.015
		YA	CH	.043*	0.013	0.002	0.016	0.069

			OA	0.011	0.013	0.395	-0.015	0.038
	CO	CH	OA	-0.01	0.013	0.472	-0.036	0.017
			YA	-0.003	0.013	0.795	-0.030	0.023
		OA	CH	0.01	0.013	0.472	-0.017	0.036
			YA	0.006	0.013	0.650	-0.021	0.033
		YA	CH	0.003	0.013	0.795	-0.023	0.030
			OA	-0.006	0.013	0.650	-0.033	0.021

Table 38: Experiment 1 pairwise comparisons for the interactions between Talker Age Group and Listener-Oriented Speaking Style

	CHYAOA	(I) COCL	(J) COCL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
Speaking rate	CH	CL	CO	-.643*	0.136	0.000	-0.916	-0.370
		CO	CL	.643*	0.136	0.000	0.370	0.916
	OA	CL	CO	-1.022*	0.136	0.000	-1.296	-0.749
		CO	CL	1.022*	0.136	0.000	0.749	1.296
	YA	CL	CO	-1.082*	0.136	0.000	-1.356	-0.809
		CO	CL	1.082*	0.136	0.000	0.809	1.356
Pause duration	CH	CL	CO	.054*	0.024	0.032	0.005	0.103
		CO	CL	-.054*	0.024	0.032	-0.103	-0.005
	OA	CL	CO	.154*	0.024	0.000	0.105	0.203
		CO	CL	-.154*	0.024	0.000	-0.203	-0.105
	YA	CL	CO	.086*	0.024	0.001	0.037	0.135
		CO	CL	-.086*	0.024	0.001	-0.135	-0.037
HNR	CH	CL	CO	2.929*	0.635	0.000	1.678	4.181
		CO	CL	-2.929*	0.635	0.000	-4.181	-1.678
	OA	CL	CO	0.011	0.635	0.986	-1.240	1.263
		CO	CL	-0.011	0.635	0.986	-1.263	1.240
	YA	CL	CO	-0.346	0.635	0.586	-1.598	0.905
		CO	CL	0.346	0.635	0.586	-0.905	1.598
Shimmer	CH	CL	CO	-.026*	0.011	0.016	-0.046	-0.005
		CO	CL	.026*	0.011	0.016	0.005	0.046
	OA	CL	CO	-0.004	0.011	0.730	-0.025	0.017
		CO	CL	0.004	0.011	0.730	-0.017	0.025
	YA	CL	CO	0.014	0.011	0.203	-0.007	0.035
		CO	CL	-0.014	0.011	0.203	-0.035	0.007

Table 39: Experiment 1 estimates for the interactions between Environment- and Listener-Oriented Speaking Style

	QN	COCL	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
F0 range	N	CL	243.026	11.323	220.353	265.7
		CO	217.992	11.323	195.318	240.666
	Q	CL	216.617	11.323	193.943	239.29
		CO	202.429	11.322	179.756	225.103
F0 mean	N	CL	203.246	7.493	187.979	218.513
		CO	201.108	7.493	185.841	216.375
	Q	CL	172.352	7.493	157.085	187.62
		CO	168.379	7.493	153.111	183.646
1-3 kHz energy	N	CL	26.652	0.541	25.577	27.726
		CO	26.782	0.541	25.707	27.857
	Q	CL	23.688	0.541	22.613	24.763
		CO	21.663	0.541	20.588	22.737
Speaking rate	N	CL	3.004	0.112	2.781	3.227
		CO	3.825	0.112	3.602	4.048
	Q	CL	3.317	0.112	3.094	3.54
		CO	4.328	0.112	4.105	4.55
Pause duration	N	CL	0.131	0.016	0.098	0.164
		CO	0.023	0.016	-0.01	0.056
	Q	CL	0.106	0.016	0.073	0.139
		CO	0.018	0.016	-0.014	0.051
Vowel duration	N	CL	0.249	0.017	0.211	0.287
		CO	0.217	0.017	0.179	0.254
	Q	CL	0.208	0.017	0.17	0.246
		CO	0.164	0.017	0.127	0.202

Table 40: Experiment 1 pairwise comparisons for the interactions between Environment- and Listener- Oriented Speaking Style

	QN	(I) COCL	(J) COCL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
F0 range	N	CL	CO	25.034*	8.612	0.005	7.828	42.240
		CO	CL	-25.034*	8.612	0.005	-42.240	-7.828

	Q	CL	CO	14.188	8.612	0.104	-3.018	31.393
		CO	CL	-14.188	8.612	0.104	-31.393	3.018
F0 mean	N	CL	CO	2.138	2.891	0.463	-3.656	7.932
		CO	CL	-2.138	2.891	0.463	-7.932	3.656
	Q	CL	CO	3.973	2.891	0.175	-1.820	9.767
		CO	CL	-3.973	2.891	0.175	-9.767	1.820
1-3 kHz energy	N	CL	CO	-0.13	0.532	0.807	-1.196	0.935
		CO	CL	0.13	0.532	0.807	-0.935	1.196
	Q	CL	CO	2.025*	0.532	0.000	0.959	3.091
		CO	CL	-2.025*	0.532	0.000	-3.091	-0.959
Speaking rate	N	CL	CO	-.821*	0.081	0.000	-0.984	-0.658
		CO	CL	.821*	0.081	0.000	0.658	0.984
	Q	CL	CO	-1.011*	0.081	0.000	-1.173	-0.848
		CO	CL	1.011*	0.081	0.000	0.848	1.173
Pause duration	N	CL	CO	.108*	0.014	0.000	0.079	0.137
		CO	CL	-.108*	0.014	0.000	-0.137	-0.079
	Q	CL	CO	.088*	0.014	0.000	0.059	0.116
		CO	CL	-.088*	0.014	0.000	-0.116	-0.059
Vowel duration	N	CL	CO	.032*	0.006	0.000	0.021	0.043
		CO	CL	-.032*	0.006	0.000	-0.043	-0.021
	Q	CL	CO	.043*	0.006	0.000	0.032	0.055
		CO	CL	-.043*	0.006	0.000	-0.055	-0.032

Table 41: Experiment 1 pairwise comparisons for the interactions between Environment- and Listener- Oriented Speaking Style

	COCL	(I) QN	(J) QN	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
F0 range	CL	N	Q	26.409*	8.612	0.003	9.203	43.616
		Q	N	-26.409*	8.612	0.003	-43.616	-9.203
	CO	N	Q	15.563	8.612	0.075	-1.643	32.769
		Q	N	-15.563	8.612	0.075	-32.769	1.643
F0 mean	CL	N	Q	30.894*	2.891	0.000	25.100	36.688
		Q	N	-30.894*	2.891	0.000	-36.688	-25.100
	CO	N	Q	32.729*	2.891	0.000	26.935	38.523
		Q	N	-32.729*	2.891	0.000	-38.523	-26.935
1-3 kHz energy	CL	N	Q	2.964*	0.532	0.000	1.898	4.030
		Q	N	-2.964*	0.532	0.000	-4.030	-1.898

	CO	N	Q	5.119*	0.532	0.000	4.054	6.185
		Q	N	-5.119*	0.532	0.000	-6.185	-4.054
Speaking rate	CL	N	Q	-.313*	0.081	0.000	-0.476	-0.150
		Q	N	.313*	0.081	0.000	0.150	0.476
	CO	N	Q	-.503*	0.081	0.000	-0.665	-0.340
		Q	N	.503*	0.081	0.000	0.340	0.665
Pause duration	CL	N	Q	0.025	0.014	0.085	-0.004	0.054
		Q	N	-0.025	0.014	0.085	-0.054	0.004
	CO	N	Q	0.005	0.014	0.740	-0.024	0.034
		Q	N	-0.005	0.014	0.740	-0.034	0.024
Vowel duration	CL	N	Q	.041*	0.006	0.000	0.030	0.052
		Q	N	-.041*	0.006	0.000	-0.052	-0.030
	CO	N	Q	.052*	0.006	0.000	0.041	0.063
		Q	N	-.052*	0.006	0.000	-0.063	-0.041

Table 42: Experiment 1 estimates for the interactions between Talker Age Group, Environment- Oriented Speaking Style, and Listener-Oriented Speaking Style

	QN	COCL	CHYAOA	Mean	Std. Error	95% Confidence Interval	
						Lower Bound	Upper Bound
F0 mean	N	CL	CH	256.297	12.957	229.889	282.704
			OA	177.699	12.957	151.292	204.107
			YA	175.742	12.957	149.334	202.149
		CO	CH	249.006	12.957	222.599	275.414
			OA	179.951	12.957	153.543	206.358
			YA	174.367	12.957	147.959	200.774
	Q	CL	CH	207.959	12.957	181.552	234.367
			OA	160.157	12.957	133.750	186.565
			YA	148.94	12.957	122.533	175.348
		CO	CH	201.912	12.957	175.505	228.320
			OA	157.007	12.957	130.599	183.414
			YA	146.217	12.957	119.810	172.625
1-3 kHz energy	N	CL	CH	28.034	0.905	26.233	29.835
			OA	24.557	0.905	22.756	26.357
			YA	27.364	0.905	25.563	29.165
		CO	CH	28.325	0.905	26.524	30.126
			OA	24.75	0.905	22.950	26.551

			YA	27.27	0.905	25.470	29.071
	Q	CL	CH	25.703	0.905	23.902	27.504
			OA	21.898	0.905	20.097	23.699
			YA	23.463	0.905	21.662	25.264
		CO	CH	21.858	0.905	20.057	23.659
			OA	21.519	0.905	19.718	23.319
			YA	21.611	0.905	19.811	23.412
Pause duration	N	CL	CH	0.089	0.027	0.035	0.144
			OA	0.2	0.027	0.145	0.254
			YA	0.104	0.027	0.050	0.159
		CO	CH	0.034	0.027	-0.021	0.088
			OA	0.024	0.027	-0.031	0.078
			YA	0.012	0.027	-0.043	0.066
	Q	CL	CH	0.082	0.027	0.028	0.137
			OA	0.149	0.027	0.094	0.204
			YA	0.087	0.027	0.032	0.141
		CO	CH	0.03	0.027	-0.025	0.085
			OA	0.018	0.027	-0.037	0.072
			YA	0.007	0.027	-0.047	0.062

Table 43: Experiment 1 pairwise comparisons for the interactions between Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style

	CHYAOA	QN	(I) COCL	(J) COCL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
								Lower Bound	Upper Bound
F0 Mean	CH	N	CL	CO	7.291	5.008	0.151	-2.745	17.326
			CO	CL	-7.291	5.008	0.151	-17.326	2.745
		Q	CL	CO	6.047	5.008	0.232	-3.988	16.082
			CO	CL	-6.047	5.008	0.232	-16.082	3.988
	OA	N	CL	CO	-2.252	5.008	0.655	-12.287	7.784
			CO	CL	2.252	5.008	0.655	-7.784	12.287
		Q	CL	CO	3.15	5.008	0.532	-6.885	13.185
			CO	CL	-3.15	5.008	0.532	-13.185	6.885
	YA	N	CL	CO	1.375	5.008	0.785	-8.66	11.41
			CO	CL	-1.375	5.008	0.785	-11.41	8.66
		Q	CL	CO	2.723	5.008	0.589	-7.312	12.758
			CO	CL	-2.723	5.008	0.589	-12.758	7.312

1-3 kHz energy	CH	N	CL	CO	-0.291	0.921	0.753	-2.138	1.555
			CO	CL	0.291	0.921	0.753	-1.555	2.138
		Q	CL	CO	3.845*	0.921	0.000	1.999	5.691
			CO	CL	-3.845*	0.921	0.000	-5.691	-1.999
	OA	N	CL	CO	-0.194	0.921	0.834	-2.04	1.653
			CO	CL	0.194	0.921	0.834	-1.653	2.04
		Q	CL	CO	0.38	0.921	0.682	-1.467	2.226
			CO	CL	-0.38	0.921	0.682	-2.226	1.467
	YA	N	CL	CO	0.094	0.921	0.919	-1.752	1.94
			CO	CL	-0.094	0.921	0.919	-1.94	1.752
		Q	CL	CO	1.851*	0.921	0.049	0.005	3.698
			CO	CL	-1.851*	0.921	0.049	-3.698	-0.005
Pause duration	CH	N	CL	CO	.056*	0.025	0.029	0.006	0.105
			CO	CL	-.056*	0.025	0.029	-0.105	-0.006
		Q	CL	CO	.052*	0.025	0.040	0.003	0.102
			CO	CL	-.052*	0.025	0.040	-0.102	-0.003
	OA	N	CL	CO	.176*	0.025	0.000	0.126	0.226
			CO	CL	-.176*	0.025	0.000	-0.226	-0.126
		Q	CL	CO	.131*	0.025	0.000	0.082	0.181
			CO	CL	-.131*	0.025	0.000	-0.181	-0.082
	YA	N	CL	CO	.093*	0.025	0.000	0.043	0.142
			CO	CL	-.093*	0.025	0.000	-0.142	-0.043
		Q	CL	CO	.079*	0.025	0.002	0.03	0.129
			CO	CL	-.079*	0.025	0.002	-0.129	-0.03

Table 44: Experiment 1 pairwise comparisons for the interactions between Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style

	CHYAOA	COCL	(I) QN	(J) QN	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
								Lower Bound	Upper Bound
F0 Mean	CH	CL	N	Q	48.338*	5.008	0.000	38.303	58.373
			Q	N	-48.338*	5.008	0.000	-58.373	-38.303
		CO	N	Q	47.094*	5.008	0.000	37.059	57.129
			Q	N	-47.094*	5.008	0.000	-57.129	-37.059
	OA	CL	N	Q	17.542*	5.008	0.001	7.507	27.577
			Q	N	-17.542*	5.008	0.001	-27.577	-7.507
		CO	N	Q	22.944*	5.008	0.000	12.909	32.979

			Q	N	-22.944*	5.008	0.000	-32.979	-12.909
	YA	CL	N	Q	26.801*	5.008	0.000	16.766	36.837
			Q	N	-26.801*	5.008	0.000	-36.837	-16.766
		CO	N	Q	28.150*	5.008	0.000	18.114	38.185
			Q	N	-28.150*	5.008	0.000	-38.185	-18.114
1-3 kHz energy	CH	CL	N	Q	2.331*	0.921	0.014	0.485	4.177
			Q	N	-2.331*	0.921	0.014	-4.177	-0.485
		CO	N	Q	6.467*	0.921	0.000	4.621	8.314
			Q	N	-6.467*	0.921	0.000	-8.314	-4.621
	OA	CL	N	Q	2.659*	0.921	0.006	0.812	4.505
			Q	N	-2.659*	0.921	0.006	-4.505	-0.812
		CO	N	Q	3.232*	0.921	0.001	1.386	5.078
			Q	N	-3.232*	0.921	0.001	-5.078	-1.386
	YA	CL	N	Q	3.902*	0.921	0.000	2.055	5.748
			Q	N	-3.902*	0.921	0.000	-5.748	-2.055
		CO	N	Q	5.659*	0.921	0.000	3.813	7.505
			Q	N	-5.659*	0.921	0.000	-7.505	-3.813
Pause duration	CH	CL	N	Q	0.007	0.025	0.785	-0.043	0.057
			Q	N	-0.007	0.025	0.785	-0.057	0.043
		CO	N	Q	0.004	0.025	0.887	-0.046	0.053
			Q	N	-0.004	0.025	0.887	-0.053	0.046
	OA	CL	N	Q	.051*	0.025	0.045	0.001	0.101
			Q	N	-.051*	0.025	0.045	-0.101	-0.001
		CO	N	Q	0.006	0.025	0.803	-0.044	0.056
			Q	N	-0.006	0.025	0.803	-0.056	0.044
	YA	CL	N	Q	0.018	0.025	0.477	-0.032	0.068
			Q	N	-0.018	0.025	0.477	-0.068	0.032
		CO	N	Q	0.005	0.025	0.854	-0.045	0.054
			Q	N	-0.005	0.025	0.854	-0.054	0.045

Table 45: Experiment 1 pairwise comparisons for the interactions between Talker Age Group, Environment-Oriented Speaking Style, and Listener-Oriented Speaking Style

	QN	COCL	(I) CHYAOA	(J) CHYAOA	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
								Lower Bound	Upper Bound
F0 mean	N	CL	CH	OA	78.598*	18.308	0.000	41.277	115.918
				YA	80.555*	18.308	0.000	43.235	117.875

			OA	CH	-78.598*	18.308	0.000	-115.918	-41.277
				YA	1.957	18.308	0.916	-35.363	39.278
			YA	CH	-80.555*	18.308	0.000	-117.875	-43.235
				OA	-1.957	18.308	0.916	-39.278	35.363
		CO	CH	OA	69.056*	18.308	0.001	31.735	106.376
				YA	74.639*	18.308	0.000	37.319	111.960
			OA	CH	-69.056*	18.308	0.001	-106.376	-31.735
				YA	5.584	18.308	0.762	-31.736	42.904
			YA	CH	-74.639*	18.308	0.000	-111.960	-37.319
				OA	-5.584	18.308	0.762	-42.904	31.736
	Q	CL	CH	OA	47.802*	18.308	0.014	10.482	85.122
				YA	59.019*	18.308	0.003	21.698	96.339
			OA	CH	-47.802*	18.308	0.014	-85.122	-10.482
				YA	11.217	18.308	0.545	-26.103	48.537
			YA	CH	-59.019*	18.308	0.003	-96.339	-21.698
				OA	-11.217	18.308	0.545	-48.537	26.103
		CO	CH	OA	44.905*	18.308	0.020	7.585	82.225
				YA	55.695*	18.308	0.005	18.375	93.015
			OA	CH	-44.905*	18.308	0.020	-82.225	-7.585
				YA	10.79	18.308	0.560	-26.530	48.110
			YA	CH	-55.695*	18.308	0.005	-93.015	-18.375
				OA	-10.79	18.308	0.560	-48.110	26.530
1-3 kHz energy	N	CL	CH	OA	3.477*	1.257	0.007	0.974	5.980
				YA	0.67	1.257	0.596	-1.833	3.172
			OA	CH	-3.477*	1.257	0.007	-5.980	-0.974
				YA	-2.808*	1.257	0.028	-5.311	-0.305
			YA	CH	-0.67	1.257	0.596	-3.172	1.833
				OA	2.808*	1.257	0.028	0.305	5.311
		CO	CH	OA	3.575*	1.257	0.006	1.072	6.078
				YA	1.055	1.257	0.404	-1.448	3.558
			OA	CH	-3.575*	1.257	0.006	-6.078	-1.072
				YA	-2.520*	1.257	0.048	-5.023	-0.017
			YA	CH	-1.055	1.257	0.404	-3.558	1.448
				OA	2.520*	1.257	0.048	0.017	5.023
	Q	CL	CH	OA	3.805*	1.257	0.003	1.302	6.308
				YA	2.24	1.257	0.079	-0.263	4.743
			OA	CH	-3.805*	1.257	0.003	-6.308	-1.302
				YA	-1.565	1.257	0.217	-4.068	0.938

			YA	CH	-2.24	1.257	0.079	-4.743	0.263
				OA	1.565	1.257	0.217	-0.938	4.068
		CO	CH	OA	0.339	1.257	0.788	-2.164	2.842
				YA	0.246	1.257	0.845	-2.256	2.749
			OA	CH	-0.339	1.257	0.788	-2.842	2.164
				YA	-0.093	1.257	0.941	-2.596	2.410
			YA	CH	-0.246	1.257	0.845	-2.749	2.256
				OA	0.093	1.257	0.941	-2.410	2.596
Pause duration	N	CL	CH	OA	-.111*	0.038	0.005	-0.186	-0.035
				YA	-0.015	0.038	0.690	-0.091	0.060
			OA	CH	.111*	0.038	0.005	0.035	0.186
				YA	.096*	0.038	0.014	0.020	0.171
			YA	CH	0.015	0.038	0.690	-0.060	0.091
				OA	-.096*	0.038	0.014	-0.171	-0.020
		CO	CH	OA	0.01	0.038	0.800	-0.066	0.085
				YA	0.022	0.038	0.566	-0.054	0.098
			OA	CH	-0.01	0.038	0.800	-0.085	0.066
				YA	0.012	0.038	0.748	-0.063	0.088
			YA	CH	-0.022	0.038	0.566	-0.098	0.054
				OA	-0.012	0.038	0.748	-0.088	0.063
	Q	CL	CH	OA	-0.067	0.038	0.083	-0.142	0.009
				YA	-0.004	0.038	0.912	-0.080	0.072
			OA	CH	0.067	0.038	0.083	-0.009	0.142
				YA	0.062	0.038	0.104	-0.013	0.138
			YA	CH	0.004	0.038	0.912	-0.072	0.080
				OA	-0.062	0.038	0.104	-0.138	0.013
		CO	CH	OA	0.012	0.038	0.746	-0.063	0.088
				YA	0.023	0.038	0.548	-0.053	0.099
			OA	CH	-0.012	0.038	0.746	-0.088	0.063
				YA	0.011	0.038	0.781	-0.065	0.086
			YA	CH	-0.023	0.038	0.548	-0.099	0.053
				OA	-0.011	0.038	0.781	-0.086	0.065

Table 46: Experiment 1 estimates for Pairwise comparisons for the interactions between Vowel and Talker Age Group

	CHYAOA	Vowel	Mean	Std. Error	df	95% Confidence Interval

						Lower Bound	Upper Bound
F1	CH	a	901.017	29.326	14.112	838.166	963.869
		ae	967.151	29.326	14.112	904.3	1030.002
		i	415.901	29.326	14.112	353.05	478.752
		u	454.746	29.326	14.112	391.895	517.597
	OA	a	820.17	29.326	14.112	757.319	883.021
		ae	750.446	29.326	14.112	687.595	813.297
		i	336.246	29.326	14.112	273.395	399.097
		u	369.76	29.326	14.112	306.909	432.611
	YA	a	822.108	29.326	14.112	759.256	884.959
		ae	830.916	29.326	14.112	768.065	893.767
		i	321.531	29.326	14.112	258.68	384.382
		u	373.159	29.326	14.112	310.308	436.01
F2	CH	a	1370.214	110.326	5.723	1097.063	1643.364
		ae	1849.965	110.326	5.723	1576.814	2123.116
		i	2618.749	110.326	5.723	2345.598	2891.899
		u	1868.015	110.326	5.723	1594.864	2141.166
	OA	a	1221.338	110.326	5.723	948.187	1494.488
		ae	1840.831	110.326	5.723	1567.681	2113.982
		i	2419.822	110.326	5.723	2146.672	2692.973
		u	1579.453	110.326	5.723	1306.302	1852.603
	YA	a	1244.048	110.326	5.723	970.897	1517.198
		ae	1760.129	110.326	5.723	1486.978	2033.279
		i	2536.425	110.326	5.723	2263.274	2809.576
		u	1716.649	110.326	5.723	1443.499	1989.8

Table 47: Experiment 1 pairwise comparisons for the interactions between Vowel and Talker Age Group

	Vowel	(I) CHYAOA	(J) CHYAOA	Mean Difference (I-J)	Std. Error	df	Sig.	95% Confidence Interval for Difference	
								Lower Bound	Upper Bound
F1	a	CH	OA	80.848*	29.891	32.157	0.011	19.973	141.722
			YA	78.910*	29.891	32.157	0.013	18.036	139.784
		OA	CH	-80.848*	29.891	32.157	0.011	-141.722	-19.973
			YA	-1.938	29.891	32.157	0.949	-62.812	58.937
		YA	CH	-78.910*	29.891	32.157	0.013	-	-18.036

								139.784	
			OA	1.938	29.891	32.157	0.949	-58.937	62.812
	ae	CH	OA	216.705*	29.891	32.157	0.000	155.831	277.579
			YA	136.235*	29.891	32.157	0.000	75.361	197.109
		OA	CH	-216.705*	29.891	32.157	0.000	- 277.579	- 155.831
			YA	-80.470*	29.891	32.157	0.011	- 141.344	-19.596
		YA	CH	-136.235*	29.891	32.157	0.000	- 197.109	-75.361
			OA	80.470*	29.891	32.157	0.011	19.596	141.344
	i	CH	OA	79.655*	29.891	32.157	0.012	18.781	140.529
			YA	94.370*	29.891	32.157	0.003	33.496	155.244
		OA	CH	-79.655*	29.891	32.157	0.012	- 140.529	-18.781
			YA	14.715	29.891	32.157	0.626	-46.159	75.589
		YA	CH	-94.370*	29.891	32.157	0.003	- 155.244	-33.496
			OA	-14.715	29.891	32.157	0.626	-75.589	46.159
	u	CH	OA	84.986*	29.891	32.157	0.008	24.112	145.86
			YA	81.587*	29.891	32.157	0.010	20.713	142.461
		OA	CH	-84.986*	29.891	32.157	0.008	-145.86	-24.112
			YA	-3.399	29.891	32.157	0.910	-64.273	57.475
		YA	CH	-81.587*	29.891	32.157	0.010	- 142.461	-20.713
			OA	3.399	29.891	32.157	0.910	-57.475	64.273
F2	a	CH	OA	148.876*	66.967	44.801	0.031	13.982	283.771
			YA	126.166	66.967	44.801	0.066	-8.728	261.061
		OA	CH	-148.876*	66.967	44.801	0.031	- 283.771	-13.982
			YA	-22.71	66.967	44.801	0.736	- 157.605	112.185
		YA	CH	-126.166	66.967	44.801	0.066	- 261.061	8.728
			OA	22.71	66.967	44.801	0.736	- 112.185	157.605
	ae	CH	OA	9.134	66.967	44.801	0.892	- 125.761	144.028
			YA	89.836	66.967	44.801	0.187	-45.058	224.731
		OA	CH	-9.134	66.967	44.801	0.892	- 144.028	125.761
			YA	80.703	66.967	44.801	0.234	-54.192	215.597
		YA	CH	-89.836	66.967	44.801	0.187	- 224.731	45.058
			OA	-80.703	66.967	44.801	0.234	-	54.192

								215.597	
	i	CH	OA	198.926^{*}	66.967	44.801	0.005	64.032	333.821
			YA	82.324	66.967	44.801	0.225	-52.571	217.218
		OA	CH	-198.926^{*}	66.967	44.801	0.005	-333.821	-64.032
			YA	-116.602	66.967	44.801	0.089	-251.497	18.292
		YA	CH	-82.324	66.967	44.801	0.225	-217.218	52.571
			OA	116.602	66.967	44.801	0.089	-18.292	251.497
	u	CH	OA	288.562^{*}	66.967	44.801	0.000	153.668	423.457
			YA	151.366^{*}	66.967	44.801	0.029	16.471	286.26
		OA	CH	-288.562^{*}	66.967	44.801	0.000	-423.457	-153.668
			YA	-137.197^{*}	66.967	44.801	0.046	-272.092	-2.302
		YA	CH	-151.366^{*}	66.967	44.801	0.029	-286.26	-16.471
			OA	137.197^{*}	66.967	44.801	0.046	2.302	272.092

Table 48: Experiment 1 estimates for the interactions between Vowel and Environment-Oriented Speaking Style

	QN	Vowel	Mean	Std. Error	df	95% Confidence Interval	
						Lower Bound	Upper Bound
F1	N	a	876.816	24.135	7.235	820.119	933.512
		ae	882.875	24.135	7.235	826.179	939.571
		i	372.501	24.135	7.235	315.804	429.197
		u	413.735	24.135	7.235	357.039	470.432
	Q	a	818.714	24.135	7.235	762.018	875.411
		ae	816.134	24.135	7.235	759.438	872.831
		i	343.285	24.135	7.235	286.589	399.981
		u	384.708	24.135	7.235	328.012	441.405

Table 49: Experiment 1 pairwise comparisons for the interactions between Vowel and Environment-Oriented Speaking Style

	Vowel	(I) QN	(J) QN	Mean Difference (I-J)	Std. Error	df	Sig.	95% Confidence Interval for Difference	
								Lower Bound	Upper Bound

F1	a	N	Q	58.102*	9.012	312.16	0.000	40.369	75.834
		Q	N	-58.102*	9.012	312.16	0.000	-75.834	-40.369
	ae	N	Q	66.741*	9.012	312.16	0.000	49.008	84.473
		Q	N	-66.741*	9.012	312.16	0.000	-84.473	-49.008
	i	N	Q	29.216*	9.012	312.16	0.001	11.483	46.948
		Q	N	-29.216*	9.012	312.16	0.001	-46.948	-11.483
	u	N	Q	29.027*	9.012	312.16	0.001	11.294	46.76
		Q	N	-29.027*	9.012	312.16	0.001	-46.76	-11.294

Table 50: Experiment 1 estimates for the interactions between Vowel and Environment-Oriented Speaking Style

	COCL	Vowel	Mean	Std. Error	df	95% Confidence Interval	
						Lower Bound	Upper Bound
F1	CL	a	859.507	24.135	7.235	802.81	916.203
		ae	867.619	24.135	7.235	810.923	924.316
		i	356.025	24.135	7.235	299.329	412.721
		u	399.875	24.135	7.235	343.178	456.571
	CO	a	836.023	24.135	7.235	779.327	892.72
		ae	831.39	24.135	7.235	774.694	888.086
		i	359.761	24.135	7.235	303.064	416.457
		u	398.569	24.135	7.235	341.873	455.266

Table 51: Experiment 1 pairwise comparisons for the interactions between Vowel and Environment-Oriented Speaking Style

	Vowel	(I) COCL	(J) COCL	Mean Difference (I-J)	Std. Error	df	Sig.	95% Confidence Interval for Difference	
								Lower Bound	Upper Bound
F1	a	CL	CO	23.483*	9.012	312.16	0.010	5.751	41.216
		CO	CL	-23.483*	9.012	312.16	0.010	-41.216	-5.751
	ae	CL	CO	36.229*	9.012	312.16	0.000	18.497	53.962
		CO	CL	-36.229*	9.012	312.16	0.000	-53.962	-18.497
	i	CL	CO	-3.736	9.012	312.16	0.679	-21.468	13.997
		CO	CL	3.736	9.012	312.16	0.679	-13.997	21.468
	u	CL	CO	1.305	9.012	312.16	0.885	-16.427	19.038
		CO	CL	-1.305	9.012	312.16	0.885	-19.038	16.427

Table 52: Production-perception summary of random effects

Regression No.	Feature	Parameter	Variance	Std. Deviation
1	F0 range	Sentence	0.000	0.000
		Listener	0.000	0.000
		Talker	0.000	0.000
2	F0 mean	Sentence	0.587	0.766
		Listener	0.000	0.000
		Talker	1.051	1.025
3	1-3 kHz energy	Sentence	0.000	0.801
		Listener	0.000	0.000
		Talker	0.000	0.718
4	Speaking rate	Sentence	0.000	0.000
		Listener	0.000	0.000
		Talker	0.000	0.000
5	Pause duration	Sentence	0.631	0.794
		Listener	0.000	0.000
		Talker	0.483	0.695
6	Vowel duration	Sentence	0.432	0.657
		Listener	0.290	0.539
		Talker	0.187	0.433
7	F1	Sentence	0.372	0.610
		Listener	0.282	0.531
		Talker	0.118	0.344
8	F2	Sentence	0.044	0.209
		Listener	0.444	0.667
		Talker	0.219	0.468
9	HNR	Listener	0.000	0.000
		Talker	0.000	0.000
10	Jitter	Listener	0.011	0.103
		Talker	0.000	0.000
11	Shimmer	Listener	0.009	0.096
		Talker	0.000	0.000

References

Andersen, E. S. (1996). "A cross-cultural study of children's register knowledge," In D. I. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (Eds.), Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp. 125–142. Mahwah, NJ: Erlbaum.

Andersen, E. S., Brizuela, M., DePuy, B., & Gonnerman, L. (1999). "Cross-linguistic evidence of the early acquisition of discourse markers as register variables," Journal of Pragmatics, 31, 1339–1351.

Amazi, D. K., & Garber, S. R. (1982). "The Lombard sign as a function of age and task," Journal of Speech, Language and Hearing Research, 25(4), 581.

Assmann, P., Barreda, S., & Nearey, T. (2013). "Perception of speaker age in children's voices," In Proceedings of Meetings on Acoustics, 19, 060059.

Au, R., Joung, P., Nicholas, M., & Obler, L. (1995). "Naming ability across the adult life span," Aging and Cognition, 2, 300-311.

Awan, S. N. (2006). "The aging female voice: acoustic and respiratory data," Clinical linguistics & phonetics, 20(2-3), 171-180.

Benjamin, B. J. (1982). "Phonological performance in gerontological speech," Journal of Psycholinguistic Research, 11(2), 159-167.

Bent, T., Buchwald, A., & Pisoni, D. B. (2009). "Perceptual adaptation and intelligibility of multiple talkers for two types of degraded speech," The Journal of the Acoustical Society of America, 126, 2660–2669.

Boersma, P., & Weenink, D. (1995). Praat speech processing software. Institute of Phonetics Sciences of the University of Amsterdam. http://www. praat. org.

Bond, Z., & Moore, T. (1994). "A note on the acoustic- phonetic characteristics of inadvertently clear speech," Speech Communication, 14, 325–337.

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

Bradlow, A., Torretta, G., & Pisoni, D. (1996). “Intelligibility of normal speech: 1. Global and fine-grained acoustic-phonetic talker characteristics,” Speech Communication, 20, 255–272.

Brinton, B., Fujiki, M., Loeb, D. F., & Winkler, E. (1986). "Development of conversational repair strategies in response to requests for clarification," Journal of Speech and Hearing Research, 29, 75–81.

Brumm H, Voss K, Köllmer I, Todt D (2004). “Acoustic communication in noise: Regulation of call characteristics in a New World monkey,” J Exp Biol 207(3):443–448.

Brumm, H., & Todt, D. (2002). Noise-dependent song amplitude regulation in a territorial songbird. *Animal Behaviour*, *63*(5), 891-897.

Cooke, M., King, S., Garnier, M., & Aubanel, V. (2014). “The listening talker: A review of human and algorithmic context-induced modifications of speech,” Computer Speech & Language, 28(2), 543-571.

Cooke, M., & Lu, Y. (2010). “Spectral and temporal changes to speech produced in the presence of energetic and informational maskers,” The Journal of the Acoustical Society of America, 128 (4), 2059-2069.

Cooke, M., Mayo, C., & Villegas, J. (2014). “The contribution of durational and spectral changes to the Lombard speech intelligibility benefit,” The Journal of the Acoustical Society of America, 135(2), 874-883.

Dreher, J. J., & O’Neill, J. (1957). “Effects of ambient noise on speaker intelligibility for words and phrases,” The Journal of the Acoustical Society of America, 29, 1320–1323.

Endres, W., Bambach, W., & Flösser, G. (1971). "Voice spectrograms as a function of age, voice disguise, and voice imitation," The Journal of the Acoustical Society of America, 49(6B), 1842-1848.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2002), "Children's use of semantic cues in degraded listening environments," The Journal of the Acoustical Society of America, 111, 2242.

Ferguson, S. H. (2004). "Talker differences in clear and conversational speech: Vowel intelligibility for normal- hearing listeners," The Journal of the Acoustical Society of America, 116, 2365–2373.

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

Ferguson, S. H., & Kewley-Port, D. (2007). "Talker differences in clear and conversational speech: Acoustic characteristics of vowels," Journal of Speech, Language and Hearing Research, 50(5), 1241.

Ferguson, S. H., Poore, M. A., Shrivastav, R., Kendrick, A., McGinnis, M., & Perigoe, C. (2010). "Acoustic Correlates of Reported Clear Speech Strategies," Journal of the Academy of Rehabilitative Audiology, 43, 45-64.

Ferguson, S. H., & Quené, H. (2014). "Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners," The Journal of the Acoustical Society of America, 135(6), 3570-3584.

Ferrand, C. T. (2000). "Harmonics-to-noise ratios in normally speaking prepubescent girls and boys," Journal of Voice, 14(1), 17-21.

Ferrand, C. T. (2002). "Harmonics-to-noise ratio: an index of vocal aging," Journal of Voice, 16(4), 480-487.

Gagne, J. P., Rochette, A. J., & Charest, M. (2002). “Auditory, visual and audiovisual clear speech,” Speech Communication, 37(3), 213-230.

Garber, S., Speidel, T., & Siegel, G. (1980). "The effects of noise and palatal appliances on the speech of five-year-old children," Journal of Speech and Hearing Research, 23, 853-863.

Garber, S., Speidel, T., & Siegel, G., Miller, E., & Glass, L. (1980). “The effects of presentation of noise and dental appliances on speech,” Journal of Speech and Hearing Research, 23, 836-850.

Gerosa, M., Lee, S., Giuliani, D., & Narayanan, S. (2006). “Analyzing children's speech: an acoustic study of consonants and consonant-vowel transition,” In Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on (Vol. 1, pp. I-I). IEEE.

Gilbert, R., Chandrasekaran, B., & Smiljanic, R. (2014). “Recognition memory in noise for speech of varying intelligibility.” The Journal of the Acoustical Society of America, 135, 389.

Glaze, L. E., Bless, D. M., & Susser, R. D. (1990). “Acoustic analysis of vowel and loudness differences in children's voice,” Journal of Voice, 4(1), 37-44.

Godoy, E., Koutsogiannaki, M., & Stylianou, Y. (2014). “Approaching speech intelligibility enhancement with inspiration from Lombard and Clear speaking styles,” Computer Speech & Language, 28(2), 629-647.

Gordon-Salant, S., & Fitzgibbons, P. J. (1997). “Selected cognitive factors and speech recognition performance among young and elderly listeners,” Journal of Speech, Language and Hearing Research, 40(2), 423.

Gorham-Rowan, M. M., & Laures-Gore, J. (2006). “Acoustic-perceptual correlates of voice quality in elderly men and women,” Journal of communication disorders, 39(3), 171-184.

Goy, H., Pichora-Fuller, K., van Lieshout, P., Singh, G., & Schneider, B. (2007). "Effect of within-and between-talker variability on word identification in noise by younger and older adults," Canadian Acoustics, 35, 108-109.

Green, T., Katiri, S., Faulkner, A., & Rosen, S. (2007). "Talker intelligibility differences in cochlear implant listeners," The Journal of the Acoustical Society of America, 121, EL223–EL229.

Hage, S. R., Jiang, T., Berquist, S. W., Feng, J., & Metzner, W. (2013). "Ambient noise induces independent shifts in call frequency and amplitude within the Lombard effect in echolocating bats," Proceedings of the National Academy of Sciences, 110(10), 4063-4068.

Hanley, T. D. & Steer, M. D. (1949), "Effect of level of distracting noise upon speaking rate, duration and intensity," The Journal of Speech and Hearing Disorders, 14, 363-368.

Halle, S., & Myerson, J. (1996). "Experimental evidence for differential slowing in the lexical and non-lexical domains," Aging, neuropsychology, and Cognition, 3, 154-165.

Harnsberger, J. D., Shrivastav, R., Brown Jr, W. S., Rothman, H., & Hollien, H. (2008). "Speaking rate and fundamental frequency as speech cues to perceived age," Journal of voice, 22(1), 58-69.

Hartman, D. E., & Danhauer, J. L. (1976). "Perceptual features of speech for males in four perceived age decades," The Journal of the Acoustical Society of America, 59, 713.

Hazan, V., & Baker, R. (2011). "Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions," The Journal of the Acoustical Society of America, 130 (4), 2139-2152.

Hazan, V., & Markham, D. (2004). "Acoustic-phonetic correlates of talker intelligibility for adults and children," The Journal of the Acoustical Society of America, 116, 3108.

Higgins, M. B., & Saxman, J. H. (1991). "A comparison of selected phonatory behaviors of healthy aged and young adults," Journal of Speech, Language and Hearing Research, 34(5), 1000.

Huber, J. E. (2008). "Effects of utterance length and vocal loudness on speech breathing in older adults," Respiratory physiology & neurobiology, 164(3), 323-330.

Junqua, J. C. (1993). "The Lombard reflex and its role on human listeners and automatic speech recognizers," The Journal of the Acoustical Society of America, 93(1), 510-524.

Kang K. H. & Guion S. G. (2008). "Clear speech production of Korean stops: changing phonetic targets and enhancement strategies," The Journal of the Acoustical Society of America, 124, 3909–3917.

Krause, J. C., & Braida, L. D. (1995). "The effects of speaking rate on the intelligibility of speech for various speaking modes," The Journal of the Acoustical Society of America, 98(5), 2982-2982.

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

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

Lane, H., & Tranel, B. (1971). "The Lombard sign and the role of hearing in speech," Journal of Speech, Language, and Hearing Research, 14(4), 677-709.

Lee, S., Potamianos, A., and Narayanan, S. (1999). "Acoustics of children's speech: developmental changes of temporal and spectral parameters," The Journal of the Acoustical Society of America, 105, 1455–1468.

Linders, B., Massa, G. G., Boersma, B., & Dejonckere, P. H. (1995). "Fundamental voice frequency and jitter in girls and boys measured with electroglottography: influence of age and height. International journal of pediatric otorhinolaryngology," 33(1), 61-65.

Linville SE. (1987). "Maximum phonational frequency range capabilities of women's voices with advancing age," Folia Phoniatr. 39, 297–301.

Linville, S. E., & Fisher, H. B. (1985). "Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females," The Journal of the Acoustical Society of America, 78(1), 40-48.

Linville, S. E., & Rens, J. (2001). "Vocal tract resonance analysis of aging voice using long-term average spectra," Journal of Voice, 15(3), 323-330.

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

Lombard, E. (1911). "Le signe de l'elevation de la voix [The sign of the rise in the voice]," Ann. maladies oreille, larynx, nez, pharynx 37, 101-119.

Lu, Y., & Cooke, M. (2008). "Speech production modifications produced by competing talkers, babble, and stationary noise," The Journal of the Acoustical Society of America, 124 (5), 3261-3275.

Lu, Y., & Cooke, M. (2009). "The contribution of changes in F0 and spectral tilt to increased intelligibility of speech produced in noise," Speech Communication, 51(12), 1253-1262.

Markham, D. & Hazan, V. 2004. "The effect of talker- and listener-related factors on intelligibility for a real-word, open-set perception test," The Journal of Speech Language and Hearing Research, 47, 725-737.

Navarro R. (1996). "Effects of ear canal occlusion and masking on the perception of voice," Perceptual and Motor Skills, 82, 199–208.

Nicoladis, E., & Genesee, F. (1997). "Language development in preschool bilingual children," Journal of Speech-Language Pathology and Audiology, 21 (4), 258–270.

Pedlow, R., Sanson, A., & Wales, R. (2004). "Children's production and comprehension of politeness in requests: Relationships to behavioural adjustment, temperament and empathy," First Language, 24(3), 347-367.

Pedlow, R., Wales, R., & Sanson, A. (2001). "Children's Production and Comprehension of Politeness in Requests Relationships to Behavioral Adjustment in Middle Childhood," Journal of Language and Social Psychology, 20(1-2), 23-60.

Pettinato, M., & Hazan, V. (2013). "The development of clear speech strategies in 9-14 year olds," In Proceedings of Meetings on Acoustics, 19, 060121.

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," The Journal of Speech Language 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 difference in intelligibility between clear and conversational speech," The Journal of Speech Language and Hearing Research, 32, 600–603.

Pichora-Fuller, M. K., Goy, H., & van Lieshout, P. (2010). "Effect on speech intelligibility of changes in speech production influenced by instructions and communication environments," Seminars in Hearing 31, 77-94.

Pichora-Fuller, M., Schneider, B., & Daneman, M. (1995). "How young and old adults listen to and remember speech in noise," The Journal of the Acoustical Society of America, 97(1), 593-608.

Pittman, A. L., & Wiley, T. L. (2001). “Recognition of speech produced in noise,” The Journal of Speech Language and Hearing Research, 44, 487–496.

Ramig, L. A., & Ringel, R. L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. The Journal of Speech Language and Hearing Research, 26(1), 22-30.

Redford, M. A., & Gildersleeve-Neumann, C. E. (2009). “The development of distinct speaking styles in preschool children,” The Journal of Speech Language and Hearing Research, 52(6), 1434.

Ringel RL, Chodzko-Zajko W. J. (1987). “Vocal indices of biological age,” Journal of Voice, 1, 31–37.

Rogers, C. L., DeMasi, T. M., & Krause, J. C. (2010). “Conversational and clear speech intelligibility of /bVd/ syllables produced by native and non-native English speakers,” The Journal of the Acoustical Society of America, 128(1), 410.

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). “Listening in aging adults: From discourse comprehension to psychoacoustics,” Canadian Journal of Experimental Psychology, 56, 139–152.

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.

Scukanec, G. P., Petrosino, L., & Squibb, K. (1991). “Formant frequency characteristics of children, young adult, and aged female speakers,” Perceptual and motor skills, 73(1), 203-208.

Shatz, M., & Gelman, R. (1973). “The development of communication skills: Modifications in the speech of young children ad a function of listener,” Monographs of the Society for Research in Child Development, 38(5), 1–38.

Siegel, G., Pick, H. L., Olsen, M., & Sawin, L. (1976). “Auditory feedback in the regulation of vocal intensity of preschool children,” Developmental Psychology, 12, 255-261.

Smiljanic, R. (2013). “Can older adults enhance the intelligibility of their speech?” The Journal of the Acoustical Society of America, 133(2), EL129-EL135.

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

Smiljanic, R., & Bradlow, A. (2007). “Clear speech intelligibility: Listener and talker effects,” In Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrucken, Germany, Speech Perception II, Paper 1; http://www.icphs2007.de/

Smiljanic, R., & Bradlow, A. R. (2009). “Speaking and hearing clearly: Talker and listener factors in speaking style changes,” Language and Linguistics Compass, 3, 236–264.

Spieler, D., Horton, W., & Shriberg, E. (2004). “Language use in the wild: Analyses of conversational speech across lifespan,” In 10th Cognitive Aging conference, Atlanta, GA.

Stathopoulos, E. T., Huber, J. E., & Sussman, J. E. (2011). “Changes in acoustic characteristics of the voice across the life span: measures from individuals 4–93 years of age,” Journal of Speech, Language, and Hearing Research, 54(4), 1011-1021.

Stathopoulos, E. T., & Sapienza, C. (1993). “Respiratory and laryngeal measures of children during vocal intensity variation,” The Journal of the Acoustical Society of America, 94, 2531.

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). “Effects of noise on speech production: acoustic and perceptual analyses,” The Journal of the Acoustical Society of America, 84 (3), 917–928.

Syrett, K., & Kawahara, S. (2013). “Production and perception of listener-oriented clear speech in child language,” Journal of child language, 1-17.

Tang, J., & Stathopoulos, E. T. (1995). “Vocal efficiency as a function of vocal intensity: A study of children, women, and men,” The Journal of the Acoustical Society of America, 97, 1885.

Tartter, V. C., Gomes, H., & Litwin, E. (1993). “Some acoustic effects of listening to noise on speech production,” The Journal of the Acoustical Society of America, 94, 2437–2440.

Van Dommelen, W. A., & Hazan, V. (2012). “Impact of talker variability on word recognition in non-native listeners,” The Journal of the Acoustical Society of America, 132(3), 1690-1699.

Van Engen, K. J., & Bradlow, A. R. (2007). “Sentence recognition in native-and foreign-language multi-talker background noise,” The Journal of the Acoustical Society of America, 121(1), 519.

Van Engen, K., Chandrasekaran, B., & Smiljanic, R. (2012). “Effects of speech clarity on recognition memory for spoken sentences,” PLoS ONE 7 (9), e43753. doi:10.1371/journal.pone.0043753.

Weeks, T. E. (1971). “Speech registers in young children,” Child Development, 42, 1119–1131.

Wilcox K. A., & Horii Y. (1980). “Age and changes in vocal jitter,” Journal of Gerontology, 35, 194–198.

Xue, S. A., & Hao, G. J. (2003). “Changes in the Human Vocal Tract Due to Aging and the Acoustic Correlates of Speech Production: A Pilot Study,” Journal of Speech, Language, and Hearing Research, 46(3), 689-701.

Yumoto, E., Gould, W. J., & Baer, T. (1982). "Harmonics-to-noise ratio as an index of the degree of hoarseness," The Journal of the Acoustical Society of America, 71(6), 1544-1550.