Children's Perception of Conversational and Clear American-English Vowels in Noise

Dorothy Leone

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

ABSTRACT

Children’s Perception of Conversational and Clear American-English Vowels in Noise

Dorothy Leone

A handful of studies have examined children’s perception of clear speech in the presence of background noise. Although accurate vowel perception is important for listeners’ comprehension, no study has focused on whether vowels uttered in clear speech aid intelligibility for children listeners. In the present study, American-English (AE) speaking children repeated the AE vowels /ɛ, æ, ɑ, ʌ/ in the nonsense word /gəbVpə/ in phrases produced in conversational and clear speech by two female AE-speaking adults. The recordings of the adults’ speech were presented at a signal-to-noise ratio (SNR) of -6 dB to 15 AE-speaking children (ages 5.0-8.5) in an examination of whether the accuracy of AE school-age children’s vowel identification in noise is more accurate when utterances are produced in clear speech than in conversational speech. Effects of the particular vowel uttered and talker effects were also examined. Clear speech vowels were repeated significantly more accurately (87%) than conversational speech vowels (59%), suggesting that clear speech aids children’s vowel identification. Results varied as a function of the talker and particular vowel uttered. Child listeners repeated one talker’s vowels more accurately than the other’s and front vowels more accurately than central and back vowels. The findings support the use of clear speech for enhancing adult-to-child communication in AE, particularly in noisy environments.

Table of contents

List of tables

List of figures

List of appendices

Acknowledgements

I am indebted to my mentor, Dr. Erika S. Levy, for truly defining and modeling for me what it means to be a researcher, a professor, a mentor, and a speech scientist. Without Dr. Levy's incredible writing and research support this dissertation would not have been possible. I am forever thankful for all she has taught me and continues to teach me.

I am grateful to the members of my doctoral committee, Dr. Karen Froud, Dr. Honor O'Malley, Dr. Robert Remez, and Dr. Hansun Waring, for their contributions to my research.

I would also like to acknowledge all of my colleagues and co-workers who supported my dissertation research. I am grateful for all of the support from the women of the Speech Production and Perception Lab at Teachers College, especially Miriam Baigorri, Gemma Moya Gale, and Sih-Chiao Hsu - their support with running participants, recording talkers, and problem-solving in the lab will always be remembered. Also, I wish to thank Luca Campanelli for his invaluable statistical knowledge and Shari Berkowitz for all of her acoustic analysis support. Thank you to my colleagues at Iona College, particularly the woman of the Speech Communication Studies department, for showing me how to "fight the good fight."

A very special thank you to my family and friends for all of their support throughout my entire academic career, especially my mother, father, sister, and husband. I thank my father, Kenny Tancredi, for his continuous academic encouragement beginning from my elementary school years. I thank my sister, Amanda Tancredi, for the support of my research in so many ways – as my talker, my participant, my lab assistant, and, most importantly, my friend. An immense thank you to my husband, John Leone, for his constant encouragement and never-ending support. John's listening ear - both for practice presentations and research hiccups - was incredible. John's patience, kindness, and love during my dissertation research are unforgettable.

Last, but certainly not least, I am forever indebted to my mother, a woman who taught me the value of an education and encouraged me each and every day of my schooling. This dissertation is dedicated to one of the strongest and most loving people in my life, my mother, Dorothy Tancredi.

Chapter 1. Introduction

Background noise exists in almost every listening environment (Helfer & Wilber, 1990), and can impact a listener's ability to perceive a speech signal (Crandell & Smaldino, 2000; Helfer & Wilber, 1990). As an SNR becomes less favorable, a listener's ability to accurately perceive the signal decreases (Stelmachowicz, Hoover, Lewis, Kortekaas, & Pittman, 2000). The ability to perceive a speech signal in adverse listening conditions, such as in the presence of noise or reverberation, increases with age until early adulthood (Crandell & Smaldino, 2000; Neuman & Hochberg, 1986; Nishi, Lewis, Hoover, Choi, & Stelmachowicz, 2010; Soli & Sullivan, 1997). Nelson and Soli (2000) suggest that acoustic environments utilizing +15 dB SNR (i.e.., the speech signal presented at 15 dB above the noise level) allow children to perceive a signal fully. In the United States classroom SNRs while class is in session have been reported between +3 dB to -17.6 dB (Larsen & Blair, 2008), suggesting that many children spend a large portion of their day listening to speech in the presence of considerable background noise.

High noise levels in classrooms may cause children to miss critical acoustic cues in the speech signal (Crandell & Smaldino, 2000), increase students' anxiety, and decrease their learning (Tennessee Advisory Commission on Intergovernmental Relations, 2003). Studies of the impact of noise on children's perception difficulties and explorations of strategies for increasing intelligibility of speakers in noise for these listeners may yield findings beneficial to children's learning in the United States, where educational attainment is declining in comparison to other countries (Dillon, 2010).

Accurate perception of words relies, in part, on accurate vowel perception. Kewley-Port, Burkle, and Lee (2007) reported that vowels carry more information about sentence intelligibility than do consonants. They compared sentence intelligibility in a consonant-only condition (in which adults listened to sentences with vowels replaced by speech-shaped noise), and a vowel-only condition (in which adults listened to sentences with consonants replaced by speech-shaped noise). The adults repeated significantly more words accurately after listening to vowel-only sentences than after listening to consonant-only sentences, suggesting that vowels may be more important for intelligibility than consonants.

Research has found that talkers modify their speech style in noisy environments in ways that are beneficial to the listener. Talkers reveal a Lombard effect, for example, increasing their volume in noisy environments (Garnier, Henrich, & Dubois, 2010). Clear speech, the focus of this study, is another intelligibility-enhancing style that talkers may utilize in background noise (Picheny, Durlach, & Braida, 1986; Smiljanic & Bradlow, 2009; Uchanski, 2005). Talkers also utilize clear speech when speaking to listeners with hearing loss (Uchanski, 2005). Clear speech is often contrasted with "conversational speech" (also known as "typical speech," or "plain speech"), the speech style used when conversing with someone highly familiar, such as a friend or family member (Bradlow, Kraus, & Hayes, 2003). Clear speech is typically characterized by higher pitch, longer duration, and increased amplitude than conversational speech (Picheny et al., 1986).

This study examined the accuracy with which AE typically developing school-age children perceive adults' AE vowels in conversational and clear speech in noise. Additionally, the acoustic properties of clear speech vowels were examined. The following section is a review of the pertinent literature, beginning with what is known about adults' and children's speech perception in noise, followed by a description of clear speech, and lastly, information on adults' and children's perception of clear speech.

1.1. Adults' perception of speech in noise

The effects of noise on speech perception have been studied with adult populations more extensively than with children. A brief review of studies of adults is followed by a review of the handful of studies that have examined children's speech perception in noise.

In general, when adults are presented with speech in noise, their perceptual accuracy decreases as the signal becomes noisier. The adverse effect of noise has been demonstrated for normal-hearing adults and adults with hearing loss (Nabeleck & Mason, 1981; Payton, Uchanski, & Braida, 1994) listening to words (Rogers, Lister, Febo, Besing, & Abrams, 2006), sentences (Payton et al., 1994; Percy & Kei, 2008), and speech sounds in isolation (Cutler, Weber, Smits, & Cooper, 2004; Gelfand, Piper, & Silman, 1986). Rogers et al. (2006) presented monosyllabic words at 50 dB HL in SNRs of 0, -2, and -6 dB to adults with normal hearing. In a word-repetition task, the adults repeated significantly fewer words accurately in the -6 dB listening condition than in the 0 dB condition. Similarly, Payton et al. (1994) found that listeners with normal hearing and listeners with hearing loss identified significantly fewer key words in nonsense sentences in the presence of speech-shaped noise, reverberation, and white noise together, than in white noise alone. Percy and Kei (2008) reported that adults answered significantly more multiple choice questions accurately when listening to sentences in more beneficial SNRs than in less beneficial SNRs. Similarly, adults identified consonants significantly more accurately in quiet than in +5 dB and +10 dB SNR conditions (Gelfand et al., 1986). Cutler et al. (2004) reported similar results for adults identifying vowels in noise in 0 dB SNR produced by one female talker.

1.2. Children's perception of speech in noise

A handful of studies have been conducted on the impact of noise on children's speech perception. Children require higher SNRs than adults to identify isolated words in background noise than in quiet. Stuart, Givens, Walker, and Elangovan (2006) examined normal-hearing preschool children's and adults' word recognition in noise. They presented monosyllabic words at 50 dB HL using SNRs of +10, 0, and -10 dB in both continuous and interrupted noise. Children aged 4-5 years, had significantly lower word-recognition abilities in all three SNRs than adults. Bradley and Sato (2008) reported on phonetically-balanced noun recognition in noise by children aged 6, 8, and 11 in their classroom setting. The younger children needed significantly higher SNRs to achieve the same word identification accuracy as the older children. Specifically, to achieve 95% accuracy, 11-year-old children needed +8.5 dB SNR, 8-year-old children +12.5 dB, and 6-year-old children +15.5 dB.

Similar results have been observed at the segmental level. Johnson (2000) examined children's (ages 6-15) and young adults' consonant and vowel identification by presenting CVCV sequences in multi-talker babble noise at +13 dB SNR. Results suggested that children's ability to perceive vowels and consonants in noise increases as they age. The children identified consonants in noise with adult-like accuracy by age 14 and vowels in noise with adult-like accuracy by age 10. Nishi et al. (2010) also investigated children's versus adults' phoneme identification in noise. They presented VCV stimuli to children between the ages of 4 and 9 and adults between the ages of 19 and 41 at 0, +5, and +10 dB in speech-shaped noise. More accurate performance was found as a function of increased age and a more favorable SNR. Similarly, Ziegler, Pech-Georgel, George, and Lorenzi (2009) presented VCVs recorded by a French-speaking female in quiet and in noise to native French-speaking children with dyslexia. Children repeated what they heard and an experimenter selected the VCV from 16 alternatives on a computer screen that corresponded to the children's utterance. Results indicated that the children's consonant identification was significantly less accurate in noise than in quiet.

In summary, children have difficulty identifying speech in noise. A variety of factors contribute to this difficulty, including the child's age and the noise level. Older children tend to perceive speech in noise more accurately than younger children and a decrease in noise is beneficial to most children.

1.3. Clear speech

1.3.1. Factors affecting clear speech production and perception

When talkers are aware of a speech perception difficulty, they often modify the speech signal, usually improving the accuracy with which their speech is perceived (Smiljanic & Bradlow, 2009). Different forms of intelligibility-enhancing speech include motherese, a speaking style directed at infants, and Lombard speech, a speaking style used in the presence of background noise. Motherese, Lombard speech, and clear speech, are characterized by a higher intensity, higher fundamental frequency, and longer duration than conversational speech (Kirchhoff & Schimmel, 2005; Kuhl & Andruski, 1997; Wassink, Wright, & Frankin, 2007). Clear speech differs from motherese and Lombard speech in that it is typically directed towards non-native listeners or listeners with hearing loss, as opposed to being child-directed or spoken in the presence of background noise. Additionally, clear speech is characterized by lower intensity than Lombard speech and by lower prosody than motherese (Smiljanic & Bradlow, 2009).

Instructions on producing clear speech may impact its production. Some studies utilize a simple statement of directions such as "Speak as if you are talking to someone with a hearing impairment or a non-native listener" (Bradlow & Alexander, 2007; Bradlow et al., 2003; Uchanski, 2005). Other studies provide the talkers with more elaborate descriptions of how to produce clear speech. Krause and Braida (2002), for example, provided talkers with clear and conversational audio speech samples, discussions about the clear speech samples, and practice and feedback sessions for clear speech production.

While descriptions of the acoustic and phonetic characteristics of clear speech vary across studies, clear speech typically is characterized by higher intensity (Bradlow et al., 2003; Picheny et al., 1986; Uchanski, 2005), slower speaking rate, more pauses, and higher fundamental frequency (Bradlow et al., 2003; Krause & Braida, 2002; Picheny et al., 1986; Uchanski, 2005) than conversational speech. Factors that may affect the variability in clear speech production include talker characteristics, instructions to the talker, and the specific stimuli used. Bradlow et al. (2003) found differences between a male's and a female's clear speech productions. The female's utterances were produced with a slower speaking rate, larger consonant-to-vowel intensity ratio, a relatively higher fundamental frequency, and a larger vowel space range for both the first and second formants than the male's. Similarly, Krause and Braida (2009) found that five talkers produced clear speech at varying speaking rates, with rates ranging from 57 to 169 words per minute. In addition, Ferguson (2004) described clear speech (at the word level) produced by 41 talkers who varied in age and gender. Talkers produced clear speech in diverse ways, with some talkers increasing vowel duration more than others, for example. Because of this variability in clear speech production, Uchanski (2005) advised using more than one talker in clear speech studies.

A talker's gender impacts listeners' clear speech perception. Ferguson (2004) reported that gender was the only factor that was associated with a difference in clear speech intelligibility: adult listeners identified clear speech produced by females significantly more accurately than clear speech produced by males. There was no statistically significant difference between the identification accuracy of females' conversational speech and of males' conversational speech. Similarly, Bradlow et al. (2003) documented that child listeners benefited more from clear speech produced by a female talker than by a male talker.

Clear speech can be presented to listeners in a variety of contexts, which may impact the cues utilized by listeners to enhance intelligibility. Clear speech studies may involve stimuli presented to listeners as syllables in isolation (Gagne, Rouchette, & Charest, 2002), words (Ferguson, 2004; Ferguson & Kewley-Port, 2002; Gagne, Masterson, Munhall, Bilida, & Querengesser, 1994), nonsense sentences (Payton et al., 1994; Picheny et al., 1985) or meaningful sentences (Bradlow & Alexander, 2007; Bradlow & Bent, 2002; Bradlow et al., 2003; Ferguson & Kewley-Port, 2002;). Tasks in which meaningful sentences are presented to listeners represent more realistic listening conditions than tasks that present words in isolation (Strange & Schaffer, 2008). However, tasks in which isolated words or words in carrier phrases are presented without semantic cues allow researchers to minimize top-down influences and zero in on the effects of clear speech on the perception of specific speech sounds (Smiljanic & Bradlow, 2009).

1.3.2. Clear speech vowels

Vowels produced in clear speech have been analyzed in more depth than consonants in clear speech. Studies consistently characterize clear speech vowels as having longer duration than conversational speech vowels (Ferguson & Kewley-Port, 2002, 2007; Picheny et al., 1986). Additionally, studies report a decrease in vowel reduction during clear speech (Picheny et al., 1986; Uchanski, 2005). Greater vowel space range has been found in clear than in conversational speech (Bradlow et al., 2003). Specifically, the first formant (F1), which corresponds to tongue height, increases for most clear speech vowels, suggesting that the tongue tends to be lower in the oral cavity for the production of clear speech vowels than for the production of conversational speech vowels. The F1 increase is expected because F1 typically increases when vocal effort is increased (Liénard & Di Benedetto, 1999). The second formant (F2), which corresponds to anterior-posterior tongue movement, increases for clear speech front vowels, suggesting a more forward tongue position, whereas it decreases for clear speech back vowels, suggesting a more retracted tongue position (Ferguson & Kewley-Port, 2007). Lastly, individual clear speech vowels are more dynamic than their conversational speech counterparts, as measured by greater spectral change (Ferguson & Kewley-Port, 2002).

The clear speech literature has documented variability in adults' and children's identification accuracy of particular clear speech vowels (Ferguson & Kewley-Port, 2002; Leone, Hsu, Baigorri, Moya-Gale, & Levy, 2011; Rogers, DeMasi, & Braida, 2010). Ferguson and Kewley-Port (2002) reported that when normal-hearing AE-speaking adults identified AE clear speech vowels in /hVd/ context in noise, /u/ and /ɛ/ were identified with less accuracy (74.7% and 76.4%, respectively) than / i, ɪ, e, æ, ɑ, ʌ, o, ʊ/. In contrast, Rogers et al. (2010) found that when AE adults identified AE vowels in /bVd/ context in noise, clear speech /ɪ/ was identified with less accuracy than /ɛ, ɑ, i, e, æ/. Similar to Ferguson and Kewley-Port's study about adult listeners, a preliminary study in which one child listener repeated vowels in /gəbVpə/ trisyllables in sentences embedded in noise found lower accuracy for the repetition of clear speech /ɛ/ than clear speech /i, æ, ɑ, ʌ/ (Leone et al., 2011).

Although, as described above, identification accuracy of clear speech vowels is discussed in the literature, to the author's knowledge, the particular vowel confusions that arise when identification is incorrect have not been reported for adult listeners and only preliminary data have been reported for child listeners. Thus, the following is a review of conversational speech vowel confusions in adult listeners and clear speech vowel confusion pilot data reported for a child listener. Adults' vowel confusions in conversational speech often include vowels that are proximal in vowel space (Bunton & Story, 2009; Cutler et al., 2004; Neel, 2008). Neel (2008) reported that AE-speaking adults identified AE /ɛ, æ, ɑ, ʌ/ with less accuracy than AE /i, ɪ, e, o, ʊ, u/ in /hVd/ context (in quiet). When confusions occurred, the adults most often identified /ɛ/ as /æ/ and /æ/ as /ɛ/. They confused /ʌ/ with /ɑ/ most frequently and /ɑ/ with either /ɪ, ɛ, æ, o/. Cutler et al. (2004) reported similar confusions when AE-speaking adults identified syllables that contained the AE target vowels /i, ɪ, eɪ, ɛ, æ, ɑ, ʌ, ɔ, oʊ, ʊ, u, aɪ, ɔɪ, aʊ/ in VC segments produced by one female AE-talker and presented in multispeaker babble at 0 dB SNR. Adults confused /ɛ/ and /æ/ and /ʌ/ with /ɑ/. Additionally, adults frequently identified /ɑ/ as /ɔ/ (Cutler et al., 2004). Bunton and Story (2004) documented AE adult listeners' identification of synthetic vowels in quiet. Adults most frequently identified /ɪ/ as /e/ or /ɛ/ and identified /ɑ/ as /æ/ or /ɔ/. Adults also identified /ɛ/ as /æ/ but less frequently identified /æ/ as /ɛ/. Consistent with the adult conversational speech vowel confusions reported, preliminary data regarding child clear speech vowel confusions demonstrate confusions among vowels close in vowel space. Leone et al. (2011) reported on the performance of a single child listener, who repeated /æ/ for /ɛ/ in 33% of the clear speech trials, but repeated /ɛ/ for /æ/ in only 3% of the trials. In 22% of trials, the child listener repeated /ɑ/ for /ʌ/ and in 11% of trials, he repeated /ʌ/ for /ɑ/. He repeated /i/ as /ʌ/ most frequently (39%). Particular vowel confusions were explored in the present study and compared with this literature.

1.3.3. Clear-speech benefit

Several studies have documented a clear speech intelligibility benefit for a variety of talkers, utterances, and listening groups in various listening situations. Regardless of who is speaking clearly, how clear speech is being elicited, or what is being said, most studies agree that a clear speech advantage is present during many listening situations. The following is a description of studies that have documented a clear-speech benefit for adults with normal hearing and hearing loss, and children with and without learning disabilities. For a summary of clear speech studies through 2004, see table 9.1 in Uchanski (2005), which lists clear speech studies with corresponding clear speech advantage percentage points.

In real-life scenarios, speech is seldom listened to in a completely quiet environment. Clear speech has shown to benefit adult listeners with normal hearing and with hearing loss in different types of noise and SNRs (Ferguson, 2004; Payton, Uchanski, & Braida, 1994; Uchanski, 2005). Payton et al. (1994) presented conversational and clear nonsense sentences embedded in white noise in SNRs of 9.5, 5.3, and 0 dB to normal-hearing adults, who were asked to identify key words. The adults scored 21 percentage points higher when listening to clear speech than when listening to conversational speech. A similar benefit was reported by Bradlow and Bent (2002), who presented adults with semantically-intact sentences embedded in white noise in -4 dB and -8 dB SNRs. Bradlow and Alexander (2007) also found a clear-speech benefit when testing normal-hearing adult listeners listening to sentences in a slightly more favorable SNR (-2 dB).

Clear speech has also been found to be advantageous when it is presented at a normal rate of speech, as opposed to the slower rate at which it is typically produced. Krause and Braida (2002) elicited clear speech in slow and normal speaking rates from talkers with public speaking experience in order to create clear speech stimuli with different speech rates. Acoustic cues could thus be analyzed without factoring in speech rate. Normal-hearing adult listeners were presented nonsense sentences, embedded in speech shaped noise at -4 dB SNR. The clear speech at a slow rate was most advantageous to listeners, averaging 63% key words correct; however, the clear speech at a normal rate was also significantly advantageous to normal hearing listeners, averaging 59% key words correct, suggesting that rate may not be crucial to the clear-speech benefit for normal-hearing listeners. Using the same clear speech stimuli, Krause and Braida (2009) performed a study on normal hearing adults' perception in nonsense sentences embedded in a SNR of -1.8 dB. The listeners identified key words significantly more accurately for the sentences produced in clear speech mode than the sentences in conversational mode. Krause and Braida also reported a significant clear speech advantage when three adults with hearing loss listened to the same clear speech stimuli as the normal-hearing adults. Similarly, additional studies have reported a significant clear speech advantage over conversational speech for adults with normal hearing and with hearing loss listening to clear speech nonsense sentences in noise (Payton et al., 2004; Picheny et al., 1985). Overall, studies have demonstrated that as the listening environment becomes more degraded, the clear-speech benefit becomes greater (Bradlow & Bent, 2002; Payton et al., 2004).

A clear-speech benefit specifically for vowels in noise has also been reported for normal-hearing adults (Ferguson, 2004; Ferguson & Kewley-Port, 2002; Rogers et al., 2010). As previously stated, vowels contribute significantly to intelligibility (Kewley-Port et al., 2007). In a study by Ferguson and Kewley-Port (2002), ten /bVd/ words in 12-talker babble in -10 dB SNR spoken by one male talker were recorded. The normal-hearing adults selected the key word (in a field of 10 sets of key words) that represented the vowel they heard. Normal-hearing adults showed a clear speech advantage over conversational speech for vowels with some vowels more aided by clear speech than others. Ferguson and Kewley-Port (2002) reported that /e, ɛ, æ, ʌ, o, ʊ, u/ are more intelligible when produced in clear speech than in conversational speech, whereas /ɪ, ɑ, i/ do not show a clear-speech benefit due to the highly accurate performance on conversational speech stimuli. Ferguson (2004) extended the results of this study by increasing the number of talkers to forty-one while using the same methods as described previously. Results again indicated an overall clear-speech vowel intelligibility benefit. Using the clear-speech database recorded in Ferguson (2004), Ferguson and Kewley-Port (2007) investigated which acoustic features were associated with the documented clear speech-benefit for vowels. Longer vowel duration, a sufficient increase in vowel space, and a raised F2 for front vowels were the measures that were associated with greater intelligibility. Furthermore, Rogers et al. (2010) presented six /bVd/ isolated syllables in multi-talker babble in a -8 dB SNR spoken by thirteen AE monolingual female talkers to a group of adult listeners. The adults indicated the vowel they heard on a computer screen. The adults identified clear speech vowels in noise significantly more accurately than conversational speech vowels in noise. Rogers et al. (2010) also found that some clear speech vowels benefit listeners more than others, but reported different vowels than Ferguson & Kewley-Port (2002). Rogers et al. (2010) reported that /ɛ, ɑ/ were identified significantly more accurately than conversational speech /ɛ, ɑ/, but a significant clear-speech benefit was not found for /i, ɪ, e, æ/. Furthermore, Ferguson and Kewley-Port (2002) reported the largest clear-speech benefit for adults was for /æ/ whereas Rogers et al. (2010) reported adults' benefiting the most from clear speech /ɑ/.

To the author's knowledge, Ferguson and Kewley-Port (2002) provide the only study that does not suggest a benefit of clear speech. In this study, ten /bVd/ words in 12-talker babble in -3 dB SNR spoken by one male talker were presented. Adults with sensorineural hearing loss selected the key word (in a field of 10 sets of key words) to indicate the vowels they perceived. The study found no clear-speech benefit for the elderly participants with hearing loss. The authors posit that individuals with hearing loss may rely on different acoustic cues for accurate vowel identification from those relied on by normal-hearing listeners. Additionally, the participants' sensorineural hearing loss may have impacted their ability to perceive high F2 frequencies . Because clear speech increases F2 values for some vowels, the participants in the study may have missed some acoustic cues and therefore did not demonstrate a clear-speech benefit.

Only two studies, to the author's knowledge, have examined how children perceive clear speech in noise: Bradlow et al. (2003) and Riley and McGregor (2012). Bradlow et al. (2003) examined key word repetition accuracy of school-age children listening to sentences in noise, at -4 dB and -8 dB SNRs in both clear and conversational speech. Broadband white noise was used to mask the signal at all frequencies. The investigators presented adults' simple declarative sentences containing three to four key words to children with and without learning disabilities. Children were instructed to repeat each sentence while experimenters noted key words that were repeated accurately. Children with learning disabilities and typically-developing children performed significantly more accurately (8.8% and 9.2% respectively) when the sentences were presented in clear speech (Bradlow et al., 2003) than when they were presented in conversational speech. Both groups benefited more from clear speech in the -8 dB SNR condition than in the -4 dB SNR condition, suggesting that as the listening condition becomes more adverse, the clear-speech benefit increases. Similarly, Riley and McGregor (2012) showed a clear-speech benefit for school-age children listening to conversational and clear speech narratives that contained target words embedded in white noise in a -8 dB SNR produced by one female talker. Child listeners selected the picture (from a field of four on a computer screen) that represented the target word they heard. Results revealed that children identified more accurate word productions in clear speech than conversational speech.

The cognitive and linguistic processes utilized in the perception of isolated words differ from those utilized in sentence perception (Grant & Seitz, 2000; Miller, Heise, & Lichten, 1951). To the author's knowledge, clear speech studies involving children as listeners have been performed only using sentences with contextual cues and real words. For example, Bradlow et al. (2003) presented sentences with semantic, syntactic, and pragmatic information to children aged 8.1 to 12.5 years and documented a clear-speech benefit. However, the impact of clear speech in vowel perception is not known for children. Because listeners' lexical representations of real words may influence speech sound recognition (Strand & Sommers, 2011), especially in degraded listening conditions (Linden, Stekelenburg, Tuomainen, & Vroomen, 2007), lexical effects may confound results. In contrast, nonsense items force the listener to rely on only phonetic information because lexical information cannot be retrieved (Strange & Schafer, 2008). The use of nonsense words differing only in the vowel allows for an examination of the impact of clear speech on vowel perception. Investigating the possibility of a clear-speech benefit when contextual cues are absent may provide further understanding of a potential clear-speech benefit and insight into how talkers can enhance their speaking styles in order to be understood in different communicative contexts.

1.4. Summary

In summary, a clear speech intelligibility advantage has been documented for a number of listening groups. One of the first accounts of a clear-speech benefit was demonstrated for adults with hearing loss (Picheny et al., 1985). Subsequently, researchers have demonstrated a clear-speech benefit for other listening populations. A handful of studies have documented adults' greater identification accuracy of clear speech sentences than of conversational speech sentences (Bradlow & Alexander, 2007; Bradlow & Bent, 2002; Kraus & Braida, 2002, 2009; Payton et al., 1994). For clear speech vowels, a subset of studies has been conducted examining adults' perception (Ferguson, 2004; Ferguson & Kewley-Port, 2002). Regarding children's perception of clear speech, one study provided evidence that clear speech sentences (with semantic and syntactic cues) are more intelligible than conversational speech sentences with the same cues for children with and without learning disabilities (Bradlow et al., 2003). In the research domain of speech in noise, there is some evidence that children perceive speech in noise less accurately than adults, and the younger the child and the more challenging the SNR, the poorer the perceptual accuracy (Bradley & Sato, 2008; Johnson, 2000; Stuart et al., 2006). Nishi et al. (2010) documented these results for children's consonant perception. However, little is known about children's vowel perception in noise. If clear speech vowels enhance intelligibility for children listeners, adults' use of clear speech may be supported as a strategy to enhance their communication with children.

1.5. The present study: questions and predictions

The present study addressed the following research questions:

Does AE school-age children's repetition accuracy of vowels in noise vary as a function the following:

a. the speaking style (conversational vs. clear)?

b. the particular vowel? (/ɛ, æ, ɑ, ʌ/)

c. the talker?

Acoustic differences between clear and conversational vowels were also examined through acoustic analysis. The acoustics are discussed with regard to perceptual findings in the Discussion section.

It was predicted that the repetition accuracy of AE vowels would be higher in clear speech (Bradlow et al., 2003; Ferguson & Kewley-Port, 2002). Furthermore, it was hypothesized that /ɛ/ would be identified with the least accuracy in clear speech (Ferguson & Kewley-Port, 2002; Leone et al., 2011) with confusions among vowels proximal in vowel space occurring. It was also predicted that some vowels will be aided more by clear speech than others (Ferguson & Kewley-Port, 2002; Rogers et al., 2010). Talker differences were expected, with some talkers providing a larger intelligibility benefit to child listeners than others (Bradlow et al., 2003; Ferguson, 2004; Uchanski, 2005). Acoustic analysis was predicted to reveal that vowels in clear speech would be characterized by longer durations, larger F1/F2 vowel space, and be more dynamic (Ferguson & Kewley-Port, 2002) than vowels in conversational speech.

Chapter 2. Method

2.1. Adults' identification task

2.1.1. Stimulus materials and procedures

Test materials contained 4 vowels (/ɛ, æ, ʌ, ɑ/) in /gəbVpə/ trisyllables in sentences (see Strange et al., 2007). Ferguson and Kewley-Port (2002) report that /e, ɛ, æ, ʌ, o, ʊ, u/ are more intelligible when produced in clear speech than in conversational speech, whereas /ɪ, ɑ, i/ do not show a clear-speech benefit due to the highly accurate performance on conversational speech stimuli. However, Rogers et al. (2010) reported that clear speech /ɛ, ɑ/ were identified significantly more accurately than conversational speech /ɛ, ɑ/, but a significant clear-speech benefit was not found for /i, ɪ, e, æ/. Based on these results, the vowels /ɛ, æ, ʌ, ɑ/ were selected for the current study. Nonsense words were included, as opposed to real words, to minimize any lexical effects (Neuman & Hochberg, 1983).

Four native monolingual American-English female adult talkers from the New York tri-state area were recorded producing the trisyllables /gəbVpə/ embedded in the carrier phrase "Five _____ this time." Talkers were recorded in a sound-treated booth in the Speech Production and Perception Lab at Teachers College, Columbia University, with the experimenter in an adjoining room in visible contact. The experimenter provided the talker with directions using an intercom and listened to the recording input over Sennheiser HD 280 pro headphones. Talkers were instructed to read four lists of utterances (Five /gəbVpə/ this time) in each speaking style. (See Appendix A for protocol.) Protocols consisted of randomized lists of 12 utterances. The first utterance and the last utterance contained the same target vowel and the final utterance was discarded to control for list-final intonation effects. Each utterance was preceded by an identifying number. Instructions for producing conversational sentences were "Speak at a normal rate, as if speaking with someone who is very familiar with your voice." For clear speech, talkers were instructed to, "Speak as if talking with someone with a hearing loss" (Bradlow et al., 2003; Ferguson & Kewley-Port, 2002). All conversational stimuli were recorded prior to clear speech stimuli. If an utterance contained irregular pronunciation, rate, prosody, vocal quality, or noise, the experimenter asked the talker to repeat the stimulus. Output was recorded through a Shure (SM58) microphone placed 15 cm from the talker's mouth and passed through a Shure (Prologue 200M) mixer to a Turtle Beach Riviera sound card of a Dell Pentium 4 desktop computer using Soundforge™ 8.0 software, with a sample rate of 22,050 Hz, 16-bit resolution, on a mono channel.

Female talkers were included because females' clear speech productions have been found to be more intelligible than males' (Bradlow, Torretta, & Pisoni, 1996) and females produce a larger clear-speech benefit (Bradlow et al., 2003; Uchanski, 2005). In addition, according to the National Center for Education Statistics, 76% of teachers in the United States are female (2009). Thus, the participants for the proposed study, school-age children, likely spend a large proportion of their time listening to female voices. A carrier phrase was used rather than words in isolation, as vowels in sentence materials are produced and perceived differently from vowels in words in isolation (Strange, Bohn, Nishi, & Trent, 2005; Strange et al., 2007), and sentences may be more representative of everyday speech. Sentences without semantic cues were utilized to minimize the effects of lexical knowledge (Neuman & Hochberg, 1983).

All talkers completed a language background questionnaire (see Appendix B). Talkers who participated in the study were monolingual speakers of American English and had minimal exposure to speaking and listening to other languages, had no history of speech and language disorders, and passed a bilateral hearing screening at 20 dB at 500, 1000, 2000 and 4000 Hz. A talker with a similar language profile to those in the experimental condition was recorded for the familiarization task.

For each vowel, the second and third recording was utilized unless it was characterized by background noise or dysfluent speech. Multiple tokens of the utterances were used to tap into categorial perception (Gottfried, 1984; Levy & Strange, 2008; Levy, 2009) rather than simply physical discrimination. To eliminate amplitude differences between talkers and speech styles, the mean RMS value was calculated across all stimuli and then all stimuli were scaled to this amplitude using SoundForge™ 8.0 software. Stimuli were then mixed with speech-shaped noise using the Praat v. 5.2.22 program.

To determine an appropriate SNR for the adult listeners, pilot data for this study were collected from three adults who listened to target stimuli in -6 dB and -10 dB SNRs in a sound treated booth and selected their response choices on a computer using the Paradigm v.1.0.2 program (Tagliaferri, 2011). Results, graphed in Appendix C, show a larger conversational speech versus clear speech difference for the -10 dB condition than the -6 dB condition. Results also indicate a ceiling effect for the clear -6 dB condition. The adults accurately identified 94% of the vowels in this condition. Thus, to avoid a ceiling effect and be able to measure differences, if present, between the identification of clear vs. conversational speech, -10 dB SNR was selected for the adults' identification task. A total of 256 experimental stimuli (4 talkers X 4 target vowels X 2 tokens X 2 speaking styles X 4 repetitions) were created. (See Appendix D for design of the adults' identification task.)

For the adults' identification task, stimuli were entered into the Paradigm software program. During the experiment, ten AE-speaking adult listeners identified each target stimulus and rated the clarity of the vowels on a 9-point Likert scale (Southwood & Flege, 1999) ranging from "least clear" to "most clear." (See Appendix E for instructions.) Prior to the experimental condition, the adult listeners completed a task familiarization block followed by a stimulus familiarization block. Familiarization tasks are described in Appendix D.

2.2. Children's repetition task

2.2.1. Stimulus materials and procedures

Variability in clear-speech production has been noted across an individual talker's utterances (Krause & Braida, 2002) and across talkers (Ferguson, 2004).Therefore, stimuli produced by the two talkers in the adults' identification task whose clear speech vowels were identified with the most accuracy and rated as most clear were prepared for presentation in the children's repetition task in order to provide an opportunity to detect possible differences between clear speech and conversational speech stimuli. To determine an appropriate SNR for the child listeners, pilot data were collected from two 6-year-old females who listened to target stimuli in 0 dB and +2 dB SNRs in a sound treated booth and repeated what they heard. An adult experimenter noted the response choice on a computer using Paradigm software. Results show a ceiling effect: both children accurately repeated 100% of the conversational and clear speech vowels (Leone et al., 2011). Additional preliminary data, graphed in Appendix F, reveal a 7-year-old boy's difficulty identifying the target vowels in -8 dB SNR and greater accuracy in the -4 dB condition (Leone et al., 2011). Thus, to avoid a ceiling effect and be able to measure differences, if present, between the identification of clear versus conversational speech, -6 dB SNR was selected. A 0 dB SNR condition served as a control measure, to ensure that the children were attending to the task.

During the familiarization and experimental tasks, all stimuli were presented over two loudspeakers (Altec Lansing BXR 1320) placed approximately 2 ft away from the listener at a mean output level of 65 dB SPL as verified by a Galaxy CM-140 SPL meter placed 30 cm from the microphone. This verification procedure was repeated at the end of the recording session. Child listeners' output was recorded through a Shure (SM58) microphone placed approximately

15 cm from the child listeners' mouths and was monitored by an experimenter over Sennheiser HD 280 pro headphones and recorded using SoundForge™ 8.0 software. As shown in the flowchart in Appendix G, each child listener was seated across from an experimenter, while another experimenter was seated at a computer behind the child listener. Child listeners were tested individually in a sound-treated booth.

Stimulus presentation was controlled by Paradigm v.1.0.2 program (Tagliaferri, 2011). In an 11-alternative closed-set response paradigm, the following response options were displayed on the computer monitor: gabeepa, gabuppa, gabeppa, gabappa, gabippa, gabaypa, gaboopa, gabUpa (vowel in "book"), gaboapa, gabawpa, gaboppa. The children were given the following directions: "We're going to listen to some sentences with silly-sounding words. I want you to listen and then say exactly what you heard." (See Appendix H for instructions.) The experimenter played each stimulus using Paradigm software, the child listener repeated the stimulus, and the experimenter clicked on the perceived response on Paradigm. (See Appendix I for screen diagram.) An additional adult listener identified the children's vowels at a subsequent time. Child listeners were provided as much time as needed to respond. If repetition of a stimulus was required because the child listener was not attending or because external noise was present, the experimenter seated at the computer replayed the stimulus on the computer. Between blocks, another experimenter provided encouragement both verbally and through games.

Prior to the onset of the experimental condition, task and stimulus familiarization procedures trained listeners to repeat appropriate responses and become familiar with the stimuli. (See Appendix J for design of the children's repetition task.) In the task familiarization, child listeners completed the same 24-trial task familiarization block that the adult listeners had completed, which included conversational and clear vowels /o, u, ɪ/ in trisyllables /gəbVpə/ in quiet produced by a talker different from the talkers who produced the experimental stimuli (1 talker X 3 target vowels X 2 tokens X 2 speaking styles X 2 repetitions). Child listeners were required to achieve 90% accuracy during the 24-trial task familiarization block. (All met this requirement.)

Each child listener then completed a 36-trial stimulus familiarization block consisting of one representation of each stimulus. Data from this block were discarded. After the stimulus familiarization, child listeners heard 4, 36-trial blocks (4 vowels X 2 talkers X 1 experimental SNR X 2 speaking styles X 2 tokens X 4 repetitions + 1 vowel X 2 talkers X 1 control SNR X 2 speaking styles X 2 tokens X 2 repetitions), totaling 144 responses. Each child listener completed 16 repetitions of each vowel, 8 repetitions for each talker. All stimuli were randomized within the blocks. Child listeners were given a break between blocks, in which they played an age-appropriate game with the experimenter for approximately 5 minutes. Total testing time was approximately 2 hours.

2.3. Stimulus verification

Two forms of stimulus verification were performed to corroborate the hypothesized differences between clear and conversational speech. In the first stimulus verification task, eleven monolingual AE adult listeners were presented with recordings of clear and conversational speech from four talkers; the adult listeners performed a forced-choice identification task to evaluate the difference between the recorded clear and conversational speech. Furthermore, as described, the results of the adult listeners were used to identify two talkers who showed a large intelligibility difference between their clear and conversational speech. The recordings of the two most differentiated talkers were then presented to the child listeners. The reduction from four talkers to two talkers aimed to shorten the task as children are typically less able to attend to a cognitive-demanding task for a long period of time than adults (McKay et al., 1994). In the second stimulus verification task, acoustic analysis was completed to verify that the stimuli recorded as conversational speech were acoustically different from the stimuli recorded as clear speech. The following section is a description of the stimulus verification outcomes, beginning with results from the adults' identification task and followed by details about the acoustic analysis methods and findings.

2.3.1. Adults' identification task results

Eleven monolingual AE adults with normal hearing completed the adult forced-choice identification task. (See Appendix K for adult participant characteristics.) Performance was evaluated based on the total identification correct out of 256 possible points (288 total responses – 32 control responses). The total identification correct score was then calculated for each speaking style subcategory (conversational speech and clear speech). Control data were tallied separately. All participants identified all control data with 100% accuracy. Descriptive results, graphed in Appendix L, indicate that clear speech vowels were identified more accurately than conversational speech vowels (77.9% and 45.1% for clear and conversational speech, respectively). McNemar's test for paired proportions indicated that clear speech was identified significantly more accurately than conversational speech for all talkers (talker 1: χ^2= 105.35(1), $p < .001$, *OR*= 8.11; talker 2: χ^2= 41.86(1), $p < .001$, *OR*= 3.45; talker 3: χ^2= 97.59(1), $p < .001$, *OR*= 6.36; talker 4: χ^2= 75.63(1), $p < .001$, *OR*= 4.66). Talkers 1 and 3 demonstrated the largest clear speech effect, that is, the largest differences between conversational and clear speech (*OR* = 8.11 for talker 1 and *OR* = 6.36 for talkers 3; see Appendix M). Additionally, median perceptual ratings (where 1 = least clear and 9 = most clear) were higher for talkers 2 and 4 (mdn. rating = 7 and 8, respectively) than talkers 1 and 3 (mdn. rating = 4 and 5, respectively). In summary, both perceptual rating and forced-choice identification revealed that talker 1 and talker 3 exhibited the largest clear speech effect. Thus, stimuli from talker 1 and talker 3 were used as the talkers for the children's repetition task.

2.3.2. Acoustic analysis: method

Acoustic analysis was performed using two speech analysis programs: Wavesurfer v. 1.8.5 and Praat v. 5.2.22, to examine differences between clear and conversational speech. Each utterance (i.e., of a total of 32 utterances) was analyzed for utterance duration and vowel duration. Fundamental (F_0), F1, and F2 frequencies of the target vowel were also examined.

Each vowel token was obtained after manually determining the beginning and end of each phrase. The beginning of each phrase "Five gaCVCa this time" was defined by the first mark of frication energy for /f/ and the end of each phrase was defined by the end of the voicing bar for /m/. The onset and offset of the syllable containing the target vowel were also manually determined on the basis of the following definitions: Syllable onset was defined as the release burst of the /b/, which was visually correlated with the spike of acoustic energy on the spectrogram, and syllable offset was defined as the beginning of closure of /p/, which was noted by a decrease in periodic energy in the higher formants (F2, F3) on the spectrogram. Thus, vowel duration included the entire gesture from release of the preceding consonant to the beginning of full closure of the following consonant.

To obtain a measure of fundamental frequency, F_0 was measured at the vowel's temporal midpoint and mean F_0 was calculated using Praat, which provides detailed information about F_0 contour. Values for the first three formants (F1, F2, F3) for a 25 ms window were calculated using formant tracking in Wavesurfer at the temporal midpoint (50% point) between onset and offset of the syllable using linear predictive coding (LPC) analysis. Comparisons were made to normative data (Hillenbrand, Getty, Clark, & Wheeler, 1995; Peterson & Barney, 1952; Strange et al., 2007). When estimated formants were judged to be erroneous, hand corrections were made. Corrections were based on the experimenter's judgments from comparisons of Fast Fourier Transform (FFT) spectra and LPC formant tracks superimposed on the spectrographic display. All stimuli were checked by a second judge, who provided a second estimate for all formant values. If a discrepancy was found, a third judge resolved the conflict.

2.3.2. Acoustic analysis: results

Table 1 lists a comparison of 7 acoustic-phonetic measurements for both talkers (talker 1 and talker 3) used in the children's repetition task for each vowel. The mean utterance and vowel durations in clear and conversational speech and the clear-minus-conversational difference for each talker are shown. Consistent with other clear speech studies (Bradlow et al., 2003; Ferguson & Kewley-Port, 2007) both talkers showed a large increase in utterance duration and vowel duration for clear speech relative to conversational speech. Talker 3's utterance length and vowel length for clear speech were greater than those of talker 1's for all tokens. Both talkers produced clear speech /æ/ (talker 1: .210 seconds, talker 3: .235 seconds) with longer duration than the other 3 target vowels (/ɛ, ʌ, ɑ/).

Acoustic analysis showed that the mean F_0 and F_0 at the vowel's temporal midpoint increased for clear speech relative to conversational speech for all tokens. In conservational speech, talker 3's mean F_0 and average F_0 at midpoint were higher than talker 1. In clear speech, talker 1 and talker 3's clear speech mean F_0 and average F_0 at midpoint were comparable. Thus, differences between clear speech and conversational speech were larger for talker 1 than talker 3.

Acoustic analysis for both talkers showed that the F1 was higher in clear speech than conversational speech for all stimuli. This is consistent with clear speech vowel studies (e.g., Ferguson & Kewley-Port, 2007) suggesting that the tongue lowers during clear speech vowel production. Acoustic analysis of both talkers showed that the F2 was higher for clear speech front vowels (/ε, æ/) than conversational speech front vowels and lower for clear speech back vowels (/æ/) than conversational back vowels. This is consistent with clear speech vowel studies (e.g., Ferguson & Kewley-Port, 2007) suggesting that the tongue moves farther forward during clear speech front vowel production and farther back for clear speech back vowel production. Acoustic analysis of /ʌ/ showed varying F2 values. Talker 1's F2 for /ʌ/was lower for clear speech when compared to conversational speech, while talker 3 produced a higher F2.

Both talkers increased both F1 and F2 range for clear speech vowels more than for conversational speech vowels, revealing a larger vowel space for clear speech vowels. Figure 1 presents the mean F1/F2 values for talker 1 and talker 3 and shows a comparison of vowel space for conversational and clear vowels. Overall, the acoustic analysis revealed duration increases, F_0 increases, and a larger vowel space for both talkers when they produced clear speech than when they produced conversational speech.

2.4. Participants: child listeners

Child listeners' legal guardians were given Institutional Review Board (IRB) consent forms for review and signature. (See Appendix N for IRB approval letter.) Questions were answered by the primary investigator. A language background questionnaire (Appendix O) was completed by the guardians. Listeners were 15 AE-speaking, school-age children, ages 5.0-8.5, with normal hearing. (See Appendix P for child listener characteristics.) This age range was selected to tap into young children's speech perception skills. Johnson (2000) indicated that children's ability to perceive vowels and consonants increases as they age and concluded that they can identify vowels in noise with adult-like accuracy by age 10. AE was the child listeners' native and only language spoken at home. All child listeners passed a hearing screening at 20 dB at 500, 1000, 2000 and 4000 Hz and had no history of a speech or language disorder. The *Arizona Articulation Proficiency Scale* (Fudala, 2000), which tests vowel and consonant production, was administered prior to data collection and used to identify children with articulation disorders. All child listeners had typical articulation, as evidenced by a score no more than one standard deviation below the mean on the *Arizona Articulation Proficiency Scale* (Fudala, 2000).

Chapter 3. Results

3.1. Data analysis

A total of 2,160 responses were collected from the 15 child listeners (144 trials from each listener). All listeners' responses were totaled and a percent correct score was computed (i.e., number of accurate responses/total responses). Control data were then tallied separately (16 trials per listener). A repetition correct score using all 1,920 experimental trials (2,160 trials – 240 control trials) was obtained. A repetition correct score was also calculated for each subcategory of the two testing conditions (i.e. conversational speech and clear speech). Data were further divided by talker as well as by each vowel. Each repetition correct score was converted to a percentage correct score. Descriptive and nonparametric statistics were performed on all data. All statistical analyses were performed using R software, version 2.15.1, the lme4 package (R Development Core Team, 2010).

To test whether vowel repetition accuracy differed significantly between clear and conversational speech or as a function of talker or a particular vowel, children's repetition correct scores were analyzed using a mixed effects logistic model with crossed random effects. Mixed effects models have been reported to provide more reliable and authentic results for categorical outcome variables (e.g., the forced-choice variables used in this study) than analysis of variance methods (Jaeger, 2008). In addition, the use of mixed effects modeling has been documented as demonstrating higher statistical power and robustness than repeated-measures analysis of variance techniques in speech perception studies (Ferguson, 2012). (See Appendix Q for a description of mixed effects models.) The final model used in the present study included speaking style, talker, and vowel as fixed effects and listener and trial as random effects (see Appendix R).

For the mixed effects logistic model created for the data in this study, listeners and trials were considered random effects. Identifying listeners as a random effect allowed the model to consider the sample of children used in this study as a random sample from a larger population. Identifying trials as a random effect allowed the model to consider the talker vowel recordings as a sample of a larger population of possible vowel productions (i.e., each time a talker pronounces a vowel, the talker may say the vowel in a slightly different way) (Baayen, Davidson, & Bates, 2008). Appendix R lists all models attempted and the final model selected.

3.2. Organization of results section

The remaining part of this chapter is organized as follows: a summary of the results regarding effects of speaking style on children's repetition accuracy of vowels in noise is provided, followed by a description of effects on particular vowels, and a summary of the effects of the particular talker on vowel repetition accuracy. The interactions of vowel and speaking style and of talker and speaking style are described. Confusion matrices demonstrating particular vowel misperceptions are presented.

3.3. Speaking style effects on vowel repetition

Regarding Research Question 1, speaking style effects on repetition accuracy, child listeners repeated clear speech vowels more accurately than conversational speech vowels.

Clear speech vowels were repeated with 87% accuracy (*SD* = 12.0) and conversational speech vowels with 59% accuracy (*SD* = 14.5), with a mean difference between the speaking styles of 27% (see Figure 2). Mixed effects logistic regression confirmed that this effect of speaking style (clear vs. conversational speech) on vowel repetition was statistically significant ($z = 6.34$, $p < .001$; see Appendix S for the mixed effects logistic model analysis).

In a further investigation of the effects of speaking styles on vowel repetition, an analysis separated child listeners into two groups according to age (younger group = ages 5.0 – 6.7; older group = 6.8 – 8.5). The younger group, which included 5 child listeners, repeated clear speech vowels with 83% accuracy and conversational speech vowels with 54% accuracy with a mean difference between speaking styles of 29%. The older group, which consisted of 10 child listeners, repeated clear speech vowels with 89% accuracy and conversational speech vowels with 65% accuracy with a mean difference of 25%. This finding suggests that clear speech benefited both younger and older child listeners similarly.

The analysis also separated the children into groups according to gender. The 8 male child listeners repeated clear speech vowels with 85% accuracy and conversational speech vowels with 58% accuracy with a mean difference between speaking style of 23%. The 7 female child listeners repeated clear speech vowels with 89% accuracy and conversational speech vowels with 61% accuracy, with a mean difference between speaking styles of 28%. This finding suggests that clear speech benefited both male and female child listeners similarly.

3.4. Particular vowel effects on vowel repetition

In response to Research Question 2, particular vowel effects on repetition accuracy, listeners repeated /ɛ/ with the most accuracy (83%, *SD* = 11.1), followed by /æ/ (79%, *SD* = 9) then /ʌ/ (67%, *SD* =16.4), and /ɑ/ with the least accuracy (63%, *SD* = 12.2). The control vowel /i/ was repeated without difficulty (100% accuracy for all listeners), indicating that all child listeners were on task. Pairwise comparisons revealed statistically significant differences in children's repetition accuracy of the four target vowels. (See Appendix S for pairwise comparisons.) The findings, represented as percent correct values in Figure 3, indicated that /ɛ/ was repeated significantly more accurately than /ʌ/ (z = 2.02, $p < .05$) and /ɑ/ (z = 3.04, $p < .001$), but not significantly more accurately than /æ/ (z = 0.47, p = .641). Also, /æ/ was repeated significantly more accurately than /ɑ/ (z = 2.13, $p < .05$) and /ʌ/ (z = 2.30, $p < .05$). The repetition accuracy difference between /ʌ/ and /ɑ/ was not statistically significant (z = 0.47, p = .637).

3.4.1. Particular vowel effects for each speaking style

Overall vowel repetition accuracy varied as a function of speaking style (see Figure 4). Listeners repeated /ɛ/ with the most accuracy and /ɑ/ with the least accuracy in both speaking styles (92%, 78% for clear speech vowels and 73%, 41% for conversational speech vowels, respectively). Listeners' repetition accuracy increased the most for /ʌ/ (clear speech vowel – conversational speech vowel difference = 40%) and thus showed the largest clear-speech benefit. There was no significant interaction between speaking style and individual vowels (x^2 (3) = 2.471, p = .481). (See Appendix R, model m8s2.) These results indicate that the clear speech benefit is not vowel-specific.

Confusion matrices for conversational and clear speech vowels indicate which vowels were frequently confused with one another. Table 2 is a confusion matrix that represents the listeners' responses (options are in the top row) to each stimulus (listed in the left column) as percentages of the total repetitions for each target vowel. Consistent with preliminary studies (Leone et al., 2011), vowels close in vowel space were most often confused. When conversational and clear speech data were combined, child listeners most frequently produced the front vowels /ɛ/ as /æ/ (7%) and /æ/ as /ɛ/ (11%). When clear and conversational speech vowel repetitions were separated (Tables 3 and 4), confusions between /ɛ/ and /æ/ remained, especially in conversational speech. Listeners repeated /ɛ/ as /æ/ (12% for conversational vowels; 3% for clear vowels) and /æ/ as /ɛ/ for each speaking style (17% for conversational vowels; 6% for clear vowels).

Because listeners repeated /ɑ/ with the lowest accuracy, greater variability was noted in the response options when compared to the other target vowels. Listeners repeated /ɑ/ as /ʌ/ for 16% of the trials and as /æ/ for 11% of the trials. (See Figure 4.) The confusion for /ɑ/ as /ʌ/ was primarily unidirectional, as /ʌ/ was seldom repeated as /ɑ/ (6%). This unidirectional confusion was consistent for both clear and conversational speech; listeners repeated /ɑ/ as /ʌ/ more frequently in conversational speech (27%) and clear speech (6%) than /ʌ/ as /ɑ/ in conversational speech (7%) and clear speech (5%) (Tables 3 and 4).

3.5. Talker effects on vowel repetition

Regarding talker effects on vowel repetition accuracy (Research Question 3), child listeners, on average, repeated talker 1's vowels with 79% accuracy and talker 3's vowels with

67% accuracy. This finding indicates that, on average, talker 1 was more intelligible for listeners in both speaking styles more than talker 3. The mixed effects logistic model confirmed a significant fixed effect of the talker on vowel repetition ($z = 3.13$, $p < .001$). Child listeners repeated talker 1's vowels significantly more accurately than talker 3's vowels.

3.5.1. Talker effects for each speaking style

The mean difference between repetition accuracy of clear speech vowels and conversational speech vowels was 28% for talker 1 and 27% for talker 2. Figure 5 shows child listeners' repetition accuracy differences for each talker and speaking style. Even though these empirical findings suggest a small difference (1%) between talkers, the mixed effects logistic model indicated a significant difference ($z = 2.10$, $p < .05$) after controlling for the random effects of listener and trial (see Appendix S). Acoustic analysis results, which were presented in the stimulus verification section, will be discussed in reference to the perception findings in the Discussion section.

Chapter 4. Discussion

4.1. Summary

The present study's main finding was that clear speech benefited child listeners' vowel perception in noise. That is, children repeated clear speech vowels produced by adults significantly more accurately than conversational speech vowels produced by adults. Significant vowel and talker differences were also observed. Repetition accuracy depended on the vowel. For example, front vowels (/ɛ, æ/) were repeated more accurately than central and back vowels (/ʌ, ɑ/). Furthermore, a larger clear-speech benefit was observed for talker 1 than talker 3. It was concluded that despite the observed repetition accuracy differences for particular vowels and talkers, child listeners benefited from clear speech for all vowels and for both talkers.

In this chapter, findings from the three major research questions are discussed. First, children's vowel repetition accuracy is described as a function of the speaking style, individual vowels, and the talker. Next, possible explanations and implications for the current study's results are presented. Lastly, limitations of the present study and future directions for clear speech research are suggested.

4.2. Speaking style

That finding that clear speech vowels were be repeated with higher accuracy than conversational speech vowels was predicted at the onset of the study. These results are consistent with those of Payton et al. (1994), who found that adults listening to nonsense sentences identified clear speech keywords with 21% higher accuracy than conversational speech keywords. Similarly, the present study reported a 27% repetition accuracy increase for clear speech vowels over conversational speech vowels. Bradlow et al. (2003) found that school-age children accurately repeated significantly more clear speech keywords in noise than conversational speech keywords in noise. Similarly, Riley and McGregor (2012) found a clear-speech benefit for school-age children listening to conversational and clear speech narratives in noise; children accurately identified significantly more words in clear speech than in conversational speech. The current study extended this previous research to include an examination of the repetition accuracy of vowels in noise.

A clear-speech benefit for adults identifying vowels in noise has been reported in the literature (Ferguson & Kewley-Port, 2002; Rogers et al., 2010). In this regard, Ferguson and Kewley-Port (2002) documented a 14% vowel identification increase for speaking style, Ferguson (2004) reported an 8.5% increase, and Rogers et al. (2010) reported a 5-7% increase. It should be noted that the adult studies (e.g., Ferguson & Kewley-Port, 2002) employed different procedures and used different stimuli from the methods and stimuli used in the present study. However, results from the present experiment expand the clear-speech vowel benefit to school-age children. The benefit appears to be larger for children (27%) than for adults, at least for the vowels /ɛ, æ,ʌ, ɑ/ in nonsense words. This difference in the degree of benefit provided by clear speech could be due to the reported differences between children's and adults' perception of speech in noise. Noise has a more detrimental effect on children's speech perception than on adults' perception (Crandell & Smaldino, 2000). Furthermore, children require higher SNRs than adults to identify speech segments and words in background noise than in quiet (Bradley & Sato, 2008; Johnson, 2000; Stuart et al., 2006). Because children perceive speech in noise less accurately than do adults, children may benefit to a greater extent from clear speech than do adults.

4.3. Differences among particular vowels

In the present study, when clear and conversational speech vowels were analyzed together, significant repetition accuracy differences were found amongst the four target vowels (/ɛ, æ, ɑ, ʌ/). Contrary to expectations derived from a preliminary study (Leone et al., 2011), the current study found /ɛ/ to be the vowel repeated with the highest accuracy when compared to the other three target vowels (/æ, ʌ, ɑ/). It is possible that the preliminary study's inclusion of only one child listener may have influenced these results.

The repetition accuracy differences amongst the four target vowels in this study suggest that /ɛ/ and /æ/ are easier for children to perceive accurately than /ʌ/ and /ɑ/. These results are consistent with those of adult identification studies involving conversational vowels in noise. When listening to AE vowels in CV or VC syllables embedded in 0 dB, 8 dB, and 16 dB SNRs, AE-speaking adults identified /ɛ/ and /æ/ more accurately than /ʌ/ and /ɑ/ (Cutler et al., 2004). Moreover, Bunton and Story (2009) found that when AE-speaking adults identified isolated synthetic productions of /ɛ, æ, ɑ, ʌ/ presented in quiet, /æ/ was identified with the most accuracy and /ʌ/ with the least accuracy, consistent with the present study's findings. Similarly, Neel (2008) found that when AE-speaking adult listeners identified vowels produced by 48 female talkers in /hVd/ context in quiet, the adult listeners identified /ʌ/ with the least accuracy when compared to /ɛ, æ, ɑ/. In summary, in both children and adults, /ɛ/ and /æ/ are repeated with more accuracy than /ʌ/ and /ɑ/. Thus, the adult literature and the present study indicate particular difficulty in perceiving central and back vowels.

4.4. Differences among particular clear speech vowels

The adult clear-speech vowel literature has found that certain vowels are more aided by clear speech than others. Ferguson and Kewley-Port (2002), for example, reported that clear speech /ɛ, æ, ʌ/ were identified significantly more accurately than their conversational speech counterparts, although a significant difference between clear speech /ɑ/ and conversational speech /ɑ/ was not found. Rogers et al. (2010) found similar results, reporting that /ɛ, ɑ/ were significantly aided by clear speech. In the present study, the interaction between speaking style and vowel was not found to be statistically significant for any of the target vowels (/ɛ, æ, ʌ, ɑ/). Thus, unlike for adults in previous studies, for children in the present study, no particular vowel was more aided by clear speech than any other vowel.

A different trend was also noted for adult and child listeners regarding the accuracy with which particular clear speech vowels were repeated. Ferguson and Kewley-Port (2002) reported that when adults identified clear speech vowels in /hVd/ context in noise, /ɛ/ was identified with less accuracy (76.4%) than /æ, ɑ, ʌ/. Furthermore, the largest clear-speech benefit for adults was noted for /æ/ (53.9%) when compared to /æ, ɑ, ʌ/ (Ferguson & Kewley-Port, 2002). In the present study, percent-correct performance for target vowels in each speaking style showed that child listeners repeated /ɛ/ with the most accuracy and /ɑ/ with the least accuracy in clear speech and the largest clear-speech benefit was shown when repeating /ʌ/. However, it should be noted that Ferguson and Kewley-Port (2002) used different stimuli and one talker.

Thus, although children and adults both benefit from clear speech, the difference between these populations' perceptual patterns for clear speech suggests that findings reported about adults' perception of clear speech vowels may not be applicable to children. Similarly, speech-in-noise perception findings suggest that children perceive speech in noise less accurately than adults (Nishi et al., 2010). Children may attend to different articulatory or acoustic cues from those attended to by adults when listening to clear speech vowels. Additionally, children's decreased ability to perceive speech in noise when compared to adults' may contribute to the difference in findings for children's vs. adults' perception of clear speech vowels.

4.5. Particular vowel confusions

Regarding vowel repetition confusions, as predicted, vowels that were proximal in vowel space were most often confused. These findings are consistent with those of adult studies of conversational speech vowels in quiet (Bunton & Story, 2009; Neel, 2008) and in noise (Cutler et al., 2004). The following is a summary of specific confusions described in other studies of the vowels /ɛ, æ, ɑ, ʌ/, which were targeted in the present study.

Neel (2008) reported that AE-speaking adults most often confused /ɛ/ with /æ/ and /æ/ with /ɛ/ when identifying conversational speech vowels in /hVd/ context in quiet. Similarly, adults also confused /ɛ/ and /æ/ when identifying syllables that contained the target vowels in the initial position presented in multispeaker babble at 0 dB SNR (Cutler et al., 2004). Bunton and Story (2004) reported that adult listeners identified /ɛ/ as /æ/ but less frequently identified /æ/ as /ɛ/. Children's confusion of /ɛ/ and /æ/ in the present study is similar to adult findings. The children most frequently confused /æ/ with /ɛ/ and /æ/ was repeated in error as /ɛ/ most frequently. (See Table 2 for confusion matrix.)

The present study also reported child listeners' confusion of /ɑ/ with /ʌ/ most frequently followed by /ɑ/ with /æ/. Adults' identification of vowels in quiet revealed the same trend (Neel, 2008). Furthermore, adults also identified /ɑ/ as /ɔ/ (Bunton & Story, 2009; Cutler et al., 2004). These differing results may be attributed to dialectal differences between the two studies' samples. The current study's child listener sample was from the New York area, where /ɑ/ and /ɔ/ can be classified as two distinct phonemes. In contrast, the adult listener samples included in the comparison studies were recruited from other areas of the United States (e.g., Arizona), where /ɑ/ and /ɔ/ are not distinguished (Dinkin, 2011). Lastly, the current study also reported that child listeners most often repeated /ʌ/ as /ɛ/. Adults, in contrast, most frequently confused /ʌ/ with /ɑ/ both in quiet (Cutler et al., 2004) and in noise (Neel, 2008). Overall, child listeners followed similar confusion trends as adult listeners for /ɛ, æ, ɑ/ confusions, but confusions differed for /ʌ/.

4.6. Particular vowel confusions for clear speech

As predicted, clear speech vowels close in acoustic vowel space were most frequently confused. To the author's knowledge, no previous study has provided a vowel confusion matrix for clear speech vowels, rendering comparisons to previous studies difficult. However, Leone et al. (2011) reported preliminary data for a child's confusions of clear speech vowels. (See Appendix T for confusion matrix.) The clear speech vowel confusions in the present study and those preliminary data (from a child not tested in the present study) showed that /æ/ was repeated as /ɛ/ more frequently than /ɛ/ repeated as /æ/. In the present study, clear speech /ɑ/ was most often confused with /æ/ followed by /ʌ/. In contrast, preliminary data showed /ɑ/ confused most frequently with /ʌ/. When clear speech vowel confusions are compared to conversational speech vowel confusions for child listeners in the present study, similar confusion trends between the two speaking styles are noted for /ɛ/ and /æ/. Table 4 displays a bidirectional confusion for clear speech /ɛ/ and /æ/, which is the same confusion pattern noted in Table 3 for conversational speech vowels. In contrast, a different confusion trend is noted for /ɑ/ and /ʌ/when clear speech vowel confusions are compared to conversational speech vowel confusions in the present study. Clear speech /ɑ/ was most often confused with /æ/ followed by /ʌ/, whereas conversational speech /ɑ/ was repeated as /ʌ/ more frequently than /æ/. Lastly, clear speech /ʌ/ was repeated as /ɛ/ and /ɑ/, whereas conversational speech /ʌ/ was repeated more frequently as /ɛ/ than /ɑ/. One conjecture regarding why confusions involving /ʌ/ followed a different pattern in clear speech from the pattern in conversational speech is that phonetically, the mid vowels remain mid vowels in clear speech, whereas the peripheral vowels become "more peripheral." That is, clear speech /ʌ/ remained a mid vowel in clear speech and did not change as much acoustically as did clear speech /ɛ, æ, ɑ/ (see Table 1 for acoustic analysis). Mid vowels appear to be less changed by clear speech than low vowels and therefore child listeners' confusions in clear speech for mid vowels appear to follow different patterns from those of low vowels. Acoustically, for example, F1 increased by approximately 94 Hz in clear speech for /æ/, but increased by only approximately 48 Hz for /ʌ/ in clear speech. Perhaps because both clear speech /æ/ and /ɑ/ showed a greater increase in F1 than did clear speech /ʌ/, child listeners confused /ɑ/ with the "lowered" vowel /æ/ and not with the mid vowel /ʌ/. This change in acoustic vowel space for the more peripheral vowels may also clarify why child listeners repeated /ʌ/ as /ɛ/ more frequently in conversational speech than clear speech. Because clear speech /ʌ/ remained a mid vowel and clear speech /ɛ/ was "lowered," acoustic vowel space between these two vowels increased and, thus, child listeners less frequently confused /ʌ/ with /ɛ/.

4.6.1. Clear speech vowel confusions and vowel duration

Durational differences were noted between clear speech vowels and conversational speech vowels. (See Table 1 for acoustic analysis results.) As predicted, both talkers produced longer clear speech vowels ($\bar{x}$ = .18 seconds) than conversational speech vowels ($\bar{x}$ = .11 seconds). The vowel perception literature has documented that lengthening or shortening vowel duration increases conversational speech vowel confusions for adult listeners. For example, in Hillenbrand, Clark, & Houde (2000), AE adults were asked to identify four sets of synthetic

/hVd/ syllables that varied only in duration. The findings indicate that when vowel duration increased, a concomitant change in vowel perception occurred; /ɛ/ was identified as /æ/ and /ʌ/ was identified as /ɑ/ and /ɔ/. That is, vowels that were manipulated to be longer or shorter in duration than the original recording were identified less accurately than the vowels with neutral duration. These results are consistent with the current study's. The vowel pairs /ɛ, æ/ and /ʌ, ɑ/ are spectrally similar vowels whose members differ in duration in conversational speech (Crystal & House, 1988). Thus, clear speech vowel confusions within the pairs might be expected because the duration was increased during clear speech vowel production.

4.7. Talker differences

Differences in the children's repetition accuracy for the present study's two talkers were observed. Child listeners had more difficulty repeating clear and conversational speech vowels produced by talker 3 than by talker 1. Talker differences had been predicted based on conversational and clear speech vowel perception research (Ferguson, 2004; Hillenbrand, 1995; Uchanski, 2005) and on the present study's adult identification findings from the stimulus verification section.

Differences in talker identification accuracy for conversational speech vowels (Neel, 2008) and clear speech vowels (Ferguson, 2004) are consistent with vowel perception literature. Neel (2008) reported two groups of AE talkers who produced conversational vowels; one group of talkers produced easily identifiable vowels and the other group produced vowels difficult to identify. Similarly, Ferguson (2004) categorized talkers who produced clear speech vowels into two groups: a group that provided listeners with a large identification benefit and a group that provided a smaller identification benefit. Thus, the talker variability demonstrated in this study was supported by previous conversational and clear speech vowel literature; talkers can be categorized into groups with varying degree of conversational and clear speech benefit.

In the present study, results from the adult identification task for stimulus verification also revealed talker identification accuracy differences. Conversational and clear speech vowels produced by talker 1 were identified more accurately than the vowels produced by talker 3. These results are consistent with the results from the children's repetition task and suggest that adult and child listeners follow similar trends for identifying vowels in clear and conversational speech.

4.8. Talker differences for clear speech

As predicted, children's repetition accuracy for each talker's vowels differed significantly in both speaking styles. Despite the differences in clear speech between talkers, an overall clear-speech benefit was found. These findings are consistent with differences found in adults' identification of clear speech vowels produced by different talkers (Ferguson, 2004, 2012). For example, Ferguson (2012) reported that adult listeners identified clear speech vowels produced by females with significant variability, documenting up to a 27.4% difference between talkers; however, adult listeners still showed an improvement in identification of clear speech utterances produced by the less intelligible talkers. Uchanski (2005) describes the need for the inclusion of multiple talkers in clear speech research because some talkers appear to produce a larger clear speech advantage than others. Even though significant differences in talker were reported for the present study, each talker provided listeners with a clear-speech benefit. These results suggest that, although the degree of benefit varies from talker to talker, a clear-speech benefit may not be talker-specific for child listeners.

4.8.1. Comparison of talker differences and acoustic analysis

As shown in Table 1, acoustic analysis revealed that talker 1 and talker 3's clear speech productions differed in utterance and vowel length, fundamental frequency, and formant values. Talker 3 decreased her speaking rate to a far greater extent than talker 1 when producing clear speech vowels. Such a durational change has also been found in clear speech adult literature (Ferguson & Kewley-Port, 2007). It has been proposed that utterance duration may impact speech intelligibility. In a study by Ferguson and Kewley-Port (2007), talkers were divided into two groups: those who provided adult listeners with a large clear speech-benefit and those who provided adult listeners with a limited clear-speech benefit. Comparison of the two groups showed that talkers who provided a limited clear-speech benefit produced significantly longer clear speech vowels than talkers who provided a large clear speech benefit. In the present study, talker 3 provided a limited clear -speech benefit and produced longer clear speech vowels than talker 1. Therefore, the present study's findings are consistent with those of Ferguson and Kewley-Port (2007), in which utterances from talkers who produced significantly longer clear speech vowels were identified less accurately. Results from the present study suggest that the clear speech vowel duration, of these talkers, at least, may have contributed to the clear-speech benefit. When combined with results from adults' clear speech vowel research, results from the current study suggest that adults and children rely on similar durational cues when listening to clear speech vowels. Talkers with the largest clear-speech benefit for adult and child listeners produced shorter clear speech vowels. One possible explanation of why shorter clear speech vowels were identified with more accuracy than longer clear speech vowels is that the listeners were relying on durational cues during speech perception. Previous vowel identification research suggests that in degraded listening conditions, such as the presence of noise, listeners rely more on durational cues than other acoustic cues, such as formants (Winn, Chatterjee, & Idsardi, 2012). If clear speech increases a vowel's duration excessively, then listeners may not be able to rely on duration as a cue. Therefore, talkers who increased duration slightly may have enhanced the vowels' durational cues and increased listeners' ability to identify the vowel. In contrast, talkers who increased duration greatly may have altered the vowels' duration to such a large extent that duration was no longer a reliable cue and the vowel may have became increasingly more difficult to identify.

Analysis of acoustic measures indicated that differences between clear speech vs. conversational speech F_0 were larger for talker 1 than for talker 3. Child listeners repeated clear speech vowels produced by the talker with the greater clear-vs.-conversational speech F_0 difference (talker 1) significantly more accurately than clear speech vowels produced by the other talker. Bradlow et al. (2003) reported similar findings, in that children identified more clear speech keywords accurately when listening to a talker with a larger difference between clear speech F_0 and conversational speech F_0 than a talker with a smaller F_0 difference. Other factors may have played a role in this finding, however, than simply the clear vs. conversational F_0 difference. For example, in Bradlow et al.'s study, children repeated keywords in sentences produced by a male and a female talker, whereas in the present study, children repeated nonsense words in phrases produced by two female talkers.

4.9. Implications

Evidence from this study and others (e.g., Bradlow et al., 2003; Riley & McGregor, 2012) indicates that clear speech is beneficial for child listeners perceiving vowels in noise. Many children spend a large portion of their day listening to speech in classrooms with adverse listening conditions (Larsen & Blair, 2008). The speech in noise literature has documented that children have difficulty perceiving speech in the presence of background noise (Crandell &

Smaldino, 2000; Nishi et al., 2010). Because adults communicate with children in these adverse conditions, it would be beneficial to modify their speech so that children can perceive their messages more accurately. The results of this study and others (e.g., Bradlow et al., 2003) suggest that clear speech may be an effective option for enhancing adult-to-child communication in noisy environments. As more information becomes available about children's perception of clear speech, the benefits of clear speech for child listeners will be better understood.

One factor in children's perception of clear speech vowels that was explored in the present study was talker variability. Results suggest children can benefit from the use of clear speech in noisy environments when perceiving vowels produced by more than one talker. Furthermore, talkers were simply instructed to "Speak as if talking with someone with a hearing loss." With these very concise instructions, talkers in the present study were able to modify their speech signal to attain better intelligibility. These promising results suggest that simple instructions, such as the directions used in this study, are warranted and that clear speech production training may not be needed for adult talkers to produce a clear-speech benefit for child listeners. Thus, the incorporation of clear speech to enhance a speech signal requires a minimal amount of time and may therefore be cost-effective. Acoustic analysis of the stimuli used in the study revealed that duration of vowels in clear speech could potentially affect intelligibility. However, if talkers modify their speech signal excessively and increase the utterance or vowel's duration excessively, child listeners may not receive as great a clear-speech benefit as they would have from a shorter utterance or vowel.

Results from the present study add to an understanding of the impact of clear speech on children's vowel perception in noise. Similarly to adults' identification of clear speech vowels in noise (Ferguson & Kewley-Port, 2002; Rogers et al., 2010), school-age children's vowel repetition in noise is aided by clear speech.

A clear-speech benefit was consistent throughout this study. Regardless of talker or specific vowel, child listeners repeated clear speech more accurately than conversational speech. A handful of studies (Bradlow et al., 2003; Riley & McGregor, 2012) have provided initial evidence that clear speech is an effective speaking style for enhancing adult-to-child communication in noisy environments. The present study adds to the clear speech literature by documenting a clear speech advantage for children's vowel perception.

4.10. Limitations

Some limitations of the study should be noted. Because of the need to restrict the number of trials for the child listeners, only a subset of AE vowels (/ɛ, æ, ɑ, ʌ/) and only two talkers were included in the study. Results from adult studies that include a large subset of AE vowels may not be comparable with those of the present study, which examined a smaller number of target vowels. Additionally, stimuli produced by more talkers would be more representative of the population at large, who produce clear speech in diverse ways (Ferguson, 2004; Uchanski, 2005). Furthermore, results from the repetition task used in the present study, unlike those from identification tasks, may have been somewhat confounded by the children's production skills. Lastly, child listeners were tested in a sound treated booth, which added the control of noise to the study, but resulted in a less naturalistic listening setting than a classroom, for example.

4.11. Conclusion and future directions

Findings from the present study suggest that clear speech is an effective method for enhancing children's perception of vowels in adverse listening conditions. These promising results suggest the use of clear speech for enhancing adult-to-child communication in AE, particularly in noisy environments. Because clear speech requires simple instructions and is effective when produced by many talkers (Bradlow et al., 2003), the incorporation of clear speech by adults communicating with children in noisy settings is promising. Future studies may begin to investigate further the advantages of clear speech for child listeners by including more vowel pairs and more talkers. In addition, because children's ability to perceive speech in noise changes with age (Bradley & Sato, 2008; Stuart et al., 2006), different age groups of listeners and different SNRs should be incorporated into future studies. Further investigation of the acoustic properties of clear speech vowels that can aid children's perception is necessary in order to better understand the acoustic characteristics that increase vowel identification for child listeners. Lastly, the extension of clear speech vowel perception studies to school-age children with disabilities is warranted. Approximately 2.4 million of school-age children in the United States have some type of learning disability (National Center for Learning Disabilities, 2012) with this percentage typically increasing from year to year. The only study thus far on this topic reported that children with learning disabilities demonstrated a strong clear-speech benefit when listening to sentences (Bradlow et al., 2003). Clear speech may be any easy, cost-effective means for enhancing adult-to-child communication. Research is just in the beginning stages of documenting its benefits for child listeners.

5. Tables

Table 1

Acoustic Analysis of the Conversational and Clear Speech Utterances as Produced by the Two Talkers (talker 1 and talker 3) Included in the Children's Repetition Task for Each Target Vowel

/ɛ/						
	Talker 1			Talker 3		
Acoustic measurement	**Conv.**	**Clear**	**Difference**	**Conv.**	**Clear**	**Difference**
1. T1 Utterance duration (s)	1.304	2.600	1.296	1.168	3.627	2.459
T2 Utterance duration (s)	1.236	2.930	1.694	1.180	3.740	2.560
Average utterance duration (s)	1.270	2.765	1.495	1.174	3.684	2.510
2. T1 Vowel duration (s)	0.085	0.132	0.047	0.096	0.172	0.076
T2 Vowel duration (s)	0.092	0.133	0.041	0.095	0.178	0.083
Average vowel duration (s)	0.089	0.133	0.044	0.096	0.175	0.080
3. T1 Duration ratio (vowel/utterance) (s)	0.065	0.051		0.082	0.047	
T2 Duration ratio (vowel/utterance) (s)	0.074	0.045		0.081	0.048	
Average duration ratio (vowel/utterance) (s)	0.070	0.048		0.081	0.048	
4. T1 F0 mean (Hz)	182	215	33	281	293	12
T2 F0 mean (Hz)	190	221	31	271	303	31
Average F0 mean (Hz)	186	218	32	276	298	22
5. T1 F0 at 50% point (Hz)	181	286	105	213	276	63
T2 F0 at 50% point (Hz)	187	266	79	219	296	77
Average F0 at 50% point (Hz)	184	276	92	216	286	70
6. T1 F1 value at 50% point (Hz)	715	855	140	739	862	123
T2 F1 value at 50% point (Hz)	758	803	45	809	903	94
Average F1 value at 50% point (Hz)	737	829	93	774	883	109
7. T2 F2 value at 50% point (Hz)	2116	2205	89	2094	2228	134
T2 F2 value at 50% point (Hz)	2118	2303	185	2036	2257	221
Average F2 value at 50% point (Hz)	2117	2254	137	2065	2243	178

Note. T1 = token 1; T2= token 2

/æ/						
	Talker 1			Talker 3		
Acoustic measurement	**Conv.**	**Clear**	**Difference**	**Conv.**	**Clear**	**Difference**
1. T1 Utterance duration (s)	1.275	2.620	1.345	1.197	3.490	2.293
T2 Utterance duration (s)	1.279	2.750	1.471	1.230	3.180	1.950
Average utterance duration (s)	1.277	2.685	1.408	1.214	3.335	2.122
2. T1 Vowel duration (s)	0.135	0.191	0.056	0.139	0.210	0.071
T2 Vowel duration (s)	0.134	0.228	0.094	0.144	0.260	0.116
Average vowel duration (s)	0.135	0.210	0.075	0.142	0.235	0.094
3. T1 Duration ratio (vowel/utterance) (s)	0.106	0.073		0.116	0.060	
T2 Duration ratio (vowel/utterance) (s)	0.105	0.083		0.117	0.082	
Average duration ratio (vowel/utterance) (s)	0.105	0.078		0.117	0.071	
4. T1 F0 mean (Hz)	171	259	88	201	234	32
T2 F0 mean (Hz)	168	282	114	206	234	28
Average F0 mean (Hz)	169	270	101	204	234	30
5. T1 F0 at 50% point (Hz)	180	245	65	196	228	32
T2 F0 at 50% point (Hz)	186	274	88	200	295	95
Average F0 at 50% point (Hz)	183	260	77	198	262	64
6. T1 F1 value at 50% point (Hz)	1021	1083	62	967	1126	159
T2 F1 value at 50% point (Hz)	994	1101	107	963	1007	44
Average F1 value at 50% point (Hz)	1008	1092	85	965	1067	102
7. T2 F2 value at 50% point (Hz)	1751	2244	493	1827	2225	398
T2 F2 value at 50% point (Hz)	1815	2193	378	1936	2159	223
Average F2 value at 50% point (Hz)	1783	2219	436	1882	2192	311

Note. T1 = token 1; T2= token 2

/ɑ/						
		Talker 1			**Talker 3**	
Acoustic measurement	**Conv.**	**Clear**	**Difference**	**Conv.**	**Clear**	**Difference**
1. T1 Utterance duration (s)	1.210	1.730	0.520	1.300	3.660	2.360
T2 Utterance duration (s)	1.210	1.680	0.470	1.220	3.840	2.620
Average utterance duration (s)	1.210	1.705	0.495	1.260	3.750	2.490
2. T1 Vowel duration (s)	0.121	0.176	0.055	0.143	0.225	0.082
T2 Vowel duration (s)	0.132	0.178	0.046	0.140	0.206	0.066
Average vowel duration (s)	0.127	0.177	0.051	0.142	0.216	0.074
3. T1 Duration ratio (vowel/utterance) (s)	0.100	0.102		0.110	0.061	
T2 Duration ratio (vowel/utterance) (s)	0.109	0.106		0.115	0.054	
Average duration ratio (vowel/utterance) (s)	0.105	0.104		0.113	0.058	
4. T1 F0 mean (Hz)	179	272	93	199	224	25
T2 F0 mean (Hz)	177	261	84	195	226	31
Average F0 mean (Hz)	178	267	89	197	225	28
5. T1 F0 at 50% point (Hz)	174	254	80	204	219	15
T2 F0 at 50% point (Hz)	170	247	77	201	218	17
Average F0 at 50% point (Hz)	172	251	79	203	219	16
6. T1 F1 value at 50% point (Hz)	853	924	71	920	1085	165
T2 F1 value at 50% point (Hz)	936	969	33	987	1100	113
Average F1 value at 50% point (Hz)	895	947	52	954	1093	139
7. T2 F2 value at 50% point (Hz)	1373	1230	-143	1492	1471	-21
T2 F2 value at 50% point (Hz)	1358	1225	-133	1538	1501	-37
Average F2 value at 50% point (Hz)	1366	1228	-138	1515	1486	-29

Note. T1 = token 1; T2= token 2

/ʌ/						
		Talker 1			**Talker 3**	
Acoustic measurement	**Conv.**	**Clear**	**Difference**	**Conv.**	**Clear**	**Difference**
1. T1 Utterance duration (s)	1.180	1.670	0.490	1.270	3.530	2.260
T2 Utterance duration (s)	1.260	1.660	0.400	1.240	3.470	2.230
Average utterance duration (s)	1.220	1.665	0.445	1.255	3.500	2.245
2. T1 Vowel duration (s)	0.090	0.127	0.037	0.085	0.152	0.067
T2 Vowel duration (s)	0.094	0.131	0.037	0.082	0.141	0.059
Average vowel duration (s)	0.092	0.129	0.037	0.084	0.147	0.063
3. T1 Duration ratio (vowel/utterance) (s)	0.076	0.076		0.067	0.043	
T2 Duration ratio (vowel/utterance) (s)	0.075	0.079		0.066	0.041	
Average duration ratio (vowel/utterance) (s)	0.075	0.077		0.067	0.042	
4. T1 F0 mean (Hz)	194	278	84	212	241	29
T2 F0 mean (Hz)	201	261	60	213	237	24
Average F0 mean (Hz)	198	270	72	213	239	26
5. T1 F0 at 50% point (Hz)	188	210	22	255	296	41
T2 F0 at 50% point (Hz)	199	209	10	257	271	14
Average F0 at 50% point (Hz)	194	210	16	256	284	28
6. T1 F1 value at 50% point (Hz)	703	782	79	782	816	34
T2 F1 value at 50% point (Hz)	805	846	41	857	895	38
Average F1 value at 50% point (Hz)	754	814	60	820	856	36
7. T2 F2 value at 50% point (Hz)	1447	1354	-93	1563	1665	102
T2 F2 value at 50% point (Hz)	1497	1447	-50	1517	1576	59
Average F2 value at 50% point (Hz)	1472	1401	-72	1540	1621	81

Note. T1 = token 1; T2= token 2

Table 2

*Confusion Matrix of Children's **Clear and Conversational** (Combined) Vowel Repetition Responses as Percentages of the Total for Each Vowel Presented*

Vowel stimuli are listed in the first column and vowel responses are listed in the top row.

	Response										
Stimulus	**/i/**	**/ɪ/**	**/e/**	**/ɛ/**	**/æ/**	**/ɑ/**	**/ʌ/**	**/ɔ/**	**/o/**	**/u/**	**/ʊ/**
/i/ (control)	**100.0**										
/ɛ/	2.5	2.3	0.2	**82.7**	7.5	2.5	2.1	0.2			
/æ/	1.3	1.0		11.5	**79.0**	6.0	1.0		0.2		
/ɑ/	3.1	1.0		4.8	11.3	**63.3**	16.3		0.2		
/ʌ/	1.9	1.3		19.8	4.0	6.0	**66.9**		0.2		

Table 3

Confusion Matrix of Children's ***Conversational*** *Vowel Repetition Responses as Percentages of the Total for Each Vowel Presented*

Vowel stimuli are listed in the first column and vowel responses are listed in the top row.

Response

Stimulus	**/i/**	**/ɪ/**	**/e/**	**/ɛ/**	**/æ/**	**/ɑ/**	**/ʌ/**	**/ɔ/**	**/o/**	**/u/**	**/ʊ/**
/i/ (control)	**100.0**										
/ɛ/	4.6	3.8	0.4	**73.3**	11.7	3.3	2.5	0.4			
/æ/	2.5	1.3		16.7	**69.2**	8.8	1.7				
/ɑ/	5.0	2.1		7.1	10.8	**48.3**	26.7				
/ʌ/	3.8	2.1		34.6	5.4	7.1	**46.7**		0.4		

Table 4

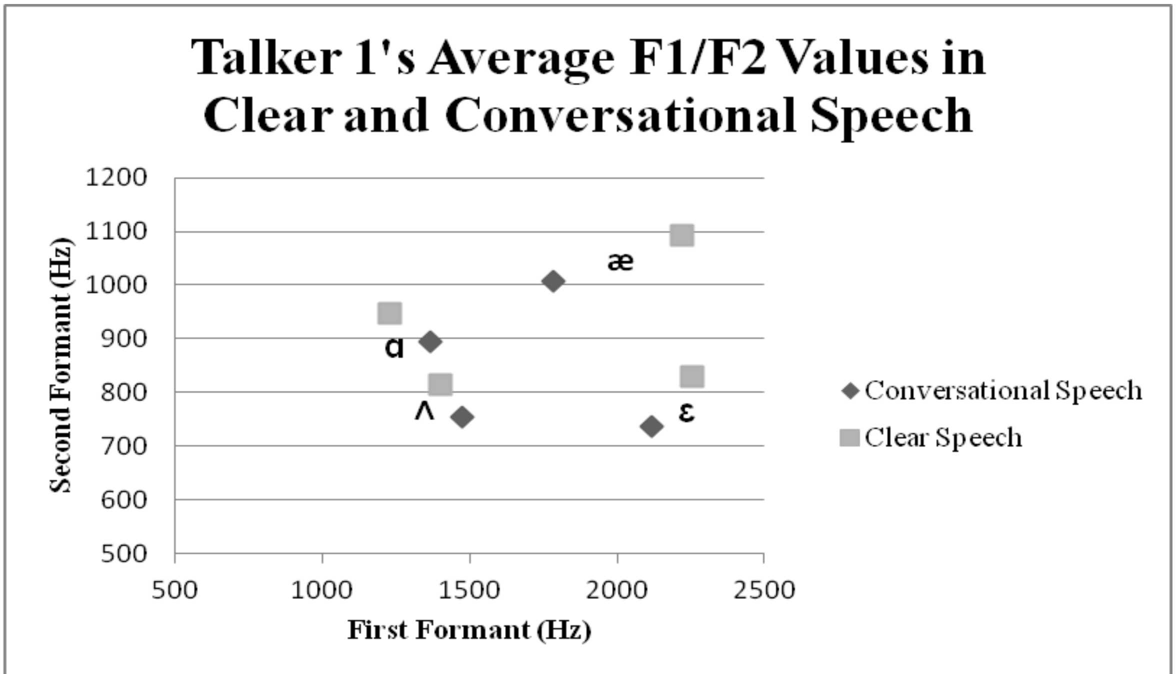
Confusion Matrix of Children's ***Clear*** *Vowel Repetition Responses as Percentages of the Total for Each Vowel Presented*

Vowel stimuli are listed in the first column and vowel responses are listed in the top row.

	Response										
Stimulus	**/i/**	**/ɪ/**	**/e/**	**/ɛ/**	**/æ/**	**/ɑ/**	**/ʌ/**	**/ɔ/**	**/o/**	**/u/**	**/ʊ/**
/i/ (control)	**100.0**										
/ɛ/	0.4	0.8		**92.1**	3.3	1.7	1.7				
/æ/		0.8		6.3	**88.8**	3.3	0.4		0.4		
/ɑ/	1.3			2.5	11.7	**78.3**	5.8		0.4		
/ʌ/		0.4		5.0	2.5	5.0	**87.1**				

6. Figures

Figure 1. Mean F1/F2 values for the two talkers (talker 1 top and talker 3 bottom) used in the children's repetition task demonstrating clear versus conversational speech vowel space differences.



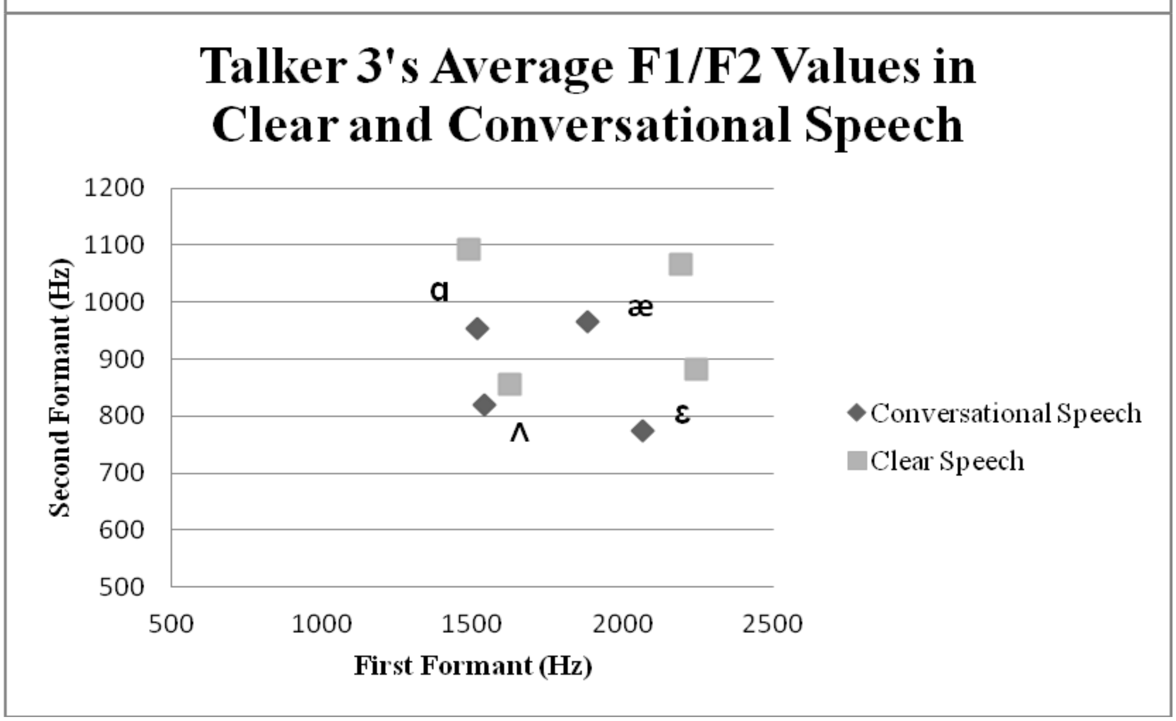


Figure 2. Children's percent correct vowel repetition scores for each speaking style. Error bars represent +/- 1 standard deviation from the mean.

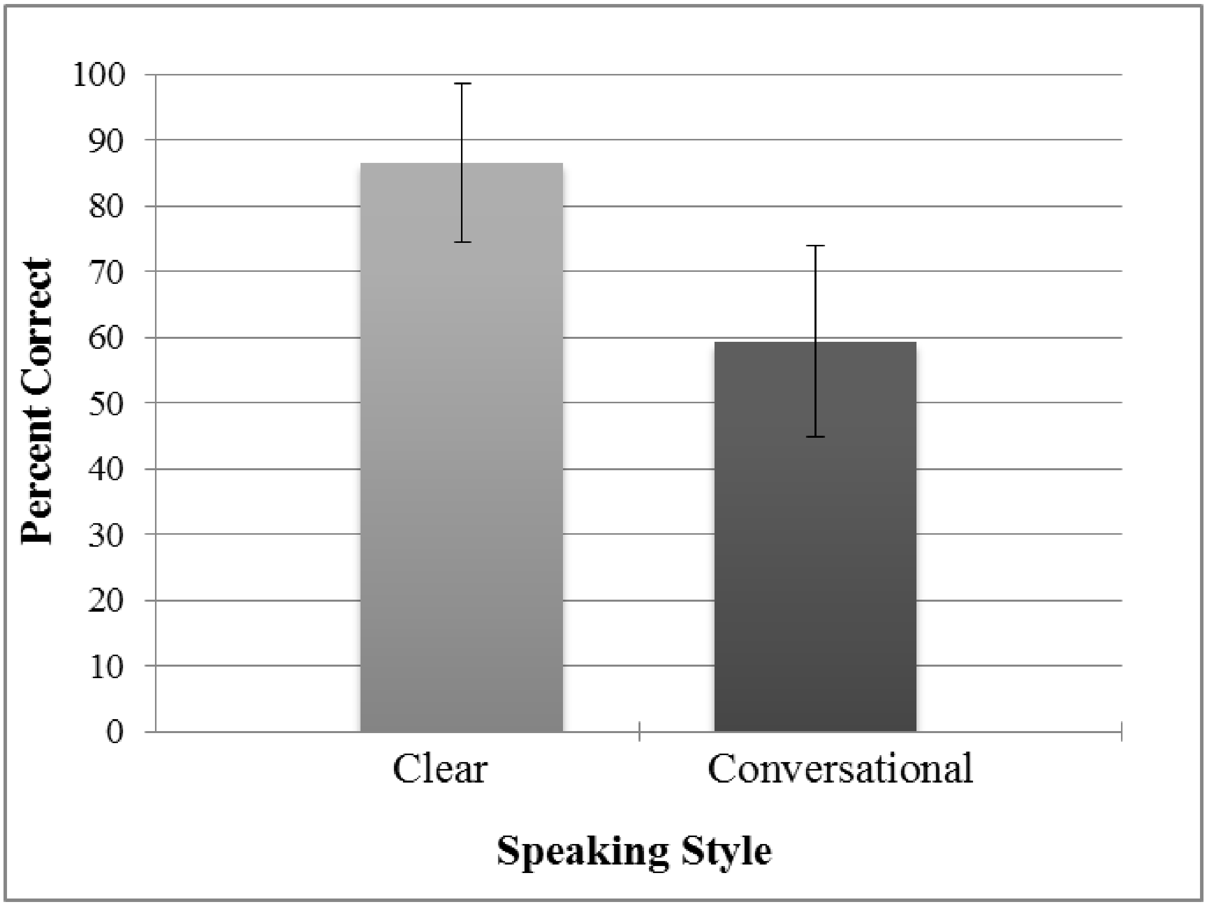


Figure 3. Children's percent correct vowel repetition scores for each target vowel with clear and conversational speech vowel results combined. Error bars represent +/- 1 standard deviation from the mean.

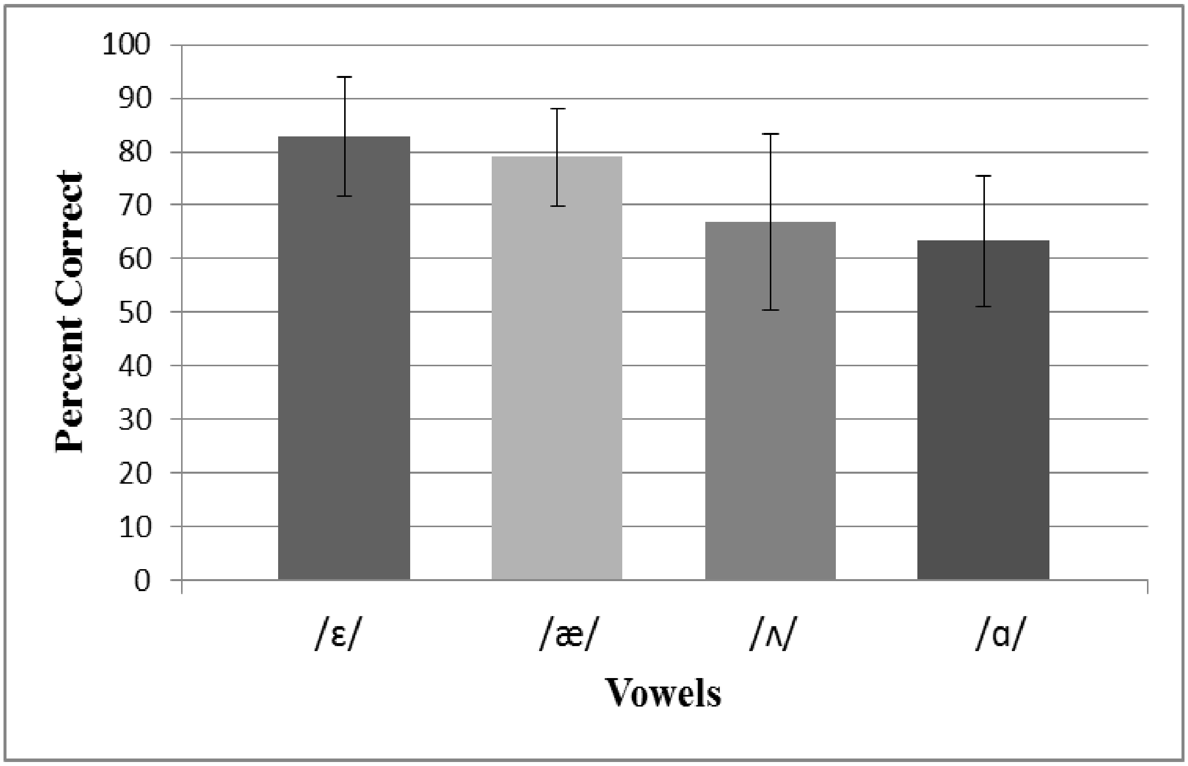


Figure 4. Children's percent correct vowel repetition scores for clear and conversational speech. Error bars represent +/- 1 standard deviation from the mean.

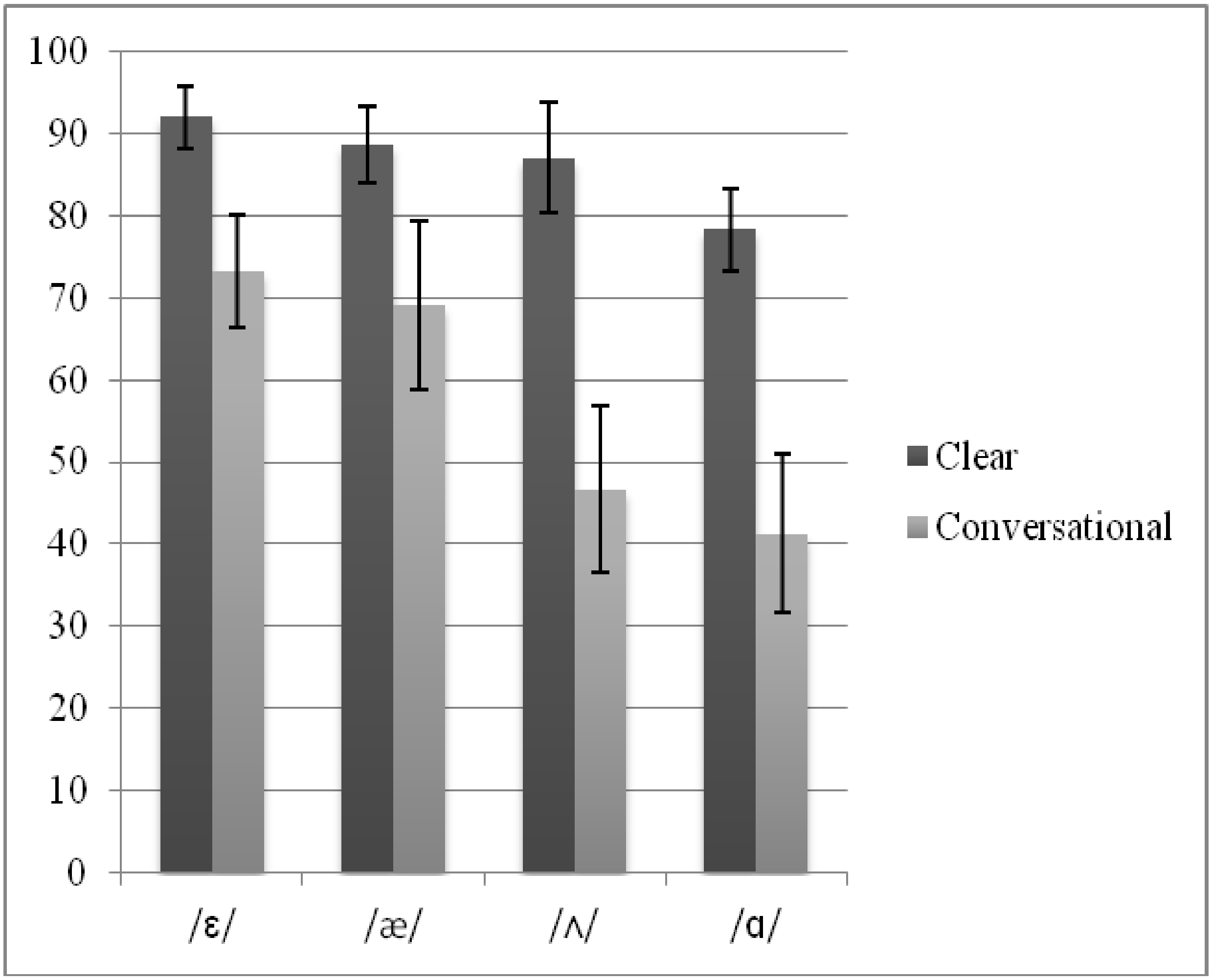
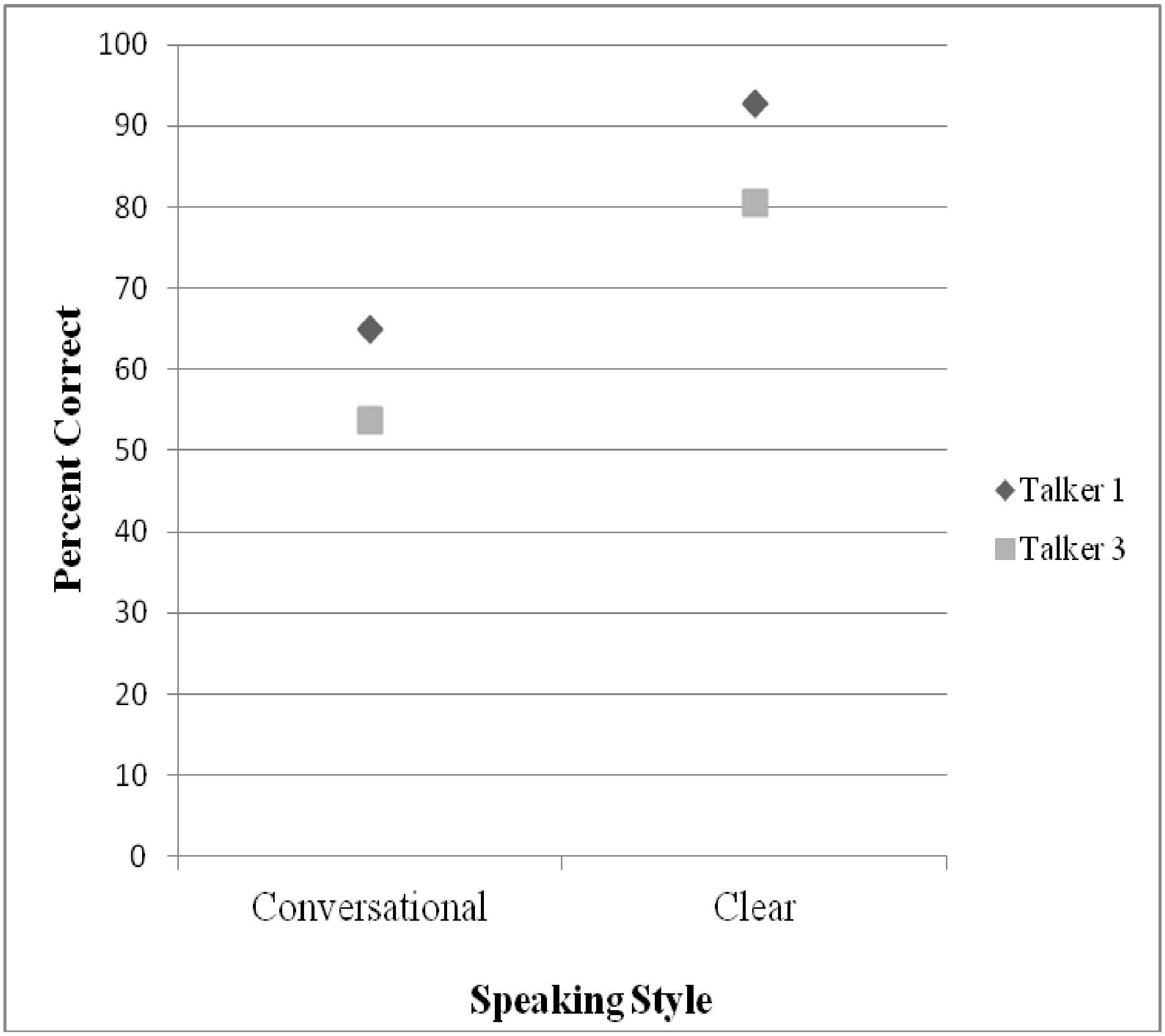


Figure 5. Children's percent correct vowel repetition scores for clear and conversational speech for each talker.



7. References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390-412.

Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *Journal of the Acoustical Society of America, 123*, 2078-2086.

Bradlow, A. R. & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America*, *112*, 272-283.

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

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996) Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20,* 255-272.

Crandell, C. C. & Smaldino, J. J. (2000). Room acoustics for listeners with normal-hearing and hearing impairment. In M. Valente, H. Hosford-Dunn, & R.J. Roeser (Eds.) *Audiology Treatment*. Thieme, New York, Chapter 20, pp. 601-637.

Crystal, T. H. & House, A. S. (1988). The duration of American-English vowels: an overview. *Journal of Phonetics, 16*, 262-284.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2011). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America*, *116,* 3668-3678.

Da Costa, V., Prada, E., Roberts, A., Cohen, S. (2012). Voice disorders in primary school teachers and barriers to care. *Journal of Voice, 26*(1), 69-76.

Dillon, S. (2010, March 10). Many nations passing U.S. in education, expert says. *New York Times*. Retrieved January 24, 2012, from http://www.nytimes.com/2010/03/10/education/10educ.html.

Dinkin, A. J. (2011). Weakening resistance: progress toward the low back merger in New York State. *Language Variation and Change, 23*(3), 315-345.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test*, 3rd ed. Circle Pines, MN: American Guidance Service.

Edwards, J., Fox, R. A., & Rogers, C. L. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and articulatory accuracy. *Journal of Speech, Language, and Hearing Research*, *45*, 231-242.

Ferguson, S. H. (2012). Talker differences in clear and conversational speech: Vowel intelligibility for older adults with hearing losss. *Journal of Speech, Language, and Hearing Research, 55*, 779-790.

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

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*, 1241-1255.

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

Fudala, J. B. (2000). *Arizona Articulation Proficiency Scale, 3rd revision.* Los Angeles, CA: Western Psychological Services.

Gagne, J. P., Masterson, V., Munhall, K. G., Bilida, N., & Querengesser, C. (1994). Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech. *The Journal of the Academy of Rehabilitative Audiology*, *27*, 135-158.

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

Garnier, M., Henrich, N., & Dubois, D. (2010). Influence of sound immersion and communicative interaction on the Lombard effect. *Journal of Speech, Langauge, and Hearing Research*, *53*, 588-608.

Gelfand, S. A., Piper, N., Silman, S. (1986). Consonant recognition in quiet and in noise with aging among normal hearing listeners. *Journal of the Acoustical Society of America, 80*, 1589-1598.

Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics, 12*, 91-114.

Grant, K. W. & Seitz, P. F. (2000). The recognition of isolated words and words in sentences: Individual variability in the use of sentence context. *Journal of the Acoustical Society of America, 107*, 1000-1011.

Helfer, K. & Wilber, L. (1990). Hearing loss, aging, and speech perception in reverberation and in noise. *Journal of Speech and Hearing Research*, *33*, 149-155.

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America,108*(6), 3013-3022.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97,* 3099-3111.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434-446.

Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research*, *43*, 144-157.

Kewley-Port, D., Burkle, T. Z., & Hee Lee, J. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *Journal of the Acoustical Society of America, 122*, 2365-2375.

Kirchhoff, K. & Schimmel, S. (2005). Statistical properties of infant-directed speech versus adult-directed speech: Insights from speech recognition. *Journal of the Acoustical Society of America*, *117*, 2238-2246.

Kuhl, P. K. & Andruski, J. E. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *227*, 684-687.

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

Krause, J. C. & Braida, L. D. (2009). Evaluating the role of spectral and envelope characteristics in the intelligibility role of clear speech. *Journal of the Acoustical Society of America, 125*, 3346-3357.

Larsen, J. B. & Blair, J. C. (2008). The effect of classroom amplification on the signal-to-noise ratio in classrooms while class is in session. *Language, Speech, and Hearing Services in Schools*, *39*, 451-460.

Leone, D., Hsu, S., Baigorri, M., Moya-Gale, G., & Levy, E.S. (2011). Children's perception of clear-speech vowels. *Journal of the Acoustical Society of America, 129,* 2625.

Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America, 125***,** 1138-1152.

Levy, E. S. & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, *36*, 141-157.

Liénard, J. S. & Di Benedetto, M.G. (1999). Effect of vocal effort on spectral properties of vowels. *Journal of the Acoustical Society of America, 106*, 411-422.

Linden, S., Stekelenburg, J.J., Tuomainen, J., & Vroomen, J. (2007). Lexical effects on auditory speech perception: An electrophysiological study. *Neuroscience Letters*, *420*, 49-52.

McKay, K. E., Halperin, J. M. Schwartz, S. T., & Sharma, V. (1994). Developmental analysis of three aspects of information processing: Sustained attention, selective attention and response organization. *Developmental Neuropsychology*, *10*(2), 121–132

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology, 41*, 329-335.

Nabelek, A. K. & Mason, D. (1981). Effect of noise and reverberation on binaural and monaural word identification by subjects with various audiograms. *Journal of Speech and Hearing Research*, *24*, 375-383.

National Center for Education Statistics. (2009). Program for international student assessment. Retrieved January 24, 2012, from the Institute for Education Sciences website: http://nces.ed.gov/surveys/pisa/

National Center for Learning Disabilities. (2012). What are learning disabilities? Retrieved October 14, 2012 from http://www.ncld.org/types-learning-disabilities/what-is-ld/what-are-learning-disabilities?start=1

Neel, A., (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research, 51*, 574-585.

Nelson, P. B. & Soli, S. (2000). Acoustical barriers to learning: Children at risk in every classroom. *Language, Speech, and Hearing Services in Schools*, *31*, 356-361.

Neuman, A. & Hochberg, I. (1983). Children's perception of speech in reverberation. *Journal of the Acoustical Society of America*, *73*, 2145-2149.

Nishi, K., Lewis, D. E., Hoover, B. M., Choi, S., & Stelmachowicz, P. G. (2010). Children's recognition of American English consonants in noise. *Journal of the Acoustical Society of America*, *127*, 3177-3188.

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

Percy, V. & Kei, J. (2008). Effect of Noise on the Comprehension of Everyday Speech by Normally Hearing Adults. *Asia Pacific Journal of Speech, Language, & Hearing*, *11*, 125-132.

Peterson, G. E. & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175-184.

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

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

R Development Core Team (2010). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, Retrieved June 15, 2012, from http://www.R-project.org/.

Riley, K. G. & McGregor, K. K. (2012). Noise hampers children's expressive word learning. *Language, Speech, and Hearing in Schools, 43*, 325-337.

Rogers, C. L., DeMasi, T. M., & Krause, J. C. (2010). Conversational and clear intelligibility of /bVd/ syllables produced by native and non-native English speakers. *Journal of the Acoustical Society of America, 128*, 410-423.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27,* 465-485.

Smiljanic, R. & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass, 10,* 236-264.

Soli, S. D. & Sullivan, J. A. (1997). Factors affecting children's speech communication in classrooms. *Journal of the Acoustical Society of America*, *101*, S3070.

Southwood M. H. & Flege, J. E. (1999). Scaling foreign accent: Direct magnitude estimation versus interval scaling. *Clinical Linguistics and Phonetics*, *13*, 335-349.

Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R., & Pittman, A. I. (2000). The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. *Journal of Speech, Language, and Hearing Research, 43*, 902-914.

Strand, J. & Sommers, M. (2011). Sizing up the competition: Quantifying the influence of the mental lexicon on auditory and spoken word recognition. *Journal of the Acoustical Society of America, 130,* 1663-1672.

Strange, W. (1999). Perception of consonants: From variance to invariance. In J.M. Pickett (Ed.) *The acoustics of speech communications: fundamentals, speech perception theory and technology (pp.* 166-182). Needham Hts., MA: Allyn & Bacon.

Strange, W., Bohn, O. -S., Nishi, K., & Trent, S. A. (2005). Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America*, *118*, 1751-1762.

Strange, W. & Shafer, V. L. 2008. Speech perception in second language learners: The reeducation of selective perception. In J. G. Hansen Edwards & M. L. Zampini (eds.), *Phonology and second language acquisition*, Philadelphia: John Benjamins.

Strange, W., Weber, A., Levy, E. S**.,** Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America, 122*, 1111-1129.

Stuart, A., Givens, G. D., Walker, L. J., & Elangovan, S. (2006). Auditory temporal resolution in normal-hearing preschool children revealed by word recognition in continuous and interrupted noise. *Journal of the Acoustical Society of America*, *119,* 1946-1949.

Tennessee Advisory Commission on Intergovernmental Relations. (2003). Do K-12 school facilities affect education outcomes? Nashville, TN: Young, E., Green, H. A., Roehrich-Patrick, L., Joseph, L., & Gibson, T.

Uchanski, R. M. (2005). Clear speech. In D.B. Pisoni & R.E. Remez (Eds), *The Handbook of Speech Perception* (pp. 207-235). Malden, MA: Blackwell.

Wassink, A. B., Wright, R. A., & Franklin, A. D. (2007). Intraspeaker variability of vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers. *Journal of Phonetics*, *35,* 363-379.

Winn, M. B., Chatterjee, M., & Idsardi, W. J. (2012). The use of acoustic cues for phonetic identification: Effects of spectral degradation and electric hearing. *Journal of the Acoustical Society of America*, *131*, 1465–1479.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science, 12*, 732-745.

8. Appendices

Appendix A. Protocol for Talkers

Condition: _____

Order: ____

1. Five gabeepa this time

8. Five gabuppa this time

4. Five gabeppa this time

5. Five gabappa this time
(hat)

2. Five gabippa this time

3. Five gabaypa this time

11. Five gaboopa this time

10. Five gabUpa this time
(should)

9. Five gaboapa this time
(road)

7. Five gabawpa this time

6. Five gaboppa this time

1. Five gabeepa this time

Appendix B. Language Background Questionnaire for Talkers

Please complete this questionnaire to the best of your knowledge and add any information you feel might be relevant (use the back of the paper if needed).

Talker's Name: ________________
Date: __________________
Date of birth: _______________ Gender: ________ Age: __________
Birthplace: __________________________ ______________________________
Town/City State/Country
What is your highest level of education? __________________
How did you find out about this study? _______________________________________

Places in which you have lived for more than 1 year:
City/State/Country Years
_______________________________ from age _______ to age _______
_______________________________ from age _______ to age _______
_______________________________ from age _______ to age _______
_______________________________ from age _______ to age _______
If you have lived in more places please check here _____ and continue on the back.

Parent 1's Birthplace: ____
Languages Parent 1 spoke fluently: _________________________________

Parent 2's Birthplace: ____
Languages Parent 2 spoke fluently: _________________________________

What languages are spoken in the home or at work? (for example, by parents, a spouse, babysitter, or relatives) Please explain.

____
____

At what age did you first hear each of these languages regularly (please explain, e.g. first heard Spanish when visiting grandparents in Peru):
English:
Spanish:
Other:

Please list the approximate percent of time you currently hear each language. (this should add up to 100%):
English:
Spanish:
Other:
If this has changed over time, please explain:__________________________________

What languages do you speak fluently and understand without effort?

1. _____________________ 2. _______________________ 3. ___________________

What language do you prefer to use?_____________________________________

Are you exposed to anyone who speaks English with a foreign accent (e.g., Spanish accent in English, other accent in English) frequently? _________________________

If yes, please describe what accent (e.g., Spanish, Italian, other [please describe]):

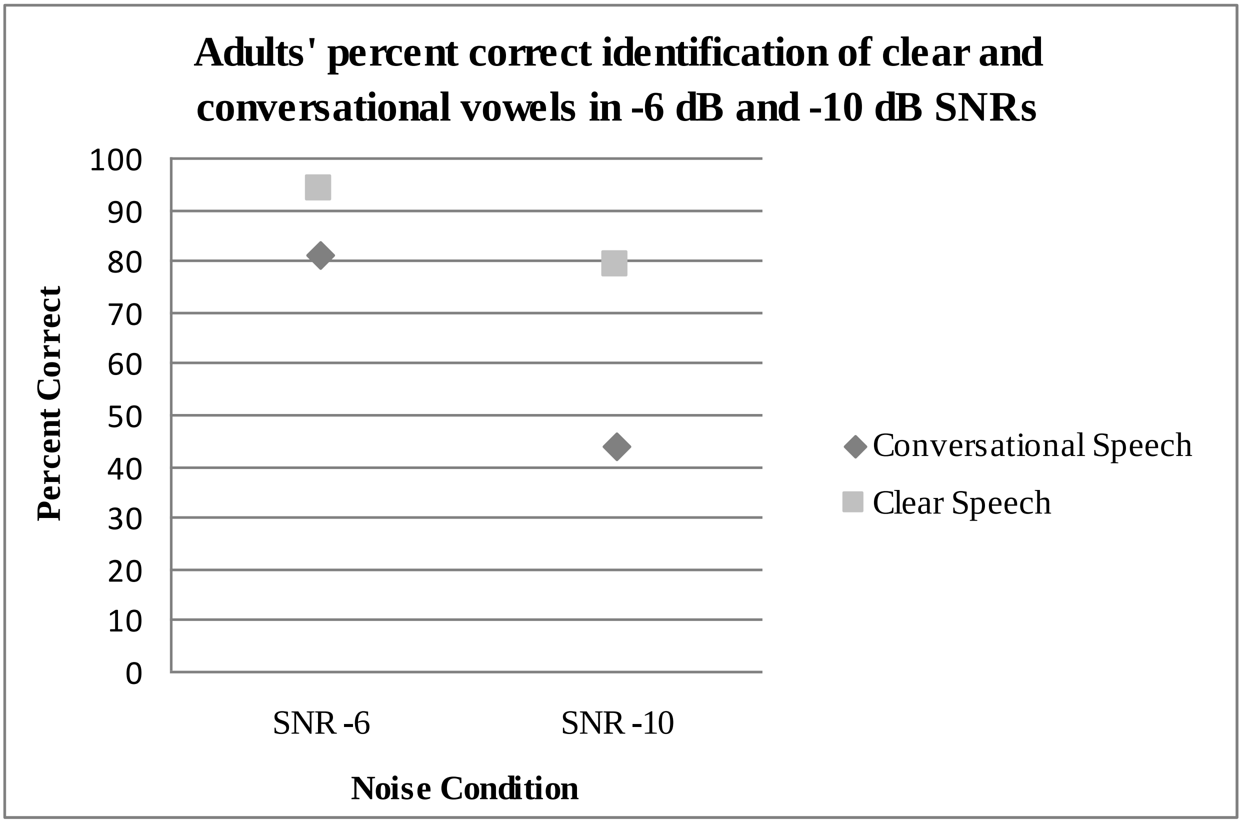
and how often, and in what context do you hear this speech: _______________________________
____

Have you had a recent hearing screening? YES___ NO____
If yes, what were the results?____

Have you ever received speech-language therapy services? YES________ NO________
If yes, when and for how many years? ______________________________
If yes, please describe (e.g., trouble producing the /r/ sounds, has trouble organizing his thoughts,etc…)____
____
____
____
____
____

Please add any comments/concerns regarding your language/speech sound development in any language: ____
____
____

Appendix C. Adults' Percent Correct Identification of Clear and Conversational Vowels in -6 dB and -10 dB SNRs



Appendix D. Design of the Adult Identification Task

Task familiarization:

1 block of /o, u, ɪ/ in gabVpa context in both conversational and clear speech produced by 1 talker, no noise; 2 tokens of each vowel; 2 repetitions of each token

6 vowels (3 conversational vowels + 3 clear vowels) x 2 tokens of each vowel x 2 trials of each token, totaling 24 trials

Inclusion criteria: no more than 1 error (96% correct)

Stimulus familiarization:

1 copy of a test block using each stimulus one time, totaling 80 trials.

Experiment:

1. 4 blocks
2. Each talker: 8 vowels in gabVpa context (4 clear vowels + 4 conversational vowels) x 2 tokens of each vowel x 4 trials of each token = 64 trials for each talker; 8 trials for each vowel in each speaking style for each talker, totaling 256 trials
3. Control: /i/ in 0 dB SNR = 2 vowels (1 clear + 1 conversational) x 2 tokens x 2 trials = 8 trials for each talker, 4 trials for each vowel in each speaking style for each talker, totaling 32 trials
4. Total stimuli: 256 experimental trials (64 trials per talker x 4 talkers) + 32 control trials (8 trials per talker x 4 talkers) = 288 trials

Sequence of blocks:

Task familiarization: 24 trials

Stimulus familiarization: 80 trials

Experimental block 1: 72 trials

Experimental block 2: 72 trials

Experimental block 3: 72 trials

Experimental block 4: 72 trials

Appendix E. Instructions for Adult Identification Task

Familiarization

This experiment is about pronunciation of speech sounds. Let's go over the vowels of American English using nonsense words in the form "gab-vowel-pa." Please read each of the following words aloud to the experimenter:

gabeepa	gabawpa
gabippa	gabuppa
gabaypa	gaboapa (road)
gabeppa	gapUpa
gabappa (hat)	gaboopa
gaboppa	

You will hear someone saying the nonsense words you just practiced in phrases. Please listen to the second vowel sound of the word and determine which American English vowel she is saying. Indicate the vowel by choosing one of the following words:

gabeepa	gabawpa
gabippa	gabuppa
gabaypa	gaboapa
gabeppa	gapUpa
gabappa	gaboopa
gaboppa	

Now that you know the vowels, here is some practice with the task. You will hear the nonsense words you just practiced. Choose the word that contains the second vowel sound in the nonsense word you heard by clicking the left mouse button. For example, if you hear "gabeppa," use the mouse to click on “gabeppa” on the screen. Try to focus only on the pronunciation of the target vowel and ignore any other factors (e.g., recording quality, rate, volume). You will now complete one 24-trial block of phrases.

Whenever you’re ready, press the left mouse button to begin.

Experiment

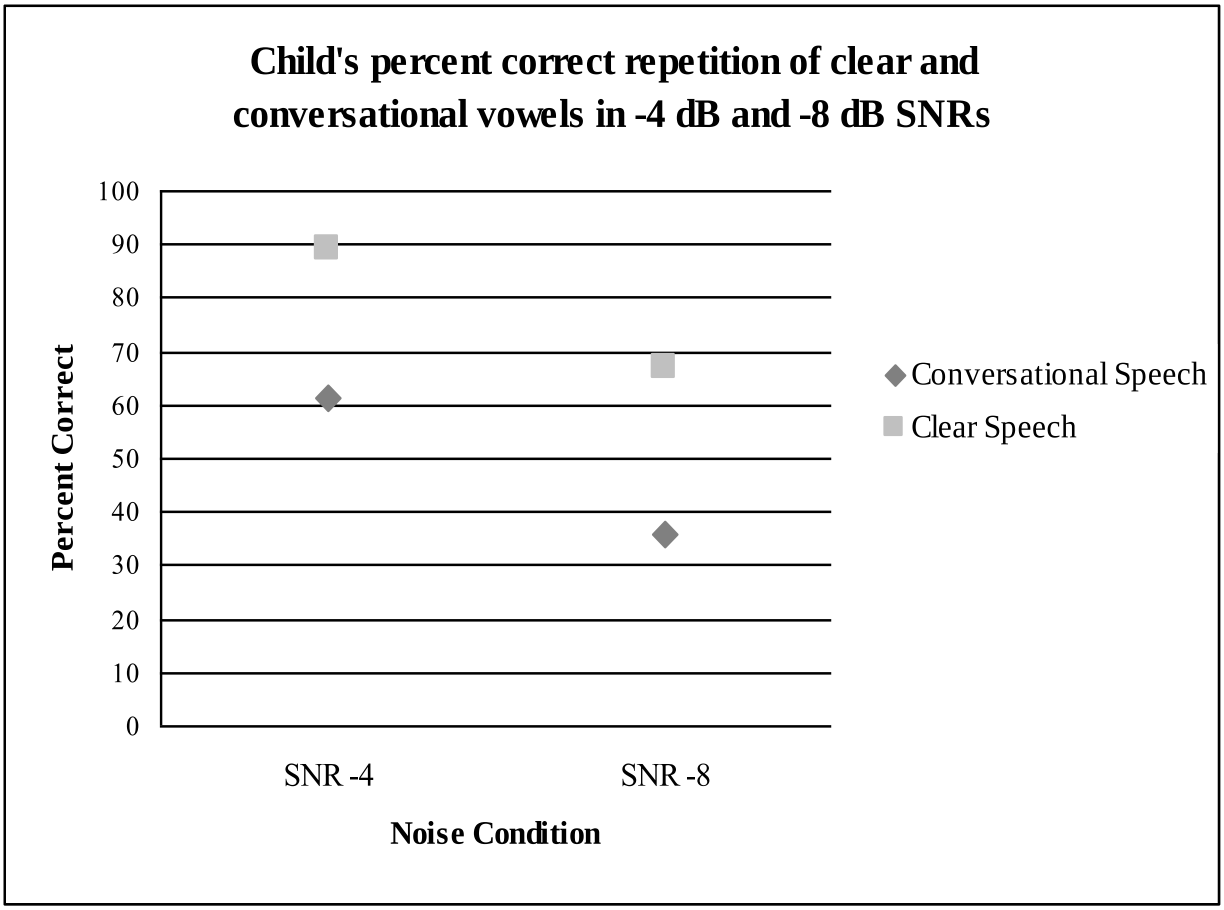
Now that you had some practice, you are going to listen to some more sounds. You will hear people saying nonsense words in the form of “gab-vowel-pa” in phrases. This time some of the phrases will be in noise and may be harder to hear. You will then see the American English nonsense words you know (e.g., gabeepa, gabippa, etc.). Just like you did before, when you hear the nonsense word, listen to the second vowel and choose the word you heard.

After you indicate the word, the same phrase will be presented again and you will see a rating scale from 1-9. The purpose of the scale is for you to indicate how clear an example of that American English vowel it is. If it was clear, choose a point on the scale near "Very clear sounding" (9). If it was unclear, select a point near the "Very unclear sounding" end of the scale (1).

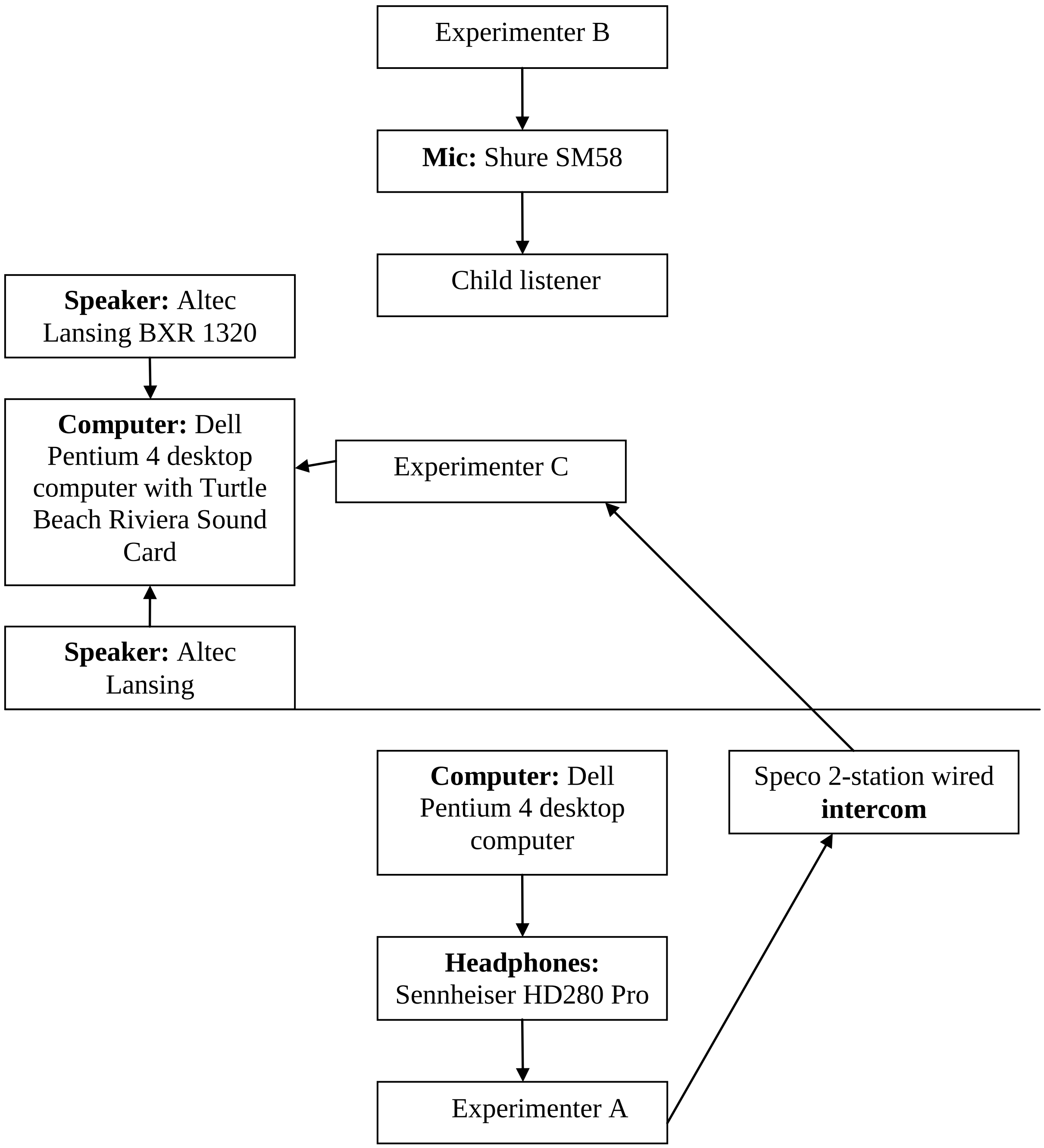
So you'll listen to the second vowel of the nonsense word (“gab-vowel-pa”) in a phrase and choose the word that has the vowel you heard. Then you will listen to it a second time and indicate how typical an example of the vowel it is. Please use the whole spectrum of the scale. You may replay the stimulus only if you did not attend to it the first time. Try to focus only on the pronunciation of the target vowel and ignore any other factors (e.g., recording quality, rate, volume). You will now complete 5 blocks of phrases. The first block has 80 trials and the rest have 72 trials. Do you have any questions?

Whenever you’re ready, press the left mouse button.

Appendix F. Typically-developing 7-year-old Child's Percent Correct Repetition of Clear and Conversational Vowels in -4 dB and -8 dB SNRs (Leone et al., 2011)



Appendix G. Flowchart of Recording Equipment



Appendix H. Instructions for Child Repetition Task

Familiarization

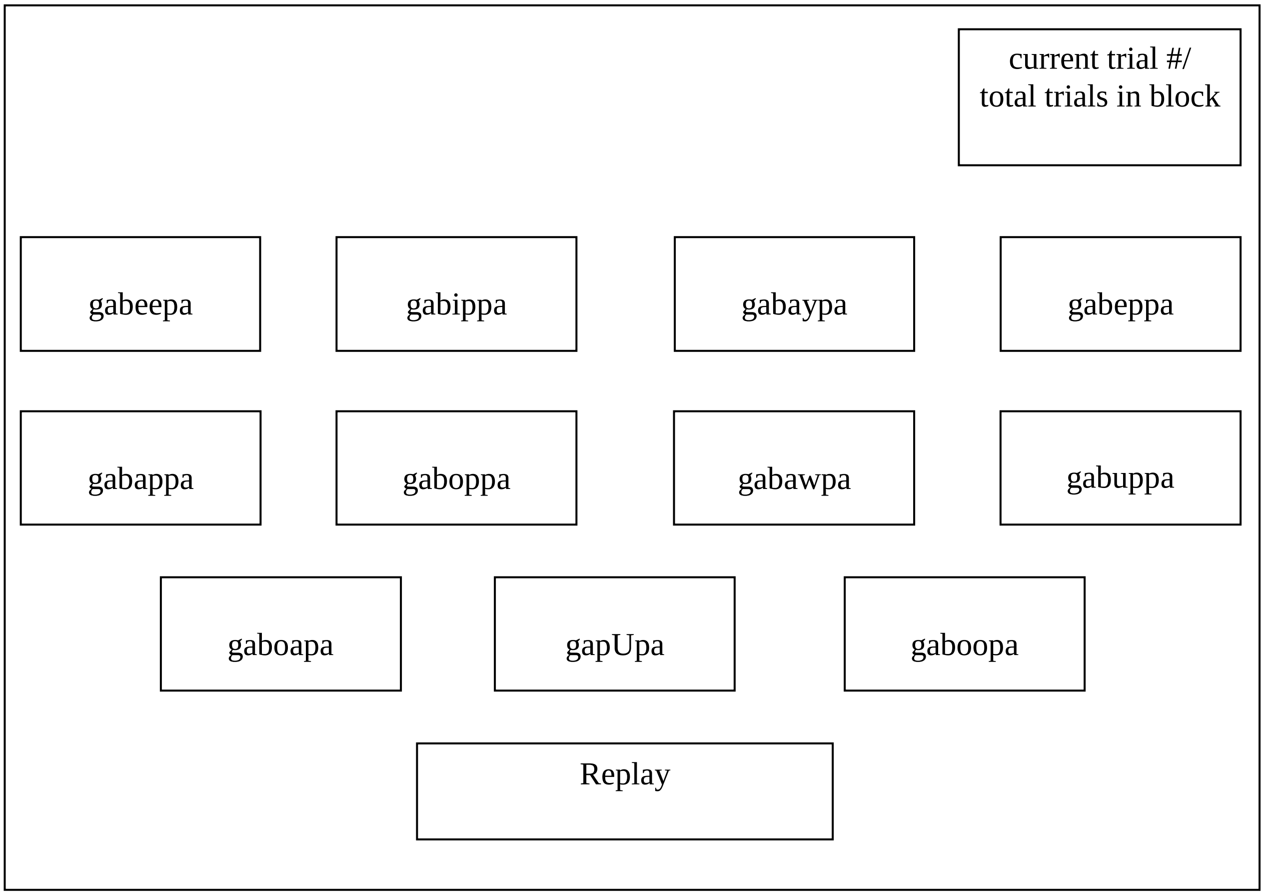
Hi! We're going to listen to some sentences with silly-sounding words. I want you to listen and then say exactly what you heard.

Get ready!

Experiment

We're going to listen to some more sentences with silly-sounding words. This time there may be other noise that we hear. Try your best to listen to the words and then say exactly what you heard. This will be harder, so just do your best.

Appendix I. Diagram of Computer Screen



Appendix J. Design of the Children's Repetition Task

Task familiarization:

1 block of /o, u, ɪ/ in gabVpa context in both conversational and clear speech produced by 1 talker, no noise; 2 tokens of each vowel; 2 repetitions of each token

6 vowels (3 conversational vowels + 3 clear vowels) x 2 tokens of each vowel x 2 trials of each token, totaling 24 trials

Inclusion criteria: no more than 2 errors (92%)

Stimulus familiarization:

1 copy of a test block using each stimulus one time, totaling 36 trials.

Experiment:

1. 6 blocks
2. Experimental -6 dB SNR: Each talker: 8 vowels in gabVpa context (4 clear vowels + 4 conversational vowels) x 2 tokens of each vowel x 4 trials of each token = 64 trials for each talker; 8 trials for each vowel
3. Total experimental stimuli (-6 dB SNR): 64 trials per talker x 2 talkers = 128 experimental trials
4. Control: /i/ in 0 dB SNR = 2 vowels (1 clear + 1 conversational) x 2 tokens x 2 trials = 8 trials for each talker, 4 trials for each vowel in each speaking style for each talker
5. Total control stimuli (0 dB SNR): 8 trials per talker x 2 talkers = 16 control trials
6. Total control stimuli (16) + Total experimental stimuli (128) = 144 trials

Sequence of blocks:

Task familiarization: 24 trials

Stimulus familiarization: 36 trials

Experimental block 1: 36 trials

Experimental block 2: 36 trials

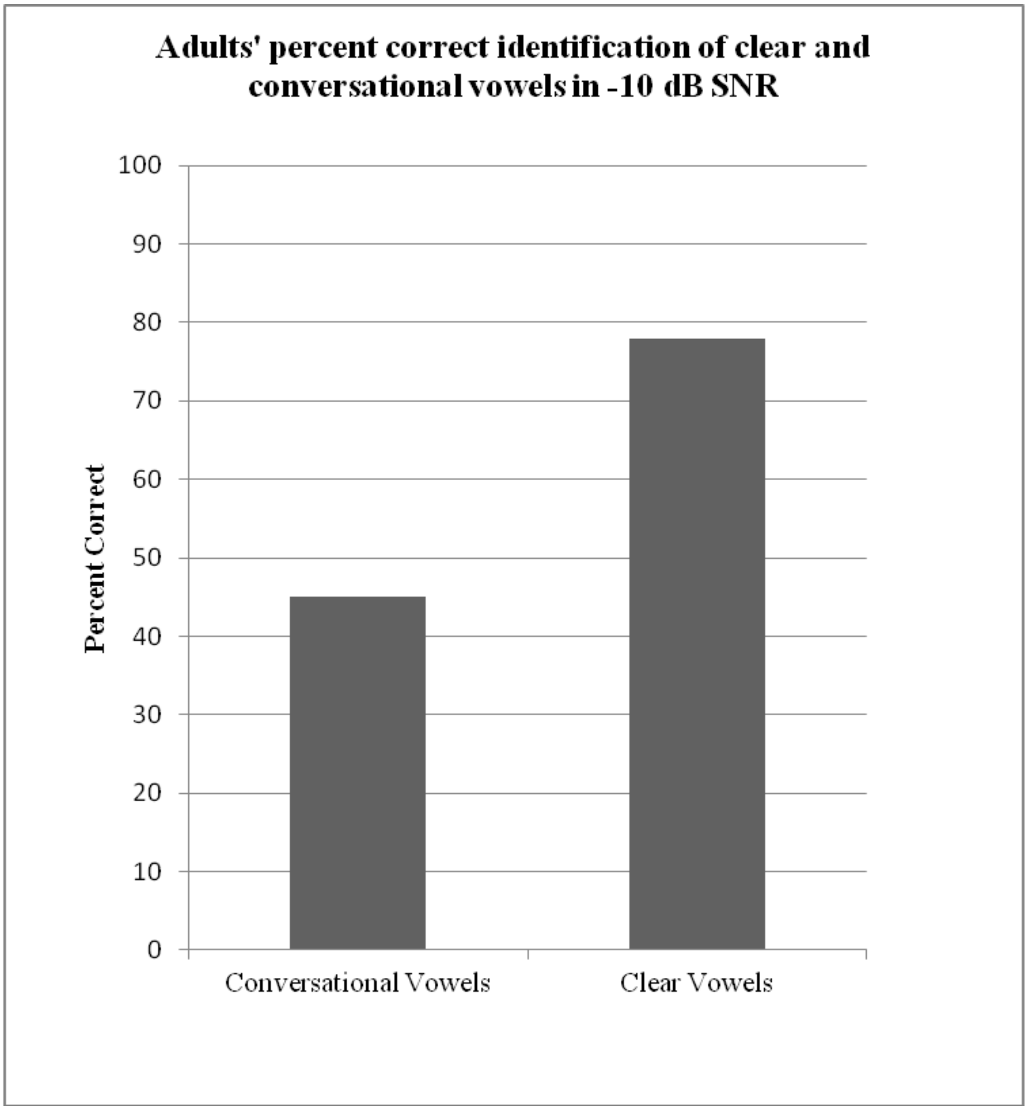
Experimental block 3: 36 trials

Experimental block 4: 36 trials

Appendix K. Adult Participant Characteristics

Participant	Age	Gender	Birthplace	Languages other than English
A1	20	Female	Yonkers, NY	none
A2	20	Female	Flushing, NY	none
A3	22	Female	New Rochelle, NY	none
A4	22	Female	Bronx, NY	Hears Spanish spoken by her grandparents (5% of time)
A5	22	Female	Elizabeth, NJ	Speaks some Spanish for class in school
A6	21	Female	Staten Island, NY	none
A7	21	Female	Staten Island, NY	none
A8	20	Female	New York, NY	Speaks some French for class in school
A9	18	Female	New York, NY	none
A10	20	Female	Brooklyn, NY	none
A11	29	Female	Marlboro, NJ	none

Appendix L. AE Adults' Identification Accuracy of Clear and Conversational Vowels in -10 dB SNR



Appendix M. McNemar's Test for Paired Proportions of Adults' Identification of Conversational and Clear Vowels for All Talkers

Tests of Within Talker Effects

	Proportion Clear	Proportion Conv.	Proportion Difference	χ^2	*df*	*p-value*	odds ratio[a]
Talker 1	.83	.446	0.384	105.35	1	0.0000	8.105
Talker 2	.847	.631	0.216	41.86	1	0.0000	3.452
Talker 3	.713	.332	0.381	97.59	1	0.0000	6.36
Talker 4	.727	.395	0.332	75.63	1	0.0000	4.656

Note. Given the number of comparisons is greater than 2, the alpha level need to be adjusted: alpha value after Bonferroni multiple-comparison correction: 0.05 / 4 = 0.0125

[a]Odds ratio (*OR*) is a measure of effect size. *OR*'s interpretation: a correct answer is "odds ratio" times more likely in clear speech than in conversational speech.

Appendix N. IRB Approval Letter

TEACHERS COLLEGE
COLUMBIA UNIVERSITY
OFFICE OF SPONSORED PROGRAMS

Institutional Review Board

February 1, 2012

Dorothy Leone
Teachers College
525 West 120th Street, Box 180
New York, NY 10027

Dear Dorothy,

Please be informed that as of the date of this letter, the Institutional Review Board for the Protection of Human Subjects in Research (IRB) at Teachers College, Columbia University has reviewed your *continuing* study entitled "*Children's Perception of Conversational and Clear American-English Vowels in Noise*" under **Expedited Review (Category 4).**

I am pleased to let you know that your *continuing* study has been fully approved.

The approval is effective until **January 31, 2013.**

The IRB Committee must be contacted if there are any changes to the protocol during this period. **Please note**: If you are planning to continue your study, a Continuing Review application must be filed six weeks prior to the expiration of the protocol. The IRB number assigned to your protocol is **11-155CR**. Feel free to contact the IRB Office [212-678-4105 or mbrooks@tc.edu] if you have any questions.

Please note that your consent form bears an official IRB authorization stamp. Copies of this form with the IRB stamp must be used for your research work.

Best wishes for your data collection.

Sincerely,

Karen Froud, Ph.D.
Associate Professor of Speech and Language Pathology
Chair, IRB

cc: File, OSP

Institutional Review Board, Office of Sponsored Programs, Box 151
525 West 120th Street, New York NY 10027 Tel: 212 678 4105 Fax: 212 678 8110

Appendix O. Language Background Questionnaire for Parents

Please complete this questionnaire to the best of your knowledge and add any information you feel might be relevant (use the back of the paper if needed).

Participant number: _______________
Date: ____________________Parent's e-mail address: ______________________________
Address: ____
Telephone Numbers: (Cell) _______________ (Work) _________________________
Child's date of birth: _________________ Gender: __________ Child's age: ____________
Birthplace: ____
Town/City State/Country
What school level is your child currently in? _____________________________________
How did you find out about this study? ____

Places in which your child has lived for more than 1 year:

City/State/Country	Years
__________________________________	from age ________ to age ________
__________________________________	from age ________ to age ________
__________________________________	from age ________ to age ________
__________________________________	from age ________ to age ________

If your child has lived in more places please check here _____ and continue on the back.

Parent 1's Birthplace: ____
Languages Parent 1 spoke fluently: ____________________________________

Parent 2's Birthplace: ____
Languages Parent 2 spoke fluently: ____________________________________

What languages are spoken in the home? (for example, by parents, guardians, grandparents, or relatives) Please explain.

____
____

At what age did your child first hear each of these languages regularly (please explain, e.g. first heard Spanish when visiting grandparents in Peru):
<u>English:</u>
<u>Spanish:</u>
<u>Other:</u>

Please list the approximate percent of time your child currently hears each language. (this should add up to 100%):
English:

Spanish:
Other:
If this has changed over time, please explain:____________________________________
____
____

What languages does your child speak fluently and understand without effort?

2. _____________________ 2. _______________________ 3. ___________________

What language does your child prefer to use?____________________________________

Has your child had a recent hearing screening? YES___ NO____
If yes, what were the results?____

Has your child received speech-language therapy services? YES________ NO________
If yes, for how many years? __________________________
If yes, please describe (e.g., trouble producing the /r/ sounds, has trouble organizing his thoughts,etc…)____
____
____
____
____

Please add any comments/concerns regarding your child's language/speech sound development in any language: ____
____
____

Appendix P. Child Listener Characteristics

Listener	**Age**	**Gender**	**Birthplace**	**Parents' Language(s)**
C1	7.8	Female	White Plains, NY	English
C2	6.11	Female	Yonkers, NY	English
C3	8.1	Female	Yonkers, NY	English
C4	5.0	Female	Staten Island, NY	English
C5	7.5	Male	Staten Island, NY	English
C6	7.4	Female	Staten Island, NY	English
C7	8.4	Male	Staten Island, NY	English
C8	5.5	Male	Brooklyn, NY	English
C9	8.5	Male	New York, NY	English
C10	6.7	Male	White Plains, NY	English
C11	8.2	Male	White Plains, NY	English
C12	6.1	Female	Tuckahoe, NY	English
C13	7.9	Male	Tuckahoe, NY	English
C14	5.4	Male	Irvington, NY	English
C15	6.0	Female	Irvington, NY	English

Appendix Q. Information about Mixed Effects Models

A logistic model was chosen for the present study because the dependent variable (i.e. correct or incorrect repetition accuracy) is binary; listeners either repeated the vowel accurately or inaccurately. A logistic model cannot analyze collapsed data because the model relies on the assumption that each trial is independent. Because the data required individual analysis, a mixed effects design was chosen. Mixed effects modeling, as opposed to linear regression, analyzes data as a whole and does not collapse the data into averages for each listener. Furthermore, mixed effects modeling builds a statistical model step by step where at each step an effect is tested. If the finding is statistically significant, the effect is included in the model, but if the finding is not significant, the effect is dropped from the model. The goal of mixed effects modeling is to find the model that best fits the data.

Appendix R. Mixed Effects Model Steps

<table>
<tr><th rowspan="2">Model</th><th rowspan="2">Fixed effects (random effects)</th><th rowspan="2">df</th><th rowspan="2">AIC</th><th rowspan="2">BIC</th><th rowspan="2">logLik</th><th colspan="4">LR test</th></tr>
<tr><th>Comparison</th><th>X^2</th><th>df</th><th>p</th></tr>
<tr><td>m0</td><td>(1|ids)</td><td>2</td><td>2142.4</td><td>2153.6</td><td>-1069.21</td><td></td><td></td><td></td><td></td></tr>
<tr><td>m1</td><td>(1|ids),(1|item)</td><td>3</td><td>1877.9</td><td>1894.6</td><td>-935.96</td><td>m0-m1</td><td>266.51</td><td>1</td><td><.001</td></tr>
<tr><td>m2</td><td>clear,(1|ids),(1|item)</td><td>4</td><td>1815.6</td><td>1837.9</td><td>-903.81</td><td>m1-m2</td><td>64.302</td><td>1</td><td><.001</td></tr>
<tr><td>m3a</td><td>clear,(1|ids),(clear|ids),(1|item)</td><td>7</td><td>1815.5</td><td>1854.5</td><td>-900.76</td><td>m2-m3a</td><td>6.0929</td><td>3</td><td>.107</td></tr>
<tr><td>m3b</td><td>clear,(1,clear|ids),(1|item)</td><td>6</td><td>1813.5</td><td>1846.9</td><td>-900.76</td><td>m2-m3b</td><td>6.0929</td><td>2</td><td>.048</td></tr>
<tr><td>m4</td><td>clear,talker2,(1, clear|ids),(1|item)</td><td>7</td><td>1788.8</td><td>1827.7</td><td>-887.38</td><td>m3b-m4</td><td>26.757</td><td>1</td><td><.001</td></tr>
<tr><td>m4c</td><td>clear,talker2,(1, clear|ids),(talker2|ids),(1|item)</td><td>10</td><td>1794.6</td><td>1850.2</td><td>-887.29</td><td>m4-m4c</td><td>0.1796</td><td>3</td><td>.981</td></tr>
<tr><td>m4d</td><td>clear,talker2,(1, clear,talker2|ids),(1|item)</td><td>10</td><td>1792.9</td><td>1848.5</td><td>-886.45</td><td>m4-m4d</td><td>1.861</td><td>3</td><td>.602</td></tr>
<tr><td>m6</td><td>clear,talker2,vowel,(1,clear|ids), (1|item)</td><td>10</td><td>1764.6</td><td>1820.2</td><td>-872.29</td><td>m4-m6</td><td>30.181</td><td>3</td><td><.001</td></tr>
<tr><td>m6b</td><td>clear,talker2,vowel,(1,clear|ids), (vowel|ids),(1|item)</td><td>20</td><td>1582.7</td><td>1693.9</td><td>-771.36</td><td>m6-m6b</td><td>201.86</td><td>10</td><td><.001</td></tr>
<tr><td>m6c</td><td>clear,talker2,vowel,(1,clear,vowel|ids),(1|item)</td><td>22</td><td>1581.5</td><td>1703.9</td><td>-768.77</td><td>m6b-m6c</td><td>5.183</td><td>2</td><td>.075</td></tr>
<tr><td>m7</td><td>clear,talker2,vowel,clear*talker2,(1,clear|ids),(vowel|ids),(1|item)</td><td>21</td><td>1580.3</td><td>1697</td><td>-769.14</td><td>m6b-m7</td><td>4.4517</td><td>1</td><td>.035</td></tr>
</table>

m7s1	clear,talker2,vowel,clear*talker2,(1,clear|ids),(1|item)	11	1762.7	1823.9	-870.37	m7s1-m7	202.46	10	<.001
m7s2	**clear,talker2,vowel,clear*talker2,(1|ids),(vowel|ids),(1|item)**	**19**	**1577.8**	**1683.4**	**-769.88**	**m7s2-m7**	**1.4793**	**2**	**.477**
m8s2	clear,talker2,vowel,clear*talker2,clear*vowel,(1|ids),(vowel|ids),(1|item)	22	1581.3	1703.6	-768.64	m7s2-m8s2	2.4714	3	.481
m8	clear,talker2,vowel,clear*talker2,clear*vowel,(1,clear|ids),(vowel|ids),(1|item)	24	1583.8	1717.3	-767.92	-m8			
m8s1	clear,talker2,vowel,clear*talker2,clear*vowel,(1,clear|ids),(1|item)	14	1764.4	1842.2	-868.18	-m8s1			

Note. Final model selected is in bold.

Appendix S. Mixed Effects Logistic Model with Crossed Random Effects Using Speaking Style, Talker, and Vowel as Fixed Effects and Listeners and Trials as Random Effects (model m7s2)

Fixed Effects

	Estimate	Standard Error	z	Significance
Intercept	-0.726	0.561	-1.294	.196
Clear (speaking style)	2.001	0.316	6.342	.001*
Talker	0.912	0.291	3.133	.001*
Pairwise Vowel Comparisons				
/ʌ/ vs. /ɑ/	.334	0.708	0.471	.637
/ɛ/ vs. /ɑ/	1.434	0.472	3.040	.001*
/æ/ vs. /ɑ/	1.727	.810	2.132	.010*
/ɛ/ vs. /ʌ/	1.100	0.545	2.017	.044*
/æ/ vs. /ʌ/	1.393	0.605	2.303	.021*
/æ/ vs. /ɛ/	0.293	0.629	0.466	.641
Clear (speaking style) X Talker	0.988	0.470	2.101	.036*

Random Effects

	Name	Variance	Standard Error
Trial	Intercept	1.048	1.023
Listener	/ɑ/	2.751	1.659
Listener	/ʌ/	0.973	0.987
Listener	/ɛ/	0.571	0.755
Listener	/æ/	3.397	1.843
Listener	Intercept	0.594	0.771

Note. N = 1920 trials; N = 15 listeners

Appendix T. Confusion Matrix of a 7-year-old Child's Clear and Conversational Vowel Repetition Responses as Percentages of the Total Vowel for Each Vowel Presented (Leone et al., 2011)

Stimulus	Response										
	/i/	/ɪ/	**/e/**	/ɛ/	**/æ/**	/ɑ/	/ʌ/	/ɔ/	**/o/**	**/u/**	/ʊ/
/i/	30	0	0	3	0	12	39	0	0	0	0
/ɛ/	0	6	0	46	33	2	19	0	0	0	0
/æ/	0	0	0	3	92	3	3	0	0	0	0
/ɑ/	0	6	0	0	14	64	22	0	0	0	0
/ʌ/	0	0	0	5	5	11	79	0	0	0	0