THE ACOUSTIC AND PERCEPTUAL EFFECTS OF SINGLE-MICROPHONE NOISE REDUCTION IN HEARING AIDS ON MANDARIN FRICATIVES AND AFFRICATES

by

Foong Yen Chong

M.Sc., The University of British Columbia, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Audiology and Speech Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2016

Abstract

Single-microphone noise reduction (SMNR) is implemented in hearing aids to suppress background noise. The noise-like feature in fricatives and affricates is susceptible to SMNR processing when background noise is present. Most SMNR studies have examined English speech materials but very few have examined Mandarin fricatives and affricates. In the present research, three studies were conducted to examine the acoustic and perceptual effects of SMNR on Mandarin fricatives and affricates. Study 1 aimed to test the validity of the inversion technique as a tool for separating speech and noise signals recorded from hearing aids in sound field. Study 1 showed that the inversion technique is a feasible and reliable tool for separating speech and noise post hearing-aid processing. However, fidelity of the retrieved speech signals showed variability between hearing aids. The acoustic effects of SMNR on Mandarin and English fricatives and affricates were examined in Study 2. Speech-plus-noise signals were presented to and recorded from one of two hearing aids mounted on a manikin, under SMNR-on and SMNR-off conditions. Speech signals were retrieved for subsequent acoustic analysis. The results showed that SMNR processing did not produce substantial acoustical changes in the temporal and spectral domain as measured in the Hearing Aid Speech Quality Index. Spectrographic analysis showed a reduction in frication-noise and release-burst intensity, and changes in the spectral mean. In Study 3, the Mandarin retroflex fricative and affricates, processed with and without SMNR, were used to examine the effects of SMNR on novel speech sound identification in noise by naïve listeners. Native English talkers might rely on bottom up processing to categorize the Mandarin retroflex sounds because these sounds were not in the English phonemic inventory. All listeners underwent five sessions of identification training and testing. The results showed that SMNR did not degrade the identification of novel speech sound in naïve listeners. Significant contributions of the present research are (i) the acoustic effects of SMNR on Mandarin and English fricatives and affricates were systematically documented and (ii) provided further evidence on SMNR having no effect on speech perception in noise.

Preface

Chapter 2, 3, and 4 were based on work conducted in UBC's Amplification Research Lab. For Chapter 2 and 3, I was responsible for the research design, development of stimulus files, collection of recordings from hearing aids, and data analysis. Part of the data analysis of Chapter 2 and 3 was conducted by the research assistants in the Amplification Research Lab under my supervision. Part of Chapter 3 was presented at the 2014 Canadian Academy of Audiology Conference. I presented the work at the podium session of the conference. The title of the abstract submitted to the conference was "Acoustic Analysis of Mandarin Retroflex Sounds Processed With and Without Noise Reduction". I was the main author and my supervisor, Dr. Lorienne Jenstad, was the second author for the abstract.

For Chapter 4, the ethics approval for the pilot study was obtained on December 8, 2014 (certificate number: H14-02854) and the ethics approval for the study was obtained on February 27, 2015 (certificate number: H14-03312) from the UBC Behavioural Research Ethics Board. I was responsible for the research design, development of the test stimuli, recruitment of subjects, test administration, and data analysis.

Please note that Chapters 2, 3, and 4 were written as separate manuscripts to be submitted for publication.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANL	Acceptable Noise Level
ANOVA	Analysis of variance
BTE	Behind-the-ear
CSII	Coherence Speech Intelligibility Index
HASQI	Hearing Aid Speech Quality Index
ICRA	International Collegium of Rehabilitative Audiology
KEMAR	Knowles Electronic Manikin for Acoustic Research
MBNR	Modulation-based noise reduction
RAU	Rationalize Arcsine Unit
RMS	Root-mean-squared
SD	Standard deviation
SII	Speech intelligibility index
SMNR	Single-microphone noise reduction
SNR	Signal-to-noise ratio
VCV	Vowel-consonant-vowel
VOT	Voice onset time

Acknowledgements

I want to thank my supervisor, Dr. Lorienne Jenstad, for her guidance and support throughout my doctoral program, and my Dissertation Committee members, Dr. Valter Ciocca and Dr. Anthony Herdman, for their invaluable input. I want to thank Dr. Anthony Herdman for taking the role as my acting supervisor while Dr. Lorienne Jenstad was unavailable. I want to thank Dr. Susan Small for being a part of my Program Advisory Committee in my doctoral program.

I want to acknowledge that the NSERC Discovery Grant, awarded to Dr. Lorienne Jenstad, provided portions of funding for the studies undertaken in this dissertation. I would also like to acknowledge my employer, National University of Malaysia (*Universiti Kebangsaan Malaysia*), as well as The Scholarship Division of Ministry of Education Malaysia, for granting study leave and providing stipend for my doctoral program for the first 36 months. I want to thank University of British Columbia (UBC) for awarding me the International Partial Tuition Scholarship and Faculty of Medicine Graduate Award.

Many thanks to the research assistants at the UBC Amplification Research Lab for helping me with data analysis: Flora Pang, Stephanie Kore, Myron Huen, Albert Chau, Lauretta Cheng, and Carney Zeng. Special thanks to Stephanie Kore, Lauretta Cheng, and Myron Huen who helped proof read my dissertation drafts. I would also like to thank manufacturers for providing hearing aids for the studies, and the participants who took part in the studies of this dissertation.

This humble work would not be successful without any of the persons and institutions mentioned above.

Dedication

I want to dedicate this humble work to my parents, Kok Khew Chong and Chooi Ngor Chan, who encouraged and supported me in every decision that I have made in pursuing my dreams.

I would also like to dedicate this work to my spouse, Chun Sing Wong, who continuously supported me throughout the roller-coaster journey of completing my PhD program, "Thank you for your sacrifices."

I want to dedicate this work to my children, Jia En Wong and Jia Rong Wong, for lighting up my days with happiness, "I hope to be your inspiration to thrive in every aspect in your lives."

Chapter 1: Introduction

Single-microphone noise reduction (SMNR), or commonly known as digital noise reduction, is implemented in hearing aids to suppress background noise. SMNR is different from directional microphone processing because it analyzes the acoustic input from one hearing-aid microphone whereas the directional microphones analyze the acoustic input from two hearing-aid microphones. Therefore, the former is termed “single-microphone noise reduction” throughout this dissertation to differentiate it from directional microphones. SMNR analyzes the physical properties of acoustic signals to determine the presence of speech or noise-like sounds and reduces hearing aid gain when noise-like sounds are present. Studies with adult hearing aid users have shown that SMNR improves listening comfort (Palmer, Bentler, & Mueller, 2006), increases acceptance of background noise (Mueller, Weber, & Hornsby, 2006; Peeters, Kuk, Lau, & Keenan, 2009; Ricketts & Hornsby, 2005), and reduces listening effort (Brons, Houben, & Dreschler, 2013; Gustafson, McCreery, Hoover, Kopun, & Stelmachowicz, 2014). Although SMNR does not improve speech intelligibility, it does not degrade intelligibility for adults or school-age children (Alcantara, Moore, Kuhnel, & Launer, 2003; Boymans & Dreschler, 2000; Nordrum, Erler, Garstecki, & Dhar, 2006; Stelmachowicz et al., 2010; Walden, Surr, Cord, Edwards, & Olson, 2000; Yuen, Kam, & Lau, 2006; Zakis, Hau, & Blamey, 2009). This evidence supports the use of SMNR in adults and school-age children (McCreery, Venediktov, Coleman, & Leech, 2012).

A cross-language paradigm was used in two studies to examine the effect of SMNR on novel speech sound discrimination in background noise by English-speaking listeners (Marcoux et al., 2006; Turgeon et al., 2009). The aim of using the cross-language paradigm was to provide a model of the effect of SMNR on language acquisition for children (Marcoux et al., 2006). One advantage of testing English-speaking listeners with non-native speech contrasts was to reduce bias related to contextual and linguistic cues (Marcoux et al., 2006). Due to the difficulty of adult listeners with normal hearing in perceiving non-native speech contrasts as compared to children, Marcoux et al. (2006) suggested that their research findings might indicate that young children fitted with hearing aids having SMNR were unlikely to exhibit problems acquiring language when the adult listeners could perform equally well when tested with non-native speech stimuli that was processed with and without SMNR. Marcoux et al. (2006) and Turgeon et al. (2009) tested native English-speaking adults and children with normal hearing using non-native Hindi speech contrasts. They reported that SMNR did not affect the discrimination of novel speech sounds among the listeners. The findings of these studies may not apply to fricatives and affricates because the stop consonants (i.e., Hindi dental vs. retroflex stops) are unlikely to be altered by SMNR processing as compared to fricative and affricate consonants that contain frication noise as a prominent feature in their acoustical waveforms. The frication noise in fricative and affricate consonants mimics the randomness of broadband noise (e.g., white noise and pink noise). In a hearing aid, this noise-like feature may result in reduced gain by SMNR because the algorithm applies to a mixture of speech and noise signals during SMNR processing. If SMNR affects the acoustics of fricatives and affricates, there is a potential that individuals with hearing loss may receive inconsistent exposure to these speech sounds when background noise is present. Studies have shown that inconsistent access to fricatives can lead to poor speech and language development (Moeller et al., 2007a; Stelmachowicz, Pittman, Hoover, Lewis, & Moeller, 2004).

As with English, fricatives are important in Mandarin. In English, the fricative /s/ is the third most commonly occurring consonant in English. The consonant /s/ has more than 15 functions such as grammatical markers for plurals (cats vs. cat), possessives (her's), and third party present tense (I drink vs. he drinks) (Denes, 1963; Rudmin, 1983). In Mandarin, about half of the consonants are fricatives and affricates (Hua & Dodd, 2000; Zhao & Li, 2009). Mandarin has the highest number of speakers worldwide, with approximately 848 million people speaking it as their first language (Lewis, Gary, & Charles, 2015). Most of the SMNR studies have tested English speech materials and only a few studies have tested speech materials from other languages such as Hindi, Swedish, and Dutch (Brons et al., 2013; Marcoux, Yathiraj, Cote, & Logan, 2006; Ng, Rudner, Lunner, Pedersen, & Rönnberg, 2013; Ng, Rudner, Lunner, & Ronnberg, 2015; Turgeon, Dostaler, Yathiraj, & Marcoux, 2009).To date, very few studies have examined the effects of noise reduction algorithms on Mandarin (Chen, Hu, & Yuan, 2015; Li et al., 2011; Liu, Zhang, Bentler, Han, & Zhang, 2012). The noise reduction algorithms examined by Chen et al. (2015) were used in cochlear implants, whereas the noise reduction algorithms examined by Li et al. (2011) was a computer-based noise reduction algorithm. Liu et al. (2012) examined a noise-reduction algorithm that targets transient noise. Hence, the results of these studies do not apply to SMNR used in hearing aids. To date, no study has examined the effect of SMNR used in hearing aids on the acoustic characteristics of Mandarin and English fricatives and affricates. Therefore, there is a need to investigate whether SMNR has acoustical effects on noise-like speech sounds such as fricatives and affricates, particularly those in the Mandarin and English language inventories. It is also important to examine whether acoustical changes due to SMNR processing have an effect on the perception of fricatives and affricates.

Mandarin retroflex fricatives and affricates provide an interesting test case for examining the effects of SMNR on novel speech sound identification in noise by listeners who have little or no exposure in Mandarin, such as native English talkers. First, retroflex fricatives and affricates are not available in English and the identification task can be more challenging when native English talkers are tested with these sounds. Testing non-native listeners with the Mandarin retroflex sounds can minimize the use of linguistic cues in a speech recognition task. For these reasons, the listeners may be forced to rely on bottom-up processing in the identification task. Second, studies have shown that non-native listeners with normal-hearing are disrupted more by background noise in recognition of foreign languages (Lecumberri, Cooke, & Cutler, 2010). Thus, the question of whether SMNR processing will have any effect on the identification of novel speech sounds in noise can be examined by using the Mandarin retroflex fricative and affricates as test stimuli. Furthermore, it is assumed that if normal-hearing naïve listeners can identify the Mandarin retroflex fricative and affricates under the SMNR-on condition better than in the SMNR-off condition, it is less likely for native Mandarin listeners to have difficulties in identifying the retroflex fricative and affricates under the same condition. This is because it is well documented in the literature that Mandarin listeners have better discrimination and identification performance on Mandarin contrasts as compared to the English listeners (Lee et al., 2012; Tsao, Liu, & Kuhl, 2006). Hence, the results will imply that SMNR has no detrimental effects on the identification of Mandarin retroflex fricative and affricates by Mandarin listeners.

In order to examine the acoustic effects of SMNR on speech signals, it is necessary to present speech and noise signals simultaneously to a hearing aid. This is because gain reduction applies to both speech and noise during SMNR processing, and the nonlinear processing of hearing aids may affect speech-only, noise-only, or speech-plus-noise signal differently. Subsequently, it is necessary to separate the speech-plus-noise output of the hearing aid into speech and noise signals for individual acoustic analysis. The introduction of the inversion technique by Hagerman and Olofsson (2004) has provided a means to isolate speech or noise from the speech-plus-noise signals post hearing-aid processing. In this technique, two speech-plus-noise signals (the noise waveform within one of the signals is 180^{o} out of phase) can be presented and recorded from a hearing aid. Subsequently, the recordings are summated or subtracted from each other to retrieve speech or noise, respectively. It is important to ensure that the implementation of the inversion technique will not affect the fidelity of the retrieved-speech signals prior to examining the acoustic effects of SMNR on these signals. For these reasons, three studies were conducted to examine the effects of SMNR on fricatives and affricates by addressing the following research questions:

i. Is the inversion technique a feasible tool for separating aided and unaided speech-plus-noise signals recorded in a sound field setting? Does the inversion technique affect the fidelity of retrieved-speech signals? (Study 1)

ii. What are the effects of SMNR on the acoustics of fricatives and affricates in Mandarin and English? (Study 2)

iii. Does SMNR affect novel speech sound identification in noise by naïve listeners? (Study 3)

1.1 Single-Microphone Noise Reduction

Background noise, defined as any unwanted or competing acoustic signal when a signal of interest such as speech or music is present, is a common nuisance for hearing aid users (Kochkin, 2010). About 66% of hearing aid owners report that they have difficulties hearing in noise: 31% report that hearing in noise is quite difficult while 35% report that hearing in noise is extremely difficult (Kochkin, 2010). Noise perceived by hearing aid users may originate from external (e.g., background noise, wind noise, or reverberation) or internal sources (e.g., hearing aid internal noise). Noise, like any other acoustic signal, is characterized by temporal, spectral, and amplitude features. Temporally, noise can be continuous (either modulated or unmodulated), interrupted, or transient. Spectrally, noise can be broadband or narrowband. Speech, which is broadband in nature, can also be regarded as noise. In terms of amplitude, noise can have high intensity or low intensity. The intensity of noise relative to a speech signal is often referred as the signal-to-noise ratio (SNR). Classification algorithms in hearing aids use these characteristics to detect noise and activate SMNR to reduce the noise. The main types of SMNR are modulation-based noise reduction (MBNR), synchrony detection, and spectral subtraction. These SMNR techniques differ based on how they determine the presence of noise or speech (Bentler & Chiou, 2006; Chung, 2004; Edwards, 2000; Hamacher et al., 2005; Kates, 2008; Schaub, 2008). Each hearing aid manufacturer may incorporate different types, or combination, of SMNR in hearing aids.

1.1.1 Modulation-based Noise Reduction (MBNR)

MBNR, or adaptive multi-channel noise reduction, is one of the most common types of SMNR implemented in commercial hearing aids (Bentler & Chiou, 2006; Chung, 2004). This algorithm uses the amplitude modulation characteristic of sounds to determine the presence of speech. The target modulation depth corresponds to the syllabic structures of speech signals (e.g., combinations of vowels and consonants) in quiet. MBNR determines that speech is present when the estimated modulation depth is 15 dB or more and the modulation rate is between 3 and 10 Hz (Schum, 2003). Signals with relatively lower modulation depths and higher modulation rates are categorized as speech-in-noise signals; signals with very low modulation depths (e.g., 0 dB) and high modulation rates are categorized as noise. Figure 1 shows examples of speech-in-quiet, speech-in-noise, and steady-state noise waveforms. In most hearing aids, MBNR processing is applied prior to feeding the digitized signals into the compression system (Schaub, 2008).

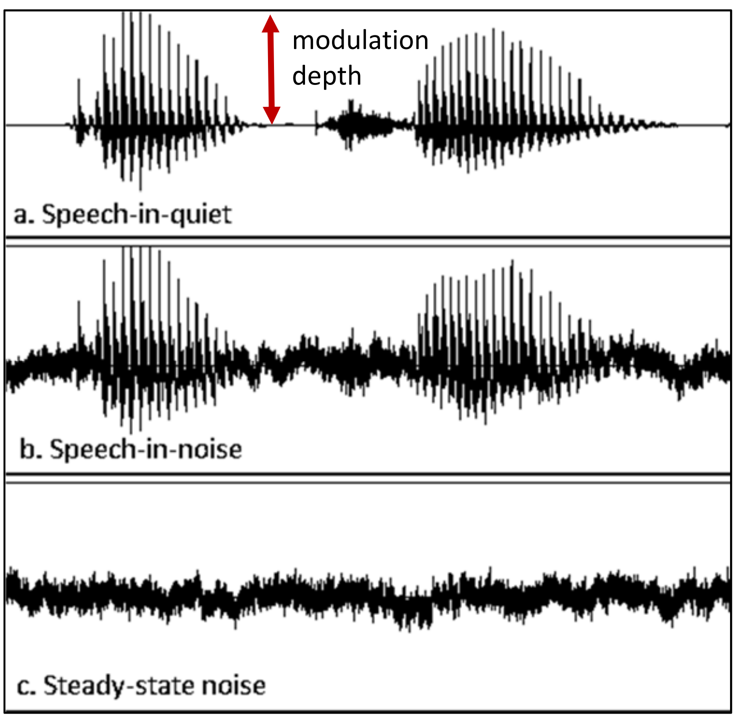


Figure 1. Acoustic waveforms of speech-in-quiet, speech-in-noise, and steady-state background noise.

In the first stage of hearing-aid processing, incoming acoustic signals are filtered into frequency channels. In each frequency channel, parameters such as the input level, modulation depth, and SNR are estimated. The input level in each frequency channel is monitored by level detectors with different time constants: a fast-level detector and a slow-level detector. It is assumed that speech has rapid level fluctuations whereas background noise (presumably steady-state noise) has slow level fluctuations over time. Thus, the fast-level detector monitors the level of a signal with rapid fluctuations whereas the slow-level detector monitors the level of a signal with slow fluctuations. The difference in the estimated level between the two level detectors is used to estimate the modulation depth of a signal within each frequency band (Zakis et al., 2009). Another way of estimating the modulation depth of a signal is by monitoring the short-term maximum level and minimum level of the signal; the difference between the maximum and minimum is the estimation of modulation depth (Schaub, 2008).

In the second stage of processing, the desired amount of gain reduction in each frequency channel is calculated. The amount of gain reduction varies with a set of pre-determined gain-decision rules. The gain-decision rules are proprietary to each hearing aid manufacturer and may take into account several factors: (i) threshold of MBNR (e.g., a specified modulation depth or input level), (ii) estimated SNR, (iii) frequency weighting function, and (iv) strength of MBNR. Therefore, different makes and models of hearing aids will exhibit different amounts of gain reduction for the same signals (Chung, 2004; Hoetink, Korossy, & Dreschler, 2009). For example, the GN ReSound Canta 780-D hearing aid has three MBNR strengths: (i) a mild strength that has a threshold of 15 dB modulation depth and a maximum gain reduction of 12 dB, (ii) a moderate strength that has a threshold of 15 dB modulation depth and a maximum gain reduction of 18 dB, and (iii) a strong strength that has a threshold of 20 dB modulation depth and a maximum gain reduction of 24 dB (Keidser, Hartley, & Carter, 2008). When noise is detected and the modulation depth of the signal is lower than 15 to 20 dB, gain reduction of 12 to 24 dB can be applied by this hearing aid. For other hearing aids, the amount of gain reduction may be dependent on the estimated SNR (e.g., maximum gain reduction is applied when the SNR is low whereas minimum or no gain reduction is applied when the SNR is high) and the frequency importance function incorporated in MBNR (e.g., minimum or no gain reduction is applied in a frequency band that conveys important speech information).

Other variable characteristics of MBNR include the time constants and the number of frequency channels (Bentler & Chiou, 2006; Chung, 2004). The “activation time”, "attack time," "engaging time," and "adaptation time" are equivalent terms that denote the time required by a MBNR to perform signal detection and gain reduction until the time at which the noise is 3 dB from its steady-state level (Bentler & Chiou, 2006; Bentler, Wu, Kettel, & Hurtig, 2008). The terms "release time" and "disengaging time" denote the speed of gain recovery to 0 dB gain reduction (Bentler & Chiou, 2006; Bentler et al., 2008). MBNR with a larger number of frequency channels (hence, smaller bandwidth in each channel) may have less effect on speech signals when a narrowband noise is present because gain reduction is restricted to the frequency channels where the noise is dominant (Bentler & Chiou, 2006). Conversely, when MBNR has a smaller number of frequency channels, the bandwidth for each channel is wider. When gain reduction is applied in a frequency channel with a wider bandwidth, this can potentially reduce the level of speech signals if the noise spectra overlap with the speech spectra. In some hearing aids, MBNR is used in conjunction with other forms of noise reduction such as Wiener filtering (Hamacher et al., 2005; see section 1.1.3 below). For example, the noise reduction in Siemens Acuris S and Siemens Triano hearing aids consists of a 16-channel MBNR and an adaptive Wiener filtering (Mueller et al., 2006; Palmer et al., 2006; Ricketts & Hornsby, 2005).

1.1.2 Synchrony Detection

The second type of SMNR is known as synchrony detection and this algorithm uses harmonics in sounds to determine the presence of speech. During speech production, opening and closing of the vocal folds generate synchronous patterns of energy (i.e., fundamental frequency and harmonics) within the speech signals. Synchrony detection uses a speech detector to detect this synchronous pattern to indicate the presence of speech within a signal. Only one hearing aid manufacturer, Oticon, implements this type of SMNR (Bentler & Chiou, 2006; Chung, 2004; Elberling, 2002; Schum, 2003) under the trademarked term, VoiceFinder (Elberling, 2002). First, the incoming signal is filtered into frequency channels and the degree of synchrony of the signal envelope across frequency channels is computed (Elberling, 2002). A high degree of synchrony indicates the presence of speech whereas a low degree of synchrony indicates the absence of speech (Elberling, 2002). No gain reduction is applied when speech is present (even at very low SNR) and the hearing aid is said to be in “speech mode.” Conversely, gain reduction is applied when the system detects that no speech is present and the hearing aid is said to be in “comfort mode” (Elberling, 2002; Schum, 2003).

1.1.3 Spectral Subtraction

The third type of SMNR is known as spectral subtraction and this algorithm estimates the noise spectrum during pauses in speech (Loizou, 2013). The presence of speech is detected by a voice activity detector. Subsequently, the estimated noise spectrum is subtracted from the speech-plus-noise spectrum to estimate the clean-speech spectrum. Then, the estimated clean-speech spectrum is fed into an inverse Fast Fourier Transform processor to generate the enhanced speech signal. Some of the algorithms that can be categorized as spectral subtraction include Wiener filtering (which works best for stationary noise), adaptive Wiener filtering (which works best for fluctuating noise), and Euphraim-Malah (Kates, 2008). Spectral subtraction is implemented in Starkey hearing aids such as the Destiny 1200 (Stelmachowicz et al., 2010). This algorithm compares the spectrum of ongoing input signals with the estimated noise spectrum; the estimation of the noise spectrum is paused when speech is detected by a voice activity detector.

In summary, SMNR processing is proprietary to manufacturers: different makes and models of hearing aids can have different types of SMNR or a combination of SMNR. Some of the limitations of the three types of SMNR discussed above include (i) the estimation of the presence of noise may not be accurate if the noise is a fluctuating signal; (ii) the efficacy of SMNR may be reduced if the unwanted noise is concurrent speech. In the next section, the effects of SMNR on speech perception are discussed.

1.2 The Effects of SMNR on Speech Perception

1.2.1 Adult Population

For the past decade, behavioural studies have been conducted to examine the efficacy and the effectiveness of SMNR. These behavioural measurements include speech-intelligibility testing, paired-comparisons, scale ratings, and self-report questionnaires (see Appendix A and Appendix B for summary). Most studies that involved adult participants (either with normal hearing or with hearing loss) showed that SMNR has no effect on speech intelligibility despite the use of different hearing aids, microphone modes, noise stimuli, stimulus presentation paradigms, types and parameters of SMNR. Very few studies showed a small but statistically significant improvement in speech intelligibility with the use of SMNR (Bray & Nilsson, 2001; Kuk et al., 2011; Peeters et al., 2009). However, some positive results were reported for acceptance of background noise and preference rating (Brons et al., 2013; Mueller et al, 2006). A few studies showed that acclimatization did not seem to be a confounding factor on speech-intelligibility performance (Alcantara et al., 2003; Bentler et al., 2008; Boymans & Dreschler, 2000; Mueller et al., 2006; Nordrum et al., 2006; Ricketts & Hornsby, 2005; Walden et al., 2000; Yuen et al., 2006; Zakis et al., 2009). Taken together, these studies suggested that conventional speech-intelligibility testing might not be sensitive enough to detect the benefit of SMNR, if any, in speech-intelligibility performance. However, three studies showed that by incorporating auditory tasks that require more cognitive resources (which is more typical of real-life situation), the positive effects of SMNR can be detected (Ng et al., 2013; 2015; Sarampalis et al., 2009). These studies are discussed in the following subsections.

1.2.1.1 No Significant SMNR Effect on Speech Intelligibility

Walden et al. (2000) tested 40 older adults (52 to 76 years old) with hearing impairment and found that activation of SMNR had no effect on speech-intelligibility performance, as well as the subjective speech understanding and sound quality ratings. In the study, each participant was fit with a pair of GN ReSound behind-the-ear (BTE) hearing aids with three listening programs: (i) SMNR-on with directional microphones, (ii) SMNR-off with directional microphones, and (iii) SMNR-off without directional microphones. The type of SMNR tested was MBNR with a maximum attenuation of 7.5 dB when no modulation was detected. The participants were given four to six weeks of field trial before the speech-intelligibility testing and subjective ratings for three domains (i.e., speech understanding, listening comfort, and sound quality) were administered. The results showed that SMNR did not improve the subjects' speech-intelligibility performance. One caveat of this study was that SMNR was not evaluated independently from directional microphones, hence, any benefit from SMNR might be obscured by the directional microphone processing.

Using Cantonese test materials, Yuen et al. (2006) also found that SMNR did not improve or degrade the speech-intelligibility performance of hearing aid users. Yuen et al. (2006) fitted nine subjects, aged 39 to 79 years, with Phonak Perseo hearing aids. The subjects were given two weeks of field trials to acclimatize to the hearing aid processing. Speech intelligibility was tested in three listening conditions: (i) SMNR-on, (ii) SMNR-on plus directional microphones, and (iii) directional microphones. The test stimuli included Cantonese sentences presented from the front loudspeaker and background noise either presented from the front (noise-front) or the side (noise-side) loudspeaker. For the noise-front condition, all of the three listening conditions yielded similar results. These findings indicated that SMNR yielded equal benefits as directional microphones when there was no spatial separation between speech and noise. For the noise-side condition, the SMNR-on plus directional microphone listening condition and the directional microphone listening condition yielded equal intelligibility scores. These results indicated that activation of SMNR did not decrease the speech intelligibility scores of the subjects when there was a spatial separation between speech and noise.

Using an improved research design, Boymans and Dreschler (2000) also found that SMNR did not improve or degrade the speech-intelligibility performance of adult listeners, but the preference score was slightly higher in the SMNR-on condition than the SMNR-off condition. The improvements in the study design as compared to the two previous studies included (i) SMNR and directional microphones were tested independently during the speech-intelligibility testing; and (ii) the speech-intelligibility testing were administered before and after each field trial. Sixteen participants, aged 40 to 75 years, were fitted with Siemens Prisma BTEs. The speech-intelligibility testing was conducted under two conditions: SMNR-on and SMNR-off. The test stimuli included (i) sentences spoken by a male talker presented in cocktail party noise and (ii) sentences spoken by a female talker presented in low-frequency car noise. The noise level was kept constant at 65 dBA, while the speech level was varied using an adaptive up-down method to obtain the speech reception threshold for each subject. The results showed that the speech-intelligibility performance in the SMNR-on condition was not significantly different than the SMNR-off condition. In addition, participants underwent a paired comparison testing to determine the preference among the four listening programs. The results showed that the SMNR-on program yielded a slightly higher preference ranking over the SMNR-off program.

In another study, Alcantara et al. (2003) found that there was no significant difference in the speech-intelligibility performance between the SMNR-on and SMNR-off conditions, as well as the subjective ratings on (i) comprehension, (ii) listening comfort, (iii) listening effort, (iv) sound clarity, and (v) sound quality. Alcantara et al. (2003) fitted each of the eight participants, aged 49 to 83 years, with Phonak Claro 21 dAZ in-the-ear hearing aids with two listening programs: (a) SMNR-on and (b) SMNR-off. The SMNR was known as the Fine-scale Noise Canceller™, which was also a MBNR. The strength of the SMNR was set to a moderate degree. Subjects were asked to use both programs equally during the three-month trial period and their speech-intelligibility performance was tested following the field trial. Sentences mixed with each of the four noises were used as the stimuli: (i) steady-state speech-shaped noise, (ii) modulated speech-shaped noise, (iii) steady-state noise with spectral dips, and (iv) modulated noise with spectral dips. Background noise was presented from 0^{o} azimuth at a fixed level of 65 dB SPL and the level of sentences was varied using an adaptive up-down method to obtain the speech reception threshold for each subject. The results showed that there was no significant difference in the speech-reception threshold among the four noise conditions; there was also no significant difference between the SMNR-on and SMNR-off conditions. Because the SMNR was set to a moderate strength in this study, it is unknown whether a higher strength of the SMNR would have yielded a different outcome.

Examining a combination of SMNR types, such as MBNR plus adaptive Wiener filter, Mueller et al. (2006) and Ricketts and Hornsby (2005) yielded similar results. Ricketts and Hornsby (2005) found that a 16-channel MBNR plus an adaptive Wiener filtering did not affect the average speech-intelligibility performance of 14 subjects. However, the subjects preferred the SMNR-on condition more than the SMNR-off condition. The subjects, aged 42 to 83 years, were tested under four listening conditions similar to the ones examined by Boymans and Drescher (2000): SMNR-on or SMNR-off with omnidirectional microphone or directional microphones. Speech stimuli were presented at two SNRs (+6 and +1 dB SNR) in a diffuse-noise condition. Using another model of hearing aid from the same manufacturer as compared to Ricketts and Hornsby (2005), Mueller et al. (2006) performed speech-intelligibility testing and the Acceptable Noise Level testing (ANL; Nabelek, Tampas, & Burchfield, 2004; Nabelek, 2006) among 22 adults, aged 23 to 76 years, with a mild-to-moderate sensorineural hearing loss. The ANL test is used to measure how much noise one can tolerate while listening to speech at a comfortable level, which is defined as the dB difference between the most comfortable level for speech and the acceptable background noise level. Mueller et al. (2006) tested the Siemens Acuris S BTE with a 16-channel MBNR and an adaptive Wiener filtering. Speech-intelligibility testing was conducted under two conditions: (i) SMNR-on and (ii) SMNR-off. The noise signal was continuous throughout the testing to ensure that the SMNR was activated during the course of testing. They found that there was no significant difference in the subjects' speech reception thresholds obtained in the SMNR-on and SMNR-off conditions, but an improvement of 4.2 dB for the ANL test was observed. Consistent with the previous studies, Mueller et al. (2006) and Ricketts and Hornsby (2005) found that there was no significant SMNR effect on speech-intelligibility performance among individuals with hearing loss, despite a combination of two types of SMNR (MBNR plus adaptive Wiener filter) was examined. However, Mueller et al. (2006) showed that a combination of MBNR and Wiener filtering was beneficial in improving acceptance of background noise.

Unlike previous studies that only tested one hearing aid, Nodrum et al. (2006) evaluated the effects of SMNR in four hearing aids (GN ReSound Canta 7, Oticon Syncro, Phonak Perseo, and Siemens Acuris) and found that SMNR in all hearing aids did not have any effect on speech intelligibility even though each device's SMNR was set to maximum strength. All four hearing aids had MBNR and one also had synchrony detection. Sixteen adults, aged 58 to 90 years, with moderate to severe sensorineural hearing loss were tested. The speech-intelligibility performance of the subjects was tested under four listening conditions similar to the studies by Boymans and Dreschler (2000) and Ricketts and Hornsby (2005): SMNR-on or SMNR-off with omnidirectional or directional microphones. Within each hearing aid, they found that speech reception thresholds obtained with SMNR-on were not significantly better than SMNR-off regardless of the microphone mode used.

In another study, Brons et al. (2013) tested normal-hearing subjects with four other hearing aids (Phonak Exelia M, ReSound Azure, Starkey Destiny 1200, Widex Mind 440) and showed that SMNR did not affect speech-intelligibility performance or listening effort ratings but decreased noise annoyance ratings (i.e., noise was rated less annoying). In addition, they also found that there was a trade-off between gain reduction and speech naturalness perception at the lower SNR condition (-4 dB SNR). They found that normal-hearing subjects preferred SMNR-on conditions more than SMNR-off conditions at the +4 dB SNR condition but not at the -4 dB SNR. Brons et al. (2013) suggested that noise annoyance contributed more to the preference ratings than speech naturalness, whereas speech intelligibility and listening effort were not correlated with the preference ratings.

1.2.1.2 Significant SMNR Effect on Speech Intelligibility

In contrast to the aforementioned studies, a few studies showed that there was a small but significant positive effect of SMNR on speech intelligibility (Bray & Nilsson, 2001; Kuk, Peeters, Lau, & Korhonen, 2011; Peeters et al., 2009). Bray and Nilsson (2001) tested 20 adults, aged 34 to 84 years, with Sonic Innovations Natura 2E directional BTE. Speech-intelligibility testing was conducted under the SMNR-on and SMNR-off conditions. There was a significant difference between the SMNR-on and SMNR-off conditions and the amount of improvement in speech reception threshold ranged from 1.0 to 1.7 dB SNR.

Examining another type of SMNR (Speech Enhancer™), Peeters et al. (2009) found that SMNR processing yielded a small but significant improvement in speech-intelligibility testing as compared to the SMNR-off condition. They also found 3.3 dB of benefit in the

ANL test. According to Peeters et al. (2009), Speech Enhancer™ takes into account the degree of hearing loss and speech intelligibility index during gain adjustment; a maximum of 12 dB gain reduction and 5 dB gain increment is provided by the algorithm. In their study, 18 adults, aged 44 to 89 years, were fitted with Widex Inteo hearing aids. The speech-intelligibility performance of the subjects was tested under four listening conditions: SMNR-on or SMNR-off with omnidirectional microphone or directional microphones. Unlike other studies, noise was presented 60 seconds prior to the beginning of sentences and continuously during the presentation of the sentences to ensure activation of Speech Enhancer™ throughout the test. Compared to the SMNR-off condition, activation of Speech Enhancer™ yielded a significant improvement (2.5 dB SNR) in the speech reception threshold. Nonetheless, consistent with previous studies, when Speech Enhancer™ was used in conjunction with directional microphones, no significant improvement (0.6 dB SNR) was obtained. Kuk et al. (2011) also tested a Widex hearing aid with Speech Enhancer™ and found that the speech reception thresholds were significantly better (5.2 dB) in the SMNR-on condition as compared to the SMNR-off condition. Peeters et al. (2009) and the Kuk et al. (2011) reported higher magnitude of benefit (2.5 and 5.2 dB, respectively) as compared to Bray and Nilsson (2001) who reported 1 dB of improvement in the speech reception threshold. This discrepancy can be attributed to the types of SMNR examined. For example, MBNR was tested in the study by Bray and Nilsson (2001) whereas the Speech Enhancer™ was tested in the studies by Peeters et al. (2009) and Kuk et al. (2011).

Two other studies examined the effect of altering SMNR parameters on speech intelligibility. One study found that SMNR with longer activation times (8 and 16 s) yielded a higher speech intelligibility score (Bentler et al., 2008) whilst the other study found that a fixed gain reduction or variable gain reduction in SMNR processing did not have an effect on speech intelligibility (Zakis et al., 2009). In the study by Bentler et al. (2008), 25 adults with mild to moderately-severe sensorineural hearing loss were fitted with Starkey Axent BTEs with MBNR. Four conditions of SMNR with various activation times were tested: (i) SMNR-off, (ii) SMNR-4s, (iii) SMNR-8s, and (iv) SMNR-16s where 4s, 8s, and 16s refer to the activation time in seconds. The gain reduction provided in the SMNR-on conditions was 1 dB for speech-in-babble stimulus presented at 65 dB SPL; the gain reduction did not differ between different activation times. During the speech-intelligibility testing, speech stimuli were presented from the front at adaptive levels whereas multi-talker babble noise was presented from a loudspeaker located in the rear at each of the two fixed levels (62 and 78 dB SPL). Their results showed that there was no significant difference in speech-intelligibility scores among the four SMNR activation times when noise was presented at 78 dB SPL. When noise was presented at a lower level (62 dB SPL), the SMNR-8s and SMNR-16s conditions yielded better speech intelligibility scores than the SMNR-off.

In another study, Zakis et al. (2009) examined the effect of maximum amount of gain reduction on speech-intelligibility performance among 10 subjects with hearing loss. The hearing aid used had an open-platform digital signal processor in which any software algorithms could be loaded. The SMNR had eight processing channels and used MBNR. The conditions for the speech-intelligibility testing were (i) fixed gain reduction, (ii) variable gain reduction, and (iii) no gain reduction. For the first two configurations, gain reduction occurred for a modulation depth of 10 dB or lower. The fixed gain configuration had a maximum gain reduction of 10 dB but the variable gain configuration had a maximum gain reduction that varied from 2 to 10 dB across frequency channels. The speech levels were varied adaptively; noise level was constant at 65 dBA. The noise signals included speech-shaped noise and babble noise. The statistical tests showed that speech-reception thresholds obtained under the three configurations did not differ significantly.

1.2.1.3 Testing SMNR with a Dual-task Paradigm

Because most of the studies discussed above showed neutral effects of SMNR on speech-intelligibility performance, further attempts were made to determine whether SMNR affects speech perception when there is a competing task (Sarampalis, Kalluri, Edwards, & Hafter, 2009) and whether SMNR affects memory processing for speech (Ng et al., 2013; 2015). In the first study, Sarampalis et al. (2009) examined the effects of SMNR on the ability of young adults with normal hearing to identify and recall spoken words while listening to sentences presented in noise. Sentences with low and high contexts were used as test stimuli. The sentences were presented in quiet and in four-talker babble at two SNR conditions (-2 and +2 dB) with the noise kept at a constant level (65 dB SPL). The subjects were presented with five blocks of 48 sentences, randomly chosen from eight lists. Their first task was to repeat the last word of each sentence; their second task was to recall all of the sentence-final words in a free-recall format after listening to every eight sentences. For the low-context sentences, SMNR did not have a significant effect on the number of words recalled in the -2 and +2 dB SNR conditions. For the high-context sentences, noise reduction had a significant effect on the numbers of words recalled in the -2 dB SNR condition. In a subsequent experiment, participants were required to repeat each sentence heard while performing a competing visual task (indicate whether a digit displayed on a screen was even or odd by using the keyboard arrows). In this dual-task paradigm, male-voice sentences were presented in quiet and in four-talker babble at -6, -2, and +2 dB SNR, and with SMNR turned on and off for each SNR condition. The results showed that speech-intelligibility performance did not differ significantly at any SNR condition between the SMNR-on and SMNR-off conditions. However, the mean reaction time for the secondary task (i.e., visual task) was significantly better with SMNR-on at poor SNR listening condition (e.g., -6 dB SNR) than at -2 and +2 dB SNR condition. The results of the first experiment indicated that SMNR facilitated memory performance that required higher level of auditory processing (i.e., auditory memory); the second experiment indicated that SMNR might reduce the cognitive load (shorter reaction time for the visual task) and this positive effect could be detected by incorporating a competing secondary task in the test paradigm.

Using a similar approach, Ng et al. (2013; 2015) tested adults, aged 32 to 65 years, with moderate to moderately severe SNHL who were experienced hearing-aid users. In the studies, the effects of different types of noise, noise reduction, and working memory capacity for speech were examined and the researchers found that the recall performance for target native speech (Swedish) was more disrupted in Swedish (native language) than in Cantonese (non-native language) four-talker babble and in stationary speech-shaped noise. Their first study published in 2013 revealed that binary masking, a type of noise reduction algorithm that is not currently available in commercial hearing aids, had a positive effect (improvement in recall performance) on individuals with high working memory capacity. In the later study, the researchers found that this positive effect of noise reduction could be extended to individuals with low working memory when the memory task was made less demanding (fewer items to remember and more favourable SNR). Both studies incorporated the reading span test to assess participants' working memory capacity and the sentence-final word identification and recall test to assess memory for speech processing. The results of the reading span test were used to group the subjects into two groups: (i) high working memory and (ii) low working memory. A participant's tasks included identification of sentence-final words after each sentence and free recall of all sentence-final words after each list. Ng et al. (2013; 2015) showed that binary masking noise reduction had positive effects on higher level of cognitive processing for speech among individuals with normal hearing and hearing loss.

Studies in the adult population generally found that SMNR had no negative effect on speech intelligibility and may provide some benefits in terms of listening comfort and release of cognitive load for a secondary task. Overall these results support the use of SMNR in the adult population. In the next section, the effects of SMNR on the pediatric population are discussed.

1.2.2 Pediatric Population

Only a few studies examined the effects of SMNR on the pediatric population and these studies showed that SMNR has no negative effect on speech recognition in noise among school-age children with hearing loss (Pittman, 2011b; Stelmachowicz et al., 2010). Similar to the results in studies with adult participants, one study showed that SMNR improved listening comfort and subjective clarity ratings in school-age children with normal hearing (Gustafson et al., 2014).

Stelmachowicz et al. (2010) tested speech-intelligibility performance of 16 children, aged 5 to 10 years, with mild to moderately-severe hearing loss, and found that SMNR had no significant positive or negative effect on the overall perception of nonsense syllables, words, and sentences when the data were collapsed across the entire group. However, there were relatively large individual differences among the 5- to 7-year-old children in the nonsense syllable and word recognition tests. In the nonsense syllable test, only two of eight children had significantly better performance when SMNR was turned on; another three children had significantly poorer performance when SMNR was turned on. In the word-recognition test, only one child had significantly better performance when SMNR was turned on; four out of eight children had significantly poorer performance when SMNR was turned on. These results indicated that it might be difficult to predict the effect of SMNR on individual performance of younger children based on the group results, due to the large individual variability in the nonsense syllable and word recognition tests. Hence, there is still a concern regarding the use of SMNR for younger children with hearing loss.

Using a dual-task paradigm, Pittman (2011b) found that SMNR did not improve or degrade children's auditory performance when there was a competing visual task. Thirty children with hearing loss, aged 8 to 12 years, were tested in her study. The children's tasks were to categorize words presented in noise while completing a complex visual task. The auditory stimuli included 50 words from each of the three noun categories (people, food, animals), giving a total of 150 words. The visual task consisted of dot-to-dot games in which each child was required to count numbers between dots in increments of three. The children were fitted with Siemens Explorer 500 BTEs and underwent the word-categorization test twice: once with SMNR turned on and another with SMNR turned off. No feedback was given to the children regarding their performance. This study showed that the categorization performance among the children was similar when SMNR was turned on or off. Although not examined in statistical analysis, the performance of the secondary task (i.e., dot-to-dot games) was very similar between the SMNR-on and SMNR-off conditions.

The aforementioned studies with pediatric population suggested that SMNR had no positive or negative effect on speech-intelligibility performance on average. One study, however, examined the effect of SMNR on listening effort and sound clarity rating among school-age children (Gustafson et al., 2014). In the study, 24 normal-hearing children (aged 7 to 12 years) participated in the study. Test stimuli included vowel-consonant-vowel (VCV) words that were recorded from two hearing aids (Oticon Agil Pro and Phonak Naida) under two SMNR conditions (SMNR-on and SMNR-off) and two SNR conditions (0 dB SNR and +5 dB SNR). The children listened to the VCV nonsense words through headphones and repeated the words; listening effort was quantified by the children's verbal response time. The children were also required to indicate the clarity of the words they heard by pointing to one of the six photos representing visual images that faded progressively in terms of visual clarity. Their results showed that the verbal response time was significantly lower and the clarity ratings were significantly higher for the VCV tokens processed with SMNR. Gustafson et al. (2014) concluded that SMNR significantly decreased listening effort among school-age children with normal-hearing.

The three studies that examined SMNR in the pediatric population showed that SMNR is not detrimental to school-age children. However, no study has tested if SMNR should be activated for infants who wear hearing aids, perhaps due to the challenges in conducting behavioural testing in infants and young children. It is unknown if SMNR will affect speech acquisition in this population. Moreover, there is lack of acoustic data to show whether SMNR affects the acoustic properties of speech sounds that are important for speech perception. More empirical evidence is required before SMNR can be recommended for infants and young children due to the unknown effects of SMNR on speech acquisition in this population (McCreery et al., 2012). In addition, all of the studies discussed above (including both adult and pediatric subjects) used speech materials that were native to the test subjects; it is unknown whether SMNR will affect novel speech sound perception in noise for adults and children.

1.2.3 The Effects of SMNR on Novel Word and Novel Speech Sound Perception

The effect of SMNR on speech acquisition has been examined using two different approaches. In one approach, children with normal hearing and children with hearing loss were tested with novel words that contained speech sounds from subjects' native language (Pittman, 2011a) whereas in another approach, adults and children with normal hearing were tested with novel speech sounds not within the subjects' native language (Marcoux et al., 2006; Turgeon et al., 2009). The second approach may also be called a cross-language approach since non-native speech materials were used to test native English-speaking listeners. The advantage of using a cross-language paradigm as opposed to conventional speech-intelligibility testing is that the results are not confounded by linguistic knowledge of a language and familiarity with the speech materials.

In the first study, Pittman (2011a) examined the effect of SMNR on novel word learning in 26 children, aged 8 to 9 years and 11 to 12 years, with mild to moderate hearing loss and found that novel word learning in noise improved significantly for the 11- to 12-year-old children with the use of SMNR; however, SMNR had no positive or negative effect for the younger age group (age 8 to 9 years). All children were fitted with BTEs that implemented MBNR and Wiener filtering as the noise reduction system. Both of these noise reduction systems worked independently but simultaneously. The speech stimuli consisted of 15 two-syllable nonsense words spoken by a female talker and were novel to the children. The noise stimulus was a steady-state broadband noise. Speech and noise were presented at 0 dB SNR to the subjects via a loudspeaker positioned at 0° azimuth. The children's task was to learn to match a novel word with an object through playing a computer game; feedback was given only for a correct response. Each child underwent three blocks of testing; each block consisted of 100 trials (20 repetition X 5 words). One block was presented in quiet, another two blocks were presented in noise with and without SMNR turned on. Pittman (2011a) showed that SMNR did not affect novel word learning in younger children (8 to 9 years old) and might be beneficial for older children with hearing loss (11 to 12 years old). Although nonsense words were used as the novel word stimuli in the Pittman (2011a) study, the phonemes within each word were selected from the phonemic system of the subjects' native language.

Unlike the study conducted by Pittman (2011a), Marcoux et al. (2006) examined the effect of SMNR on novel speech sound discrimination by native English-speaking adults and found that SMNR did not improve or degrade the discrimination ability of adult listeners. In the study, two groups of 10 adults with normal hearing were tested over a period of four sessions. In each session, the subjects were presented pairs of words that contained voiced and voiceless Hindi dental-retroflex contrasts (i.e., /d̪ - ɖ/ and /t̪ - ʈ/, respectively) in steady-state noise. The subjects' task was to indicate whether the words sounded the same or different and no feedback was given regarding their performance. The experimental group was presented with stimuli preprocessed by a hearing aid (Widex Senso Diva BTE with MBNR as the noise reduction system) with SMNR turned on whereas the control group was presented with unprocessed stimuli. The stimuli were presented via headphones at 0 dB SNR. The results showed that there was no difference in discrimination performance between the experimental group and the control group across all four test sessions. Electroacoustic measurement on the SMNR-processed and unprocessed stimuli revealed that

SMNR decreased the overall level of the processed stimuli by 5 to 6 dB but improved the overall SNR by 1 dB as compared to the unprocessed stimuli. This study showed that although the overall level of speech plus noise stimuli was decreased by SMNR processing, discrimination ability of non-native speech sounds was unaffected.

Using a subset of stimuli (9 out of 90 pairs of stimuli) from the Marcoux et al. (2006) study, Turgeon et al. (2009) tested the effect of SMNR on novel speech sound discrimination by native English-speaking children and similar findings were obtained as in the Marcoux et al. (2006) study. Nineteen English-speaking children, aged 4 to 5 years, with normal hearing were tested. Stimuli were presented to each subject via headphones at 0 and +5 dB SNR. Half of the participants listened to Hindi speech contrasts that were processed by SMNR and another half of the participants listened to unprocessed stimuli. Their task was to select the “same” button when they perceived the sounds as identical or select the “different” button when they perceived the sounds as different. Correct responses were reinforced by presenting a short video to the subjects. If the subjects responded incorrectly, they were instructed to try again for that particular trial. Turgeon et al. (2009) found that SMNR processing and SNR conditions did not affect the discrimination performance of non-native speech contrasts by the children. In summary, three of the studies discussed above suggested that SMNR did not affect novel word learning by children with hearing loss and novel speech sound discrimination by adult and children with normal hearing. Further, two of the studies demonstrated the feasibility of using the novel approach of cross-language paradigm to examine the effect of SMNR on speech acquisition (Marcoux et al., 2006; Turgeon et al., 2009).

The studies discussed in section 1.2.1 to 1.2.3 are summarized in Appendices A and B. From the list of studies, it can be observed that the ratio between adult studies and pediatric studies is large, with only five studies involving pediatric samples. In addition, all of the pediatric studies involved school-age children who have acquired basic speech and language skills. There are still questions remaining whether SMNR processing affects novel speech sound perception; listeners may rely on bottom-up processing to help them discriminate and recognize novel speech contrasts. In bottom-up processing, acoustic properties of auditory stimuli are one of the important elements to ensure that the brain can encode the auditory stimulus. As such, it is important to know how SMNR affects the acoustics of speech sounds and whether these acoustic changes will affect speech perception. However, most of the SMNR studies reported in the literature focused on behavioural measurements, and very few examined the acoustics of speech after being processed by SMNR. Therefore, the second experiment in this dissertation was designed to answer the question whether SMNR processing affects the acoustic properties of speech sounds. Among the SMNR studies, a few showed the feasibility of using a cross-language paradigm to examine the effect of SMNR on novel speech contrast discrimination (Marcoux et al., 2006; Turgeon et al., 2009). To date, no study has examined whether SMNR will affect novel speech sound identification in noise, particularly using noise-like Mandarin consonants such as fricatives and affricates even though Mandarin is the most widely used/spoken language in the world. Therefore, the third experiment in this dissertation is designed to answer the question whether SMNR affects novel speech sound identification in noise by naive listeners. An identification task was chosen over a discrimination task because the identification task forces the subjects to develop and use phonetic categories (Logan, Lively, & Pisoni, 1991), and identification is a higher level of auditory skill than discrimination (Loven, 2009).

1.3 The Importance of the Acoustic Characteristics on the Perception of Fricatives

The fricatives /s, z/ serve as important grammatical markers in English (Rudmin, 1983). The consequence of not having consistent access to English fricatives when learning speech and language can lead to poor speech and language performance in English (Moeller et al., 2007a; Stelmachowicz et al., 2004). Studies have shown that early-identified infants with hearing loss have delayed development in consonants—particularly fricatives and affricates—and vocabulary production despite having access to current hearing aid technologies, when compared to their peers with normal hearing (Moeller et al., 2007a; 2007b).

Perception of the fricatives /s, z/ relies on the acoustic characteristics of these phonemes such as overall level (and therefore audibility) and spectral properties. First, perception of fricatives by listeners with hearing loss is closely related to the audibility of frication noise (Hedrick, 1997; Pittman & Stelmachowicz, 2000; Zeng & Turner, 1990). Zeng and Turner (1990) reported that the performance of three subjects with hearing loss in recognition of the voiceless fricatives /f, Ɵ, s, ʃ/ was attributed to the audibility of the fricative noise rather than the transitional cues. In the fricative recognition task, subjects with hearing loss performed equally well as the subjects with normal hearing when they were given equivalent audibility of the fricative cue. However, when the transitional cues were provided, the subjects with hearing loss did not perform as well as the subjects with normal hearing (Zeng & Turner, 1990). Subsequently, Pittman and Stelmachowicz (2000) examined the perceptual weighting strategy used by adults and children, with and without hearing loss, to perceive English fricatives within four nonsense syllables (/us/, /uʃ/, /uf/, and /uΘ/). The short-term audibility measure was used to estimate the audibility of a short segment within a syllable, such as the vowel or fricative segment, and was calculated using a modified version of the speech intelligibility index (Pittman & Stelmachowicz, 2000). They found that the identification performance of the syllable /us/ increased as the short-term audibility of the fricative /s/ segment increased. Pittman and Stelmachowicz (2000) concluded that children and adults with hearing loss relied on the frication noise segments more than the vowel segments or the formant transition segments to perceive the fricative /s/ in the syllable /us/ context. The findings of these studies suggest that audibility of the sibilant noise is important for listeners with hearing loss to perceive fricatives.

Second, the spectral properties are important to distinguish place of articulation of fricatives (Jongman, Wayland, & Wong, 2000). Hedrick (1997) reported that three subjects with sensorineural hearing loss used the spectral cues of fricative noise to distinguish between the fricatives /s/ and /ʃ/ rather than using the consonant–vowel transitional cues. Studies have also found that the bandwidth of fricative noise is important for children to perceive the fricatives /s, z/ produced by female talkers (Kortekaas & Stelmachowicz, 2000; Stelmachowicz, Pittman, Hoover, & Lewis, 2001; Stelmachowicz, Pittman, Hoover, & Lewis, 2002). Kortekaas and Stelmachowicz (2000) tested 5- to 10-year-old children and adults with normal hearing. The subjects listened to plural words, spoken by a male talker, embedded in speech-spectrum noise at +10 and +30 dB SNR. Kortekaas and Stelmachowicz (2000) found that the fricative /s/ with a reduced bandwidth yielded poorer detection performance in children as compared to the adult subjects. In another study, Stelmachowicz et al. (2001) found that children and adults with hearing loss required a stimulus bandwidth of 0 to 9 kHz to achieve optimum performance in the identification of the fricative /s/ spoken by a female talker and a child talker as opposed to a male talker (i.e., 0 to 5 kHz stimulus bandwidth). Subsequently, Stelmachowicz et al. (2002) studied the aided perception of /s, z/ by children with hearing loss. They found that frequencies between 2 and 8 kHz were important for the perception of fricatives /s, z/ spoken by a female talker whereas frequencies between 2 and 4 kHz were important for the perception of fricatives /s, z/ spoken by a male talker. Stelmachowicz et al. (2002) reported that children with hearing loss had poorer performance on the perception of the fricative /s/ spoken by a female talker and this was attributed to the limited high frequency gain of hearing aids. These studies suggest that the spectral content of fricatives is important for listeners with hearing loss to perceive fricatives.

Due to the importance of fricatives in the English language, the importance of audibility of frication noise, and the effects of spectral properties on perception of fricatives, the acoustic characteristics of these phonemes are of utmost importance for individuals with hearing loss who are developing speech and language. Acoustic characteristics and audibility of fricatives may be affected by limitations of hearing aids that could include insufficient gain in the high frequency regions and potential adverse effects of advanced digital signal processing such as SMNR. As a result, there is a possibility that individuals with hearing loss may receive inconsistent exposure to these phonemes. Given the importance of acoustic characteristics on the perception of fricatives, there is a need to investigate whether SMNR has acoustic effects on noise-like speech sounds such as fricatives and affricates. Acoustic measurement such as spectrographic analysis on SMNR-processed speech sounds may provide valuable information such as (a) whether SMNR has acoustic effects on speech sounds, (b) which speech cues are masked by noise and if SMNR processing could reduce these masking effects, (c) which speech cues are unaffected after being processed by SMNR, and (d) how SMNR can be configured to enhance these cues to improve speech intelligibility. Thus, the aim of the second study was to determine whether SMNR has any acoustic effect on fricatives and affricates in the two most common languages in the world, Mandarin and English.

1.4 Mandarin Fricatives and Affricates

Other than in China, Mandarin is also used in 12 countries around the world such as Taiwan, Singapore, and Malaysia to name a few (Lewis, Gary, & Charles, 2015). As stated earlier, fricatives and affricates play an important role in Mandarin; about half of the Mandarin consonants are fricatives and affricates (Zhao & Li, 2009). There are a few distinctions between the fricatives and affricates of English and Mandarin languages, as shown in Table 1. First, fricatives and affricates in English can be voiced or voiceless (e.g. /s/ as in Sue and /z/ as in Zoo; /ʧ/ as in church and /ʤ/ as in juice) whereas all fricatives and affricates in Mandarin are voiceless. Second, there are alveolar /s, z/ and palato-alveolar /ʃ, ʒ/ fricatives in English whereas there are alveolar /s/, alveolo-palatal /ɕ/, and retroflex /ʂ/ fricatives in Mandarin. Third, there are only two affricates in English (i.e., palato-alveolar /ʧ, ʤ/ affricates) but there are six affricates in Mandarin (i.e., alveolar /ts, tsʰ/, alveolo-palatal /tɕ, tɕʰ/, and retroflex /tʂ, tʂʰ/ affricates). Fourth, the English affricates have no aspiration contrast but the Mandarin affricates have aspiration contrast. Fifth, Mandarin fricatives and affricates articulated at the same place of articulation exhibit a three-way contrast (e.g., /s/-/ʦ/-/ʦʰ/).

To date, most of the SMNR studies in the literature tested English speech materials and two studies tested Hindi speech materials (Marcoux et al., 2006; Turgeon et al., 2009), and limited studies have examined the effects of SMNR in hearing aids on the acoustics and perception of Mandarin fricatives and affricates in noise. For this reason, the effects of SMNR processing on the acoustics of Mandarin fricatives and affricates were examined in Study 2 while the perceptual effect was examined in Study 3. The Mandarin speech sounds of interest were the alveolar and retroflex fricatives /s, ʂ/ and affricates / ts, ts^h, tʂ, tʂ^h/. The Mandarin alveolo-palatal fricative /ɕ/ and affricates /tɕ, tɕ^h/ were excluded because they were considered to be allophonic with the alveolar and retroflex fricatives and affricates (Munro, 2008).

Table 1. Fricatives and affricates of English and Mandarin languages.

Manner of Articulation	Place of Articulation							
	Labio-dental	Dental	Alveolar	Alveolo-palatal	Retroflex	Palato-alveolar	Velar	Glottal
Fricative								
English	f v	Ɵ ð	s z			ʃ ʒ		h
Mandarin	f		s	ɕ	ʂ		x	
Affricate								
English						ʧ ʤ		
Mandarin			ʦ ʦ^h	ʨ ʨ^h	tʂ tʂ^h			

Note: English phonemes on the left and right of each place of articulation column denote voiceless sounds and voiced sounds, respectively. Mandarin phonemes on the top and bottom of each place of articulation column denote unaspirated and aspirated sounds, respectively.

1.5 Possible Acoustic Effects of SMNR on Fricatives

Acoustic measurements have shown that overall gain reduction as a result of SMNR processing varies across makes and models of hearing aids (Bentler & Chiou, 2006; Chung, 2004; Hoetink et al., 2009). Interestingly, some implementations of SMNR unintentionally reduce gain at mid and high frequencies (Hoetink et al., 2009). For MBNR, Schum (2003) hypothesized that gain reduction at high frequency channels may occur because signals filtered into the high frequency channels have lower amplitude modulations (as compared to signals filtered into the low frequency bands) and may be regarded as noise. This may suggest that high frequency speech sounds such as fricatives and affricates may be affected if noise is present. Gain reduction at the high frequency region is not desirable because many consonants, such as fricatives and affricates, have high frequency content and are important for speech intelligibility.

Preliminary studies from our lab suggested that some implementations of SMNR had measurable acoustic effects on the acoustics of fricatives /s, z/ (Chong & Jenstad, 2010; 2011). We quantified the acoustic effect of SMNR processing by using the short-term level change index. The short-term level change is calculated using a modified version of the equation for the short-term audibility (Pittman & Stelmachowicz, 2000; Stelmachowicz, 2001):

$$\text{Short term level change} = \frac{1}{20}\sum_{i=0}^{22}(fi\text{NR}off - fi\text{NR}on)\,W$$

In the short-term level change equation, (fiNRoff — fiNRon) is the level difference in each band between the SMNR-off and SMNR-on recordings; *W* is a constant (0.045) across 22 bands from 0.3 kHz to 8.0 kHz. The multiplier 1/20 was used to represent a 20 dB dynamic range, more suitable for calculating audibility of short segments of speech than the 30 dB dynamic range used to calculate the long-term speech spectrum (Pittman & Stelmachowicz, 2000; Stelmachowicz et al., 2001; Stelmachowicz et al., 2002). The level difference in any band that occurred more than 20 dB below the peak level in the measured speech segment was given a weighting of “0”.

Figure 2 and Figure 3 show the amount of short-term level change index for fricative noise /s/ and /z/ as a result of SMNR processing. These fricatives were spoken by two female and two male talkers. Three types of noise signals were used: (i) cafeteria noise, (ii) single-talker modulated International Collegium of Rehabilitative Audiology (ICRA) noise, and (iii) steady-state pink noise. The values of short-term level change index can range from -1 to +1. Positive bars indicate level increment, and negative bars indicate level reduction post SMNR processing. A value close to zero indicates no change in the short-term level whereas a value close to +1 or -1 indicates maximum level change as an effect of SMNR processing. The preliminary results show that SMNR caused relatively more level change for /s, z/ produced by female talkers in the pink noise condition than for male talkers. Whether these acoustic changes are significant for speech perception is unknown but the importance of acoustic characteristics on the perception of fricatives for individuals with normal hearing or hearing loss is well documented.

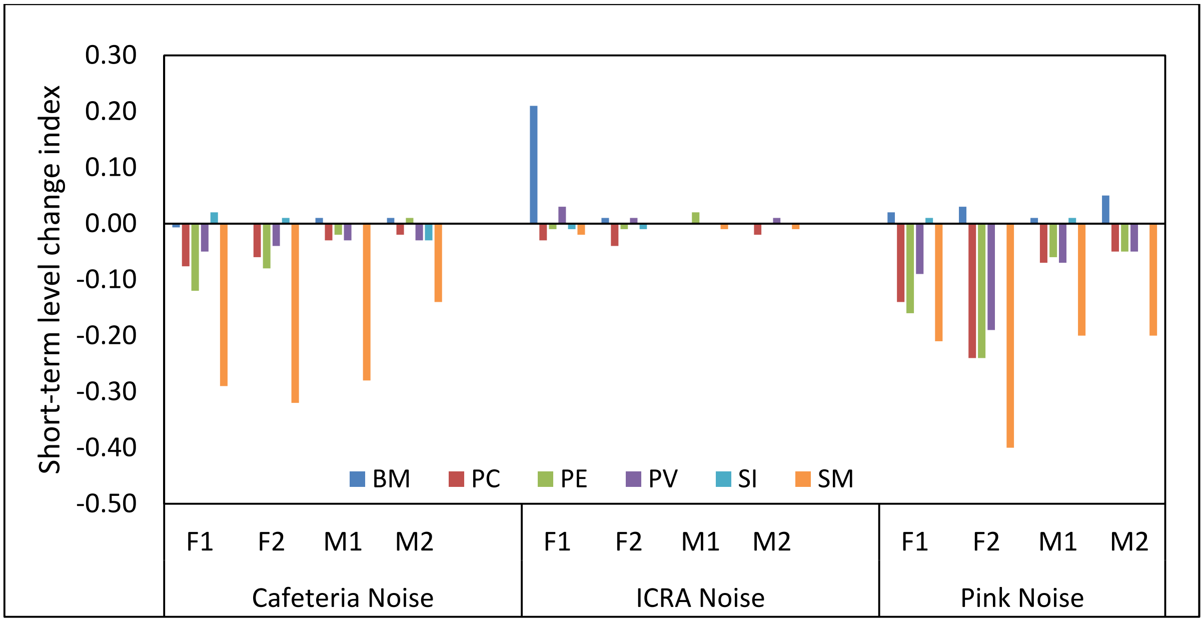


Figure 2. Short-term level change index caused by SMNR processing in six hearing aids (BM, PC, PE, PV, SI, and SM) for the phoneme /s/. This phoneme was spoken by two female (F1, F2) and two male (M1, M2) talkers. The values of the index can range from -1 to +1. A positive bar indicates level increment; a negative bar indicates level reduction.

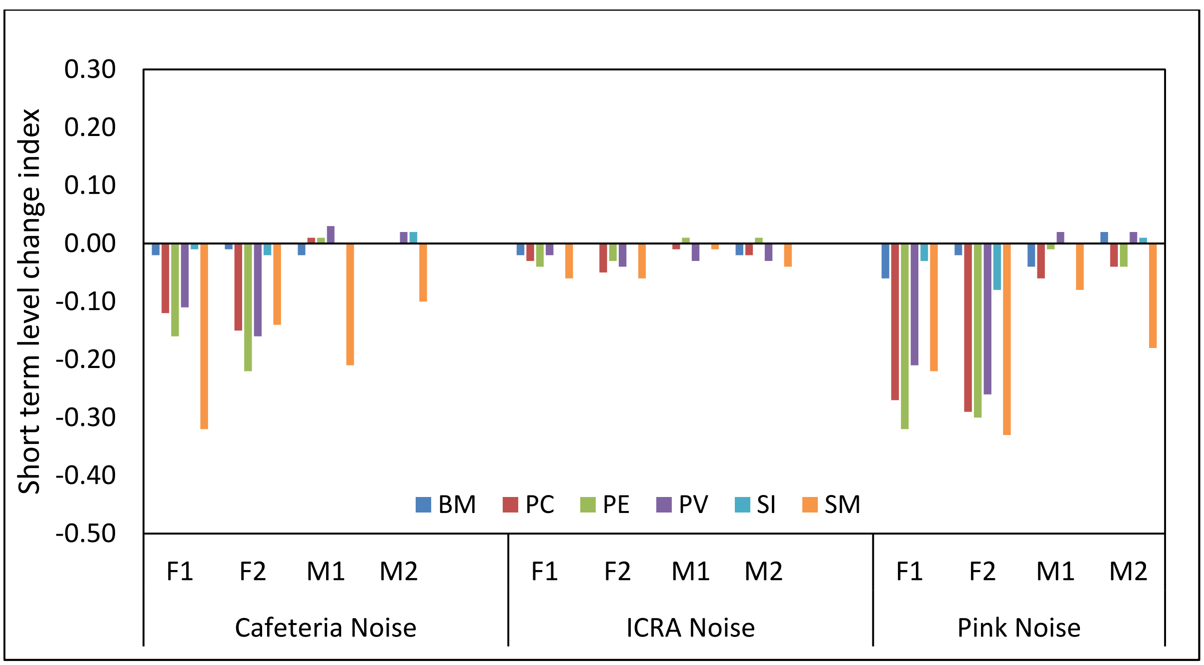


Figure 3. Short-term level change index caused by SMNR processing in six hearing aids (BM, PC, PE, PV, SI, and SM) for the phoneme /z/. This phoneme was spoken by two female (F1, F2) and two male (M1, M2) talkers. The values of the index can range from -1 to +1.A positive bar indicates level increment; a negative bar indicates level reduction.

1.6 The Inversion Technique: a Technique to Separate Speech and Noise Post Hearing-aid Processing

In order to examine the effect of SMNR processing on speech signals when noise is present, it is necessary to separate the speech signals from the SMNR-processed speech-plus-noise signals. Hagerman and Olofsson (2004) proposed a method—the "inversion technique"—that can be used to separate speech or noise signals from the speech-plus-noise signals post hearing-aid processing. Two speech-plus-noise signals (the noise waveform within one of the signals is 180^{o} out of phase) can be presented and recorded from a hearing aid. Subsequently, the recordings are summated or subtracted from each other to retrieve speech or noise, respectively. The retrieved acoustical waveforms can then be analyzed separately. It is important to ensure that the inversion technique will not affect the fidelity of the retrieved-speech signals prior to examining the acoustic effect of SMNR on these signals; this is to ensure that any acoustic change observed from the SMNR-processed speech signals can then be attributed to SMNR processing itself and not the interaction between SMNR processing and the inversion technique. Thus, the validity of the inversion technique as a tool for separating speech and noise signals recorded in a sound field setting was examined in the first study.

1.7 Summary of Studies

In this dissertation, three studies were conducted to examine the acoustic and perceptual effects of SMNR on the Mandarin and English fricatives and affricates. The purpose of the first study was to examine whether the inversion technique is a feasible tool for separating aided and unaided speech-plus-noise signals recorded in a sound field setting and whether its application affects the fidelity of retrieved-speech signals.The specific aims of this study were to quantify (i) the amount of error, (ii) changes to speech fidelity introduced by the inversion technique, and (iii) test-retest reliability of the inversion technique. The retrieved-silence method (Naylor & Johannesson, 2009) was used to achieve the specific aims of the study. Speech-plus-noise signals and speech-in-quiet signals were presented to, and recorded from, Knowles Electronic Manikin for Acoustic Research (KEMAR; G.R.A.S. Sound and Vibration, Denmark) in a double-walled sound-treated booth. The recordings were collected under two conditions: (i) without hearing aids on KEMAR (i.e., the unaided condition) and (ii) with one of the two commercially available hearing aids mounted on KEMAR (i.e., the aided condition). For the first specific aim, sets of speech-plus-noise recordings (one with both speech and noise inverted) were added to retrieve silence; the residual signals denote the amount of error from the inversion technique. For the second specific aim, retrieved-speech signals were subtracted from the corresponding speech-in-quiet recordings; the residual speech denotes the fidelity changes to speech from the inversion technique. For the third specific aim, retrieved-speech signals at Time-1 were subtracted from the retrieved-speech signals at Time-2 to retrieve silence. If the two time points were identical, then the waveforms should cancel out and leave only silence. I hypothesized that (i) the amount of attenuation achieved would be smaller when there were more sources of error, (ii) inversion technique would not affect the speech fidelity of the retrieved-speech signals, and (iii) the inversion technique would have good test-retest reliability. A detailed description of the first study can be found in Chapter 2 of this dissertation.

The aim of the second study was to determine the acoustic effects of SMNR on Mandarin fricatives and affricates with English fricatives and affricates as the control stimuli. The test materials consisted of sets of speech-plus-noise signals where the speech materials were VCV word strings and the noise was pink noise. Each word string contained one of the six Mandarin fricatives and affricates /s, ts, ts^{h}, ʂ, tʂ, tʂ^{h}/ or one of the five English fricatives and affricates /s, z, ʃ, tʃ, ʤ/ in each of the three vowel contexts /a, i, u/. These sets of speech-plus noise signals were presented to and recorded from one of the two hearing aids mounted on KEMAR under SMNR-on and SMNR-off conditions. The inversion technique was implemented on sets of speech-plus-noise recordings to separate speech and noise signals post hearing-aid processing. Acoustic measurements were performed on the VCV word-strings retrieved from the SMNR-on and SMNR-off recordings. The retrieved signals from both conditions were compared to determine if SMNR processing resulted in any acoustic changes on fricatives and affricates. The acoustic measurements included (i) amount of noise reduction, (ii) effective SNR change, (iii) the Hearing Aid Speech Quality Index (HASQI; Kates & Arehart, 2010), (iv) frication-noise intensity difference, (v) release-burst intensity difference, and (vi) spectral mean difference. I hypothesized that SMNR processing would cause (i) an increment to the amount of noise reduction when the noise level increased and input SNR decreased, (iii) an improvement on the effective SNR as the input SNR decreased, and (iii) acoustic changes in the amplitude and spectral domains. A detailed description of the second study can be found in Chapter 3 of this dissertation.

The aim of the third study was to examine the effects of SMNR on novel speech sound identification in noise by naïve listeners. Two groups of self-reported native English-speaking adults with normal-hearing participated in five sessions of identification training and testing. The control group was tested and trained with stimuli processed without SMNR; the experimental group was tested and trained with stimuli processed with SMNR. The stimuli consisted of VCV tokens embedded in pink noise; each VCV token consisted of a Mandarin retroflex consonant (e.g., either a fricative /ʂ/ or an affricate /tʂ^{h} or tʂ/ consonant) in one of the three vowel contexts /a, ɨ, u/. Mandarin retroflex fricatives and affricates, processed with and without SMNR, were used as the test stimuli because they are not within the phonemic inventory for English talkers. A classification procedure with three alternatives was used to test and train the subjects to identify the retroflex fricative and affricates. The subject's task was to label the tokens by selecting one of the three pictures displayed on a touch screen. The percentage of correct identification was recorded and transformed into Rationalized Arcsine Unit (RAU) in order to meet assumptions for statistical analysis (Studebaker, 1985). Reaction time was also recorded. A mixed-model analysis of variance was used to analyze the data with RAU or reaction time as the dependent variable, SMNR status and training voice as the between subject variables, and test sessions and test voice as the within subject variables. I hypothesized that (i) the RAU score would be different (the direction of the difference was uncertain) between the SMNR-on and SMNR-off groups and (ii) the RAU score would be lower in the baseline session than the post-test sessions, (iii) the reaction time would be different between the SMNR-on and SMNR-off groups, and (iv) the reaction time would be shorter with increasing number of test sessions. A detailed description of the third study can be found in Chapter 4 of this dissertation.

Chapter 2: Validation of the Inversion Technique

Most hearing aid processing, such as single-microphone noise reduction (SMNR), is nonlinear and will process signals that differ on spectral, temporal, and amplitude dimensions differently, such as speech-only, speech-plus-noise, or noise-only signals. To understand the effect of hearing aid processing on speech-plus-noise signals, one cannot extrapolate from the processing of speech-only signals. For example, in order to examine the effect of SMNR processing on speech signals when noise is present, it is necessary to extract the speech signals from the SMNR-processed speech-plus-noise signals. Hagerman and Olofsson (2004) proposed a method—the "inversion technique"—that can be used to separate speech or noise signals from the speech-plus-noise signals after hearing-aid processing. This technique requires pairs of speech-plus-noise signals that have identical physical parameters such as modulation and signal-to-noise ratio (SNR), but the noise waveform is phase inverted in one: (i) original-speech plus original-noise and (ii) original-speech plus phase-inverted noise. With these sound files, two recordings can be made with a hearing aid in a test box or in free-field. The recorded sound files (i) and (ii) are added digitally to retrieve speech. Conversely, the recorded sound files (i) and (ii) are subtracted digitally to retrieve noise. The inversion technique has been used in a number of studies to determine the effects of hearing-aid processing (e.g., compression and noise reduction) on the acoustics (e.g., effective SNR, amplitude envelope, and effective compression ratio) of the processed speech-plus-noise signals (Fredelake, Holube, Schlueter, & Hansen, 2012; Ghent, Nilsson, & Bray, 2007; Gustafson et al., 2014; Hagerman & Olofsson, 2004; Naylor & Johannesson, 2009; Pittman, 2011a; Pittman, 2011b; Souza, Jenstad, & Boike, 2006). However, there are variations on how the inversion technique was applied across studies.

First, some studies used software simulation to implement hearing-aid processing to the signals (Ellaham, Giguere, & Gueaieb, 2013; Fredelake et al., 2012; Souza et al., 2006) whereas other studies used signals that were recorded from hearing aids (Gustafson et al., 2014; Naylor & Johannesson, 2009; Pittman, 2011a, 2011b). Second, the recording environment and equipment were also different across studies. For example, Hagerman and Olofsson (2004) recorded the hearing aid output from an ear simulator (IEC 60711) in an anechoic room while others recorded hearing-aid output from a hearing-aid test chamber (Naylor & Johannesson, 2009; Pittman, 2011a; 2011b) or in a sound field setting (Gustafson et al., 2014). Recorded signals from hearing aids and simulations may differ in a few ways. First, the recorded signals may have more noise and distortions than the simulations. This can occur because nonlinear hearing aids and recording equipment may add noise and distortions to the recorded signals. Second, signals recorded with different hearing aids may have different, and variable, time delays. Third, signals recorded in a sound field setting have more reverberation than simulations. While recognizing that simulations have several advantages over recorded signals, recorded signals from hearing aids in sound field would be more representative of the conditions under which hearing aids are worn.

It is necessary to ensure that the inversion technique, and its particular application in the studies of this dissertation, is feasible prior to examining the acoustic effect of SMNR on these signals. Noise, distortions, time delays, and level differences may result in incomplete separation of the speech and noise signals post hearing-aid processing. Excessive residual noise (also known as error) that remains in the retrieved signals may affect subsequent acoustic analysis when these signals are used to examine the effect of SMNR processing on speech. Thus, the aim of the first study was to test the validity of the inversion technique as a tool for separating speech and noise signals recorded in sound field under the aided and unaided conditions. Validation can be achieved by quantifying the amount of residual noise (error) and changes to speech fidelity introduced by the inversion technique, as well as test-retest reliability.

Several methods can be used to validate an implementation of the inversion technique. Validation methods reported in the literature include quantifying the amount of residual noise from the inversion technique (Jenstad & Zakis, 2011; Johannesson, 2006) and the Hilbert transform method (Olofsson & Hansen, 2006). Hagerman and Olofsson (2004) used the Hilbert transform method, proposed by Olofsson and Hansen (2006), to determine the amount of distortion introduced into the signals by the recording system and hearing aids used in their study. The biggest limitation of the Hilbert transform method is that it does not test the implementation of the inversion technique itself directly, only the amount of distortion from the recording system and hearing aids. The Hilbert transform method was not used in this current study because it may not reflect whether the implementation of the inversion technique will impose any changes to the fidelity of the retrieved-speech signals.

Johannesson (2006) stated that the inversion technique can be validated by measuring the degree of signal attenuation when two speech-plus-noise recordings (one with both speech and noise phase-inverted) are summed to retrieve silence. The residual signal within a retrieved-silence file denotes the amount of error from the inversion technique. In one report, Johannesson (2006) stated that the amount of error (residual signal) was at least 15 dB lower than the retrieved-speech or retrieved-noise signals in their study and the typical amount of attenuation of the unwanted signals ranged between 20 to 30 dB. Using this method, Naylor and Johannesson (2009) found that the unwanted signals were attenuated by 20 dB in their study.

This retrieved-silence method was also used by Jenstad and Zakis (2011) as one of the measurements to validate their implementation of the inversion technique. Jenstad and Zakis (2011) also incorporated other measures to validate the inversion technique such as residual noise, speech fidelity, and test-retest reliability. The amount of residual noise was defined by Jenstad and Zakis (2011) as the amount of noise remaining in a retrieved-speech file relative to a recording of noise only. Speech fidelity was defined as the residual speech remaining when the retrieved-speech signals were subtracted from the speech-in-quiet recordings, quantified by the amount of attenuation in dB (Jenstad & Zakis, 2011). The amount of attenuation for residual noise and residual speech was approximately 20.0 to 23.6 dB in the unaided condition; the amount of attenuation for residual noise signal was 19.9 dB in the aided condition; the amount of attenuation for residual speech in the aided condition was not reported (Jenstad & Zakis, 2011). In the current study, the retrieved-silence and residual-speech measurements were incorporated as the analysis method due to their face validity. A criterion level of attenuation of 15 dB was used in the current study to indicate the acceptable amount of error and speech fidelity after signal separation (Johannesson, 2006).

In summary, the aim of the first study was to test the validity of the inversion technique as a tool for separating aided and unaided speech and noise signals recorded in a sound field setting. The specific aims of this study were to quantify (i) the amount of error introduced by the inversion technique, (ii) changes to speech fidelity introduced by the inversion technique, and (iii) test-retest reliability. I hypothesized that (i) the amount of attenuation achieved would be lower when there are more sources of error, (ii) the inversion technique would not affect the speech fidelity of the retrieved-speech signals, and (iii) the inversion technique would have good test-retest reliability. The test stimuli were fricatives and affricates of the two most common languages in the world (Mandarin and English). Mandarin and English were chosen because these two languages have the highest numbers of users in the world and fricatives and affricates are important in both of these languages. Fricatives and affricates in vowel-consonant-vowel (VCV) contexts were chosen to limit the contextual or linguistic cues when used for test stimuli in future behavioural studies.

2.1 Method

2.1.1 Recording of the Mandarin and English Speech Materials

2.1.1.1 Subjects

Two native Mandarin talkers and two native English talkers, one female and one male for each language, participated in the recording of speech materials.

2.1.1.2 Stimuli

The targeted Mandarin consonants were /s, ts, ts^h, ʂ, tʂ, tʂ^h/ in the VCV format. Three vowel contexts /a, ɨ, u/ were used. The alveolo-palatal fricative /ɕ/ and affricates /ʨ, ʨ^h/ were excluded because they are considered to be allophonic with the alveolar and retroflex fricatives and affricates in Mandarin language (Munro, 2008). The targeted English consonants were /s, z, ʃ, tʃ, ʤ/ in the /a, i, u/ vowel contexts. The Mandarin talkers were required to produce the VCV syllables with a high tone (tone 55) in a Mandarin carrier phrase “我要说___ (I will say___).” The English talkers were required to produce the English VCV syllables in a carrier phrase "I will say___."

2.1.1.3 Procedure

Prior to recording the speech stimuli, a list of VCV syllables written in Roman letters (or *PinYin* for Mandarin) was given to each talker for practice. During the recording sessions, five lists of VCV tokens written on paper were used to cue the talkers to produce the targeted VCV syllables. Each of the Mandarin lists contained all 18 VCV syllables (6 consonants X 3 vowel contexts) and each of the English lists contained all 15 VCV syllables (5 consonants X 3 vowel contexts). All VCV tokens were arranged in a randomized order in each list. All speech materials were recorded in a double-walled sound-treated booth. The speech materials were recorded monophonically using a microphone (Audio-Technica AT3035 30 series) routed through an UltraLite™ digital-to-analog pre-amplifier (Mark of the Unicorn, Inc.), using Praat (version 5.3.61; Boersma & Weenick, 2014) on an Apple iMac computer. The microphone was maintained approximately 10 cm away from a talker's lips during recording. Voice level was monitored using the VU meter within the Praat program. The recorded speech materials were digitized at 44100 Hz and stored in a computer as .wav files.

The Mandarin and English VCV syllables were validated by a native Mandarin-speaking and a native English-speaking adult, respectively, based on visual inspections of the acoustic waveforms in Praat (version 5.3.61; Boersma & Weenick, 2014) and auditory inspections through a pair of Sennheiser HD 265 linear headphones (Sennheiser Canada Inc.). The best VCV exemplar (i.e., naturalness, duration, and tone) for each consonant in each vowel context was selected and excised. Each excised VCV syllable was stored as a single .wav file in a computer and later was used to develop test stimuli for this study.

2.1.2 Validation of the Inversion Technique

2.1.2.1 Hearing Aids

Two commercial behind-the-ear hearing aids were used: Oticon Safari P300 and Phonak Solana M H20. Hereafter, these hearing aids were referred to as HA#1 and HA#2 respectively. The omnidirectional microphone mode was activated and all other advanced signal processing features such as SMNR, directional microphones, feedback cancelation, and frequency lowering were disabled for each hearing aid. All hearing aids were set to provide linear gain (the compression ratio in all processing channels was adjusted to 1:1) for a moderate to moderately-severe sensorineural hearing loss. Linear gain was chosen in order to control for compression as a confounding factor. The maximum power output for each hearing aid was set to maximum to control for output limiting as a confounding factor.

2.1.2.2 Stimuli

The speech signals consisted of six Mandarin VCV syllables and five English VCV syllables, spoken by a female and a male talker of each language. The VCV syllables with the same consonant in three vowel contexts (/a, i, u/), spoken by a single talker, were repeated five times and concatenated into a word string; in total, there were 15 VCV tokens in each word string. A total of 22 word strings (6 Mandarin syllables X 2 talkers and 5 English syllables X 2 talkers) was developed, as listed in Appendix C. Each word string (speech-only files) was then used to develop the speech-plus-noise sound files. All speech-only files were equalized in root-mean-squared (RMS) voltage prior to adding the noise. The noise consisted of steady-state pink noise generated from the Audacity program. This noise file was stored as "original noise" and its waveform was 180º phase-inverted and stored as the "inverted noise". The speech-only sound files and each pink noise ("original noise" and "inverted noise") were mixed at -10 dB SNR to +10 dB SNR in 5 dB steps to develop sets of speech-plus-noise sound files. The speech signal was kept constant at an arbitrary RMS voltage level while the RMS voltage level for the noise signal was varied to achieve the desired SNRs. Because the stimulus files were also used in the second study to examine the effect of SMNR on speech and noise, pink noise was added to the beginning of each speech-plus-noise file (see Figure

4) to ensure the activation of SMNR prior to the beginning of speech stimuli. On average, the duration of the pink noise was 22 seconds prior to the speech-plus-noise signal.

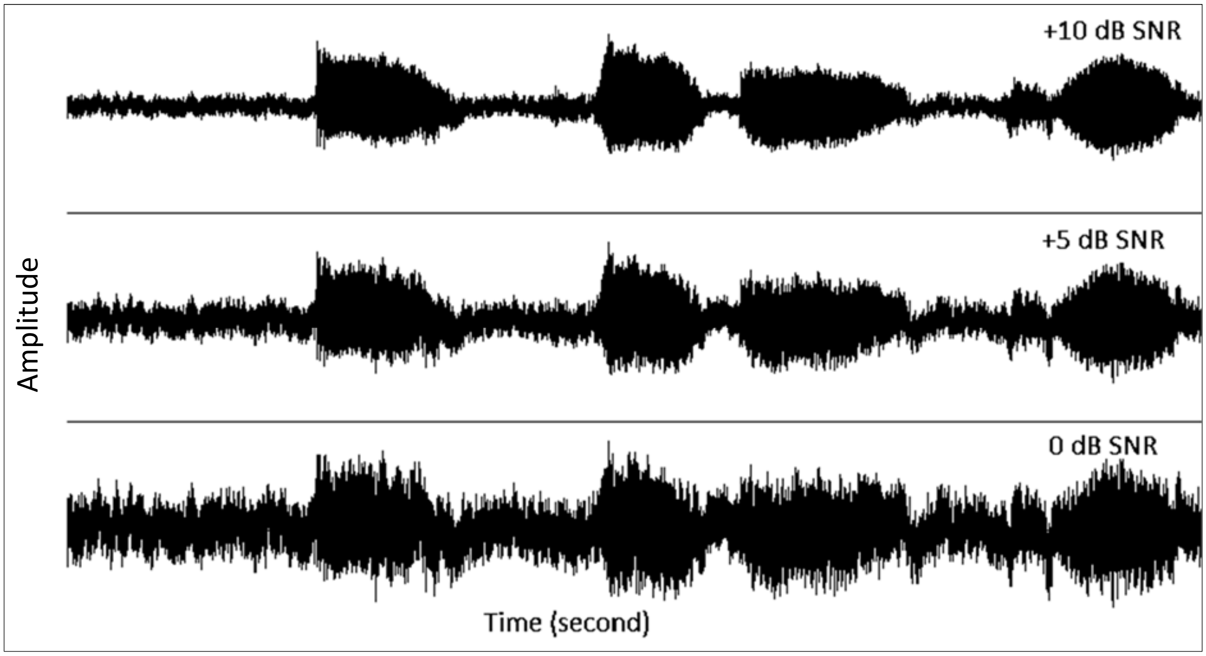


Figure 4. Examples of speech-plus-noise stimulus waveforms. Each waveform consists a noise-only portion (shown in the first few seconds) and a speech-plus-noise portion (shown in the rest of the waveform).

Two types of speech-plus-noise sound files were created, as listed in Appendix D: (i) original-speech plus original-noise, and (ii) original-speech plus phased-inverted noise. This set of two sound files allows the implementation of the inversion technique for speech signal extraction post hearing-aid processing. Half of type (i) and (ii) speech-plus-noise sound files were also phase-inverted and stored as the inverted files. The inverted sound files and non-inverted sound files enable the silence retrieval when summated. All sound files were developed using Audacity and Praat (version 5.3.61; Boersma & Weenick, 2014). These stimuli were digitized at 44.1 kHz and 32 bit resolution, and were stored in the computer as .wav files.

2.1.2.3 Procedure

All recordings were collected in a double-walled sound-treated booth. The recording system consisted of a computer, a Hammerfall DSP Multiface II Sound Card (RME, Germany), a Behringer Truth B2303A loudspeaker (MUSIC group IP Ltd., Philippine), a 45BA KEMAR (G.R.A.S. Sound and Vibration, Denmark); a microphone, a preamplifier, an earmould adapter, and a coupler (model RA0045 IEC 60711) embedded in KEMAR. The Audacity program was used as an interface to play and record the acoustic signals simultaneously. The equipment set up is illustrated in Appendix E.

The positioning of KEMAR in the double-walled sound-treated booth was determined using the sound-field calibration method suggested by Walker, Dillon, & Byrne, (1984). In the sound-field calibration method, the output of the loudspeaker for pure tones from 250 to 8000 Hz in octave frequencies was measured using a Larson-Davis 824 sound level meter, coupled with a half-inch microphone. The sound level meter was moved around systematically to determine the point in the sound booth where small variations in location would not make a large difference in the output of loudspeaker (i.e., less than 2 dB difference). Next, the head of the KEMAR was placed at the point.

Each of the sound files was presented to and recorded from KEMAR (i.e., the unaided condition) and from each of the two hearing aids mounted on KEMAR's left ear (i.e., the aided condition). This first set of recordings was marked as Time-1 recordings. There were 110 sets of speech-plus-noise recordings for each of the unaided and aided conditions; in total there were 330 sets of recordings for all conditions. Approximately 25% (n=83 sets) of these recordings were repeated for the purpose of test-retest reliability and were marked as Time-2 recordings. Stimuli were presented via the loudspeaker at the levels and input SNR conditions shown in Table 2. These levels were used to maintain the desired input SNRs and the average speech level at 65 dBA (Olsen, 1998; Pearson, Bennett, & Fidell, 1977). In addition, these levels were used to minimize the likelihood of reaching the output limiting of the hearing aids. Sound field calibration was conducted for each sound file to ensure consistent presentation level to the microphone of the hearing aid. The recordings were digitized at 44.1 kHz and 32-bit resolution and stored in a computer as .wav files for offline analysis.

Table 2. Sound field calibration level for test stimuli.

Input SNR condition	Level: noise only (dBA)	Level: speech and noise (dBA)
+10 dB SNR	54.7-55.0	63.5-66.0
+5 dB SNR	60.0-60.3	65.5-68.0
0 dB SNR	64.9-65.0	68.0-70.0
-5 dB SNR	69.8-69.9	70.0-71.0
-10 dB SNR	74.6-74.9	75.0-75.4

The inversion technique was used to extract silence ("retrieved-silence") or speech ("retrieved-speech") signals from the speech-plus-noise recordings. First, the RMS voltage of each acoustic file within a set of speech-plus-noise recordings was equalised because any amplitude difference will result in incomplete signal cancelation. Second, the waveforms were time-aligned using the Audacity program prior to mixing. Visual inspection on these waveforms was carried out: a few prominent peaks of the acoustic waveforms were selected and zoomed into the smallest sample. The time difference where these peaks occurred between the two acoustic waveforms indicated the amount of misalignment. Any misalignment was calculated and corrected by adding silence (e.g., 144 samples) at the beginning of one of the speech-plus-noise waveform. Auditory inspection was also carried out to listen to the retrieved-silence or retrieved-speech signals to determine that the unwanted signal was attenuated. Given that two merged files resulted in doubling of amplitude, 6 dB was subtracted from the overall level of each extracted signal.

The amount of error from the inversion technique was quantified by the amount of residual signal when pairs of original and phase-inverted speech-plus-noise recordings were summed to retrieve silence. The intensity levels of the retrieved-silence files and the unmerged original and phased-inverted speech-plus-noise recordings were computed using the Praat program. Next, the intensity level of each retrieved-silence file was compared to the levels of the unmerged original and phased-inverted speech-plus-noise recordings; the level difference between the retrieved-silence file and the unmerged files denotes the amount of attenuation.

There were a few assumptions in the speech fidelity measurement. First, residual noise in the retrieved-speech signals should remain low when the inversion technique is used to extract speech signals from the speech-plus-noise recordings. Second, when the residual noise in the retrieved-speech signals is small, a retrieved-speech waveform should resemble a speech-only waveform recorded with a linear system. Third, assuming that the residual noise in the retrieved-speech waveform and the speech-only waveform is small, subtracting one of the signals from the other will result in silence. Hence, residual noise in the retrieved-speech signals can be estimated by subtracting the retrieved-speech signals from the speech-only signals recorded in the same recording conditions and measuring the signals (residual-speech) that remains. Moreover, Johannesson (2006) stated that this residual noise estimate has good face validity for examining the inversion technique. Therefore, speech fidelity of the retrieved-speech signals was quantified by the amount of attenuation obtained when a retrieved-speech signal was subtracted from a speech-only signal to retrieve residual-speech signals. The intensity levels of the residual-speech signals were computed using the Praat program and were compared to the levels of the unmerged files.

Test-retest reliability was quantified by the amount of attenuation obtained when a retrieved-speech signal from Test-1 recordings was subtracted from a corresponding retrieved-speech signal from Test-2 recordings to retrieve silence. If the retrieved-speech signals from two tests were identical, then the waveforms should cancel out and leave only silence. Recall that 83 sets of speech-plus-noise signals were recorded twice; recordings from the first test were labeled as Test-1 recordings and recordings from the second test were labeled as Test-2 recordings. Speech was extracted from each set of speech-plus-noise recordings using the inversion technique. Retrieved-speech files from Time-1 were time-aligned and subtracted from the corresponding retrieved-speech files from Time-2 to retrieved silence. Next, the intensity level of each retrieved-silence file was compared to the levels of the Time-1 and Time-2 retrieved-speech files, the difference denotes the amount of attenuation for test-retest reliability.

According to Johannesson (2006), the level of the retrieved-silence files should be at least 15 dB lower than the unmerged files as an indication of successful signal cancelation from the inversion technique. Therefore, a 15 dB attenuation criterion was used an indication of acceptable (i) amount of error, (ii) speech fidelity, and (iii) test-retest with the inversion technique.

2.2 Results

2.2.1 Aim (i): Amount of Error

The average amount of attenuation for retrieved-silence signals in three recording conditions (i.e., unaided, aided HA#1, and aided HA#2) for all input SNRs is shown in Figure 5. The dashed line represents the criterion value (i.e., 15 dB attenuation). Any condition that has attenuation of 15 dB or more is considered to be a successful signal cancelation by using the inversion technique. The white bars represent the unaided recording condition, the black bars represent the HA#1 recording condition, and the grey bars represent the HA#2 recording condition. Each bar represents the amount of attenuation averaged across 22 retrieved-silence files. Each error bar represents ±1 standard deviation (SD) of each condition. The average amount of attenuation ranged from 31.1 to 38.0 dB for the unaided condition, 31.8 to 34.6 dB for HA#1, and 22.2 to 24.2 dB for HA#2. Figure 5 shows that the amount of attenuation for all conditions exceeded the 15 dB attenuation criterion.

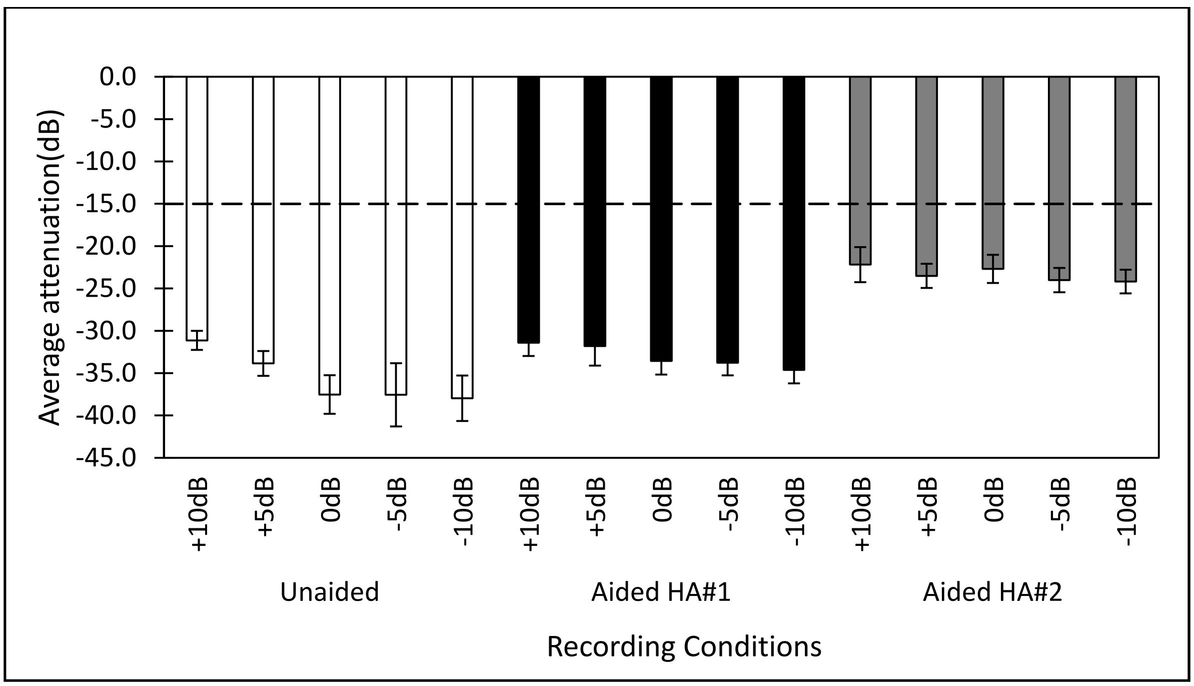


Figure 5. Amount of attenuation for retrieved-silence files. The dashed line indicates the target amount of attenuation (-15 dB), where more attenuation is better (less error). The white bars represent the unaided recording condition, the black bars represent the HA#1 recording condition, and the grey bars represent the HA#2 recording condition. The error bars represent ± 1 SD of each condition.

2.2.2 Aim (ii): Speech Fidelity

The average amount of attenuation for the residual-speech signals in three recording conditions for all input SNRs is displayed in Figure 6. The white bars represent the unaided recording condition, the black bars represent the HA#1 recording condition, and the grey bars represent the HA#2 recording condition. Each bar represents the amount of attenuation averaged across 22 residual-speech files. Each error bar represents ±1 SD of each condition. The average amount of attenuation in the unaided (range: -14.6 to -33.6 dB) and aided HA#1 conditions (range: -12.69 to -32.3 dB attenuation) at all input SNRs exceeded the 15 dB attenuation criterion. The average amount of attenuation for residual-speech files of aided HA#2 conditions at positive SNRs (range: -11.5 to 22.9 dB attenuation) exceeded the attenuation criterion. The average amount of attenuation for residual-speech files in the aided HA#2 was -14.8 dB (range: -11.4 to -24.1 dB) at 0 dB SNR, -12.3 dB (range: -9.4 to -14.8 dB) at the-5 dB SNR, and 9.6 dB (range: -7.2 to -13.5 dB) at -10 dB SNR. The amount of attenuation for HA#2 at negative SNRs did not meet the 15 dB attenuation criterion.

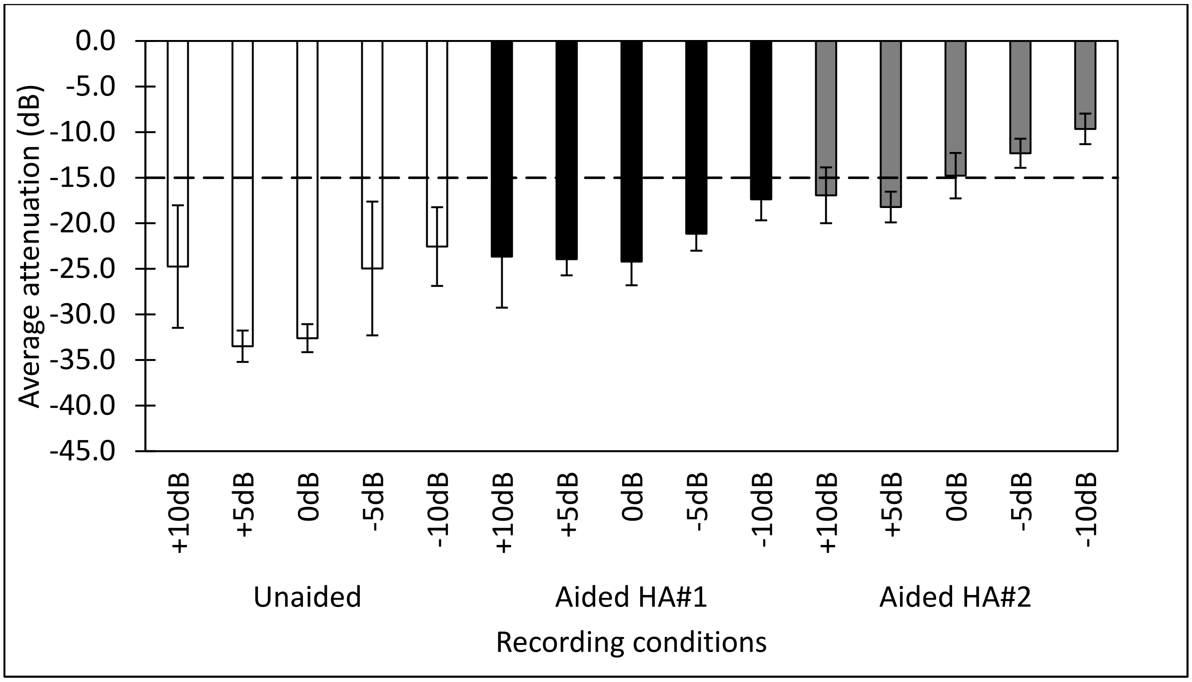


Figure 6. Amount of attenuation for residual-speech signals. The dashed line indicates the target amount of attenuation (-15 dB), where more attenuation is better fidelity. The white bars represent the unaided recording condition, the black bars represent the HA#1 recording condition, and the grey bars represent the HA#2 recording condition. The error bars represent ± 1 SD of each condition.

2.2.3 Aim (iii): Test-retest reliability

Figure 7 shows the average amount of attenuation for test-retest reliability. The white bars represent the unaided recording condition, the black bars represent the HA#1 recording condition, and the grey bars represent the HA#2 recording condition. Each bar represents the amount of attenuation averaged across 22 residual-speech files. Each error bar represents ±1 SD. The average amount of attenuation for the unaided (range: -25.1 to -35.4 dB) and aided HA#1 conditions (range: -19.6 to -34.7 dB) at all input SNRs exceeded the 15 dB attenuation criterion. The average amount of attenuation for the aided HA#2 condition at four SNRs (+10, +5, 0, and -5 dB SNR) exceeded the attenuation criterion (range: -13.0 to -24.3 dB). The average amount of attenuation for the -10 dB SNR condition for HA#2 did not meet the 15 dB criterion (mean: -12.3 dB; range: -11.3 to -13.6 dB).

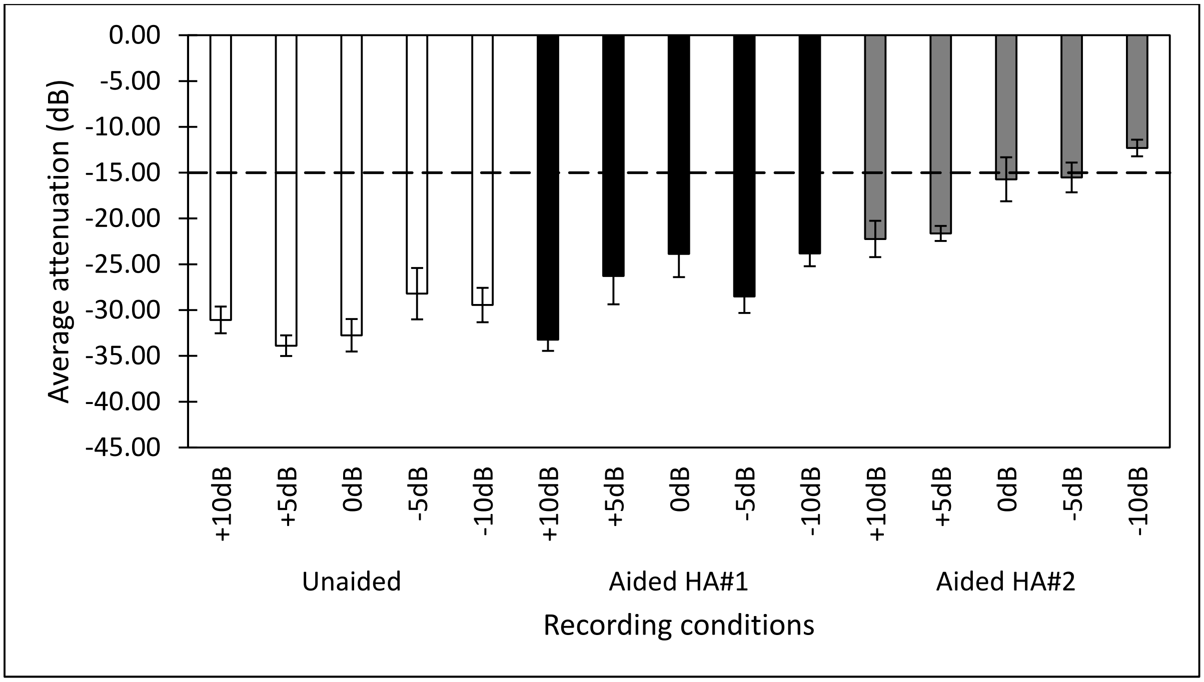


Figure 7. Amount of attenuation for test-retest reliability. The dashed line indicates the target amount of attenuation (-15 dB), where more attenuation is better test-retest reliability. The white bars represent the unaided recording condition, the black bars represent the HA#1 recording condition, and the grey bars represent the HA#2 recording condition. The error bars represent ± 1 SD of each condition.

2.3 Discussion and Conclusions

The main objective of this study was to validate this implementation of the inversion technique as a tool for separating speech and noise signals (e.g., unaided and aided) recorded in a sound field setting because this method can be very useful for examining the effects of advanced features in hearing aids on speech, such as SMNR processing. In the inversion technique, two speech-plus-noise recordings (one with noise 180^{o} phase inverted) are summed to retrieve speech and subtracted to retrieve noise. Any difference between one recording and another will reduce the ability to achieve perfect cancellation of the unwanted signal. Two of the crucial factors that could affect the success of signal cancelation by the inversion technique include (i) amplitude and (ii) time alignment of each set of speech-plus-noise recordings. Any amplitude difference or temporal misalignment of a set of recordings will result in incomplete signal cancelation. Although these factors can be controlled carefully, there will still be differences between the two recordings due to random sources of error such as random errors with the hearing aids. Hence, the first specific aim of this study was to quantify the errors due to any random sources while controlling other factors that could be controlled.

In this study, the stimulus level for all input SNRs was carefully controlled during two stages: stimulus development and stimulus presentation. In stimulus development, all stimulus files for an SNR condition were equalized in terms of RMS voltage level. In stimulus presentation, sound field calibration was conducted daily to ensure consistent presentation levels for all stimulus files at each SNR condition. In order to minimize the time alignment factor, visual inspections were carried out for waveforms of each set of speech-plus-noise recordings and any misalignment was corrected by adding silence at the beginning of one of the speech-plus-noise waveform. Subsequently, the inversion technique was applied to retrieve speech or silence once the temporal misalignments had been corrected as much as possible.

However, it is important to acknowledge that there will be residual signal due to sources of error such as noise from the recording environment (e.g., reverberation from the sound field setting), equipment noise, hearing-aid noise, and nonlinear hearing-aid processing. In addition, I hypothesized that the amount of attenuation achieved would be less when there are more sources of error. In this study, the amount of error was quantified by the amount of attenuation when two speech-plus-noise signals (e.g., one with original phase and another with inverted phase) were mixed to retrieve silence. The results showed that all retrieved-silence files met the 15 dB attenuation criterion, with typical amount of attenuation between 22 to 38 dB. These results were in line with the values of 20 to 30 dB attenuation reported by Johannesson (2006). The amount of attenuation for aided conditions in this current study was also comparable to the values reported by Wu & Bentler (2007). Wu and Bentler (2007) used a modified method of the inversion technique to test adaptive directional microphones in free field and they named this modified technique as the signal-cancelation method. In order to measure the residual noise of the test set up, white noise was presented twice (once with phase inverted) and recorded from hearing aids mounted to a 2cc coupler in free-field. The inversion technique was then applied to retrieve silence. Wu & Bentler (2007) reported that the amount of attenuation for the retrieved silence files was approximately 35 to 40 dB. The result of the current study suggests that the inversion technique is a feasible tool to cancel unwanted signals within sets of speech-plus-noise files recorded in a sound field setting.

The results of this current study also showed that the average amount of signal cancelation was the highest for the unaided recording condition (range: 31.1 to 38.0 dB attenuation), followed by HA#1 (range: 31.8 to 34.6 dB attenuation) and HA#2 (range: 22.2 to 24.2 dB) recording conditions. These results followed the trend found in Jenstad and Zakis (2011) where the amount of attenuation for aided recordings (i.e., 19.0 dB attenuation) was lower than the amount of attenuation in the unaided recordings (i.e., 23.6 dB attenuation). One possible explanation for relatively higher amount of attenuation in the unaided condition as compared to the aided conditions was that there were fewer sources of error (e.g., hearing-aid noise and nonlinear hearing-aid processing) in the unaided condition. For example, when testing the inversion technique using computer simulation, Ellaham et al. (2013) found that there was no error in the retrieved silence files for the unaided condition and the inversion technique recovered the exact signals. In short, the results of the retrieved-silence measurement were consistent with the previous findings: the amount of error was higher in the aided condition.

The second objective of this study was to quantify changes to speech fidelity introduced by the inversion technique. Fidelity of retrieved-speech signals was quantified by the amount of residual speech when speech-in-quiet waveforms were subtracted from the retrieved-speech signals. A 15 dB attenuation criterion (i.e., the residual-speech file was at least 15 dB below the recorded speech levels) was used in this measurement because Johannesson (2006) reported that the error estimation in his study was at least 15 dB lower than the retrieved-speech or retrieved-noise signals and the typical amount of attenuation was between 20 to 30 dB when the recordings were made in a test box. The results showed that the amount of attenuation for the residual-speech signals for the unaided and the aided HA#1 recording conditions met the 15 dB criterion for all input SNR conditions. These results were comparable to the amount of attenuation reported by Jenstad and Zakis (2011) in the unaided condition (i.e., 20 dB attenuation). The results also showed that the residual-speech signals for the aided HA#2 at 0 dB SNR marginally meet the criterion. However, the -5 dB and -10 dB SNR input conditions did not meet the attenuation criterion, suggesting that the retrieved-speech signals under these two conditions had the poorest fidelity relative to the other conditions tested in this study. Based on the criterion used in this current study, the results of the speech fidelity measurement indicated that the inversion technique had little effect on the fidelity of the retrieved-speech signal obtained under the 0 to +15 dB SNR conditions. However, fidelity was affected in aided recordings in poor SNR conditions (-10 and -5 dB SNR). Thus, it is recommended that verification of speech fidelity to be conducted for free-field aided recordings, particularly the ones recorded under a negative SNR condition, before proceeding with acoustic measurements.

The third objective was to examine the test-retest reliability of the inversion technique. Approximately 25% (i.e., 83 sets) of the speech-plus-noise recordings were repeated. Generally, an amount of attenuation of 15 dB or more suggests good test-retest reliability. Overall results showed that the unaided and aided HA#1 conditions at all input SNRs had good test-retest reliability. The aided HA#2 condition had good test-retest reliability at four SNR conditions (+10, +5, 0, and -5 dB SNR) but the -10 dB SNR condition did not meet the 15 dB criterion.

The variability of the results between the two hearing aids indicated that there could be a range of results across different measurements or makes and models of hearing aids. The differences in the results for the two hearing aids could be due to the differences in digital signal processing strategy or levels of technology (e.g., HA#1 has fewer processing channels thanHA#2) incorporated in each of the hearing aids by different manufacturers. Time delay in each processing channel or group delay (e.g., low-frequency channels or high-frequency channels) may introduce different amount of error in the recordings obtained with each hearing aid. However, with only one unit of hearing aid from each of the two manufacturers, it is difficult to rule out whether the particular hearing aids tested in the current study were outliers. Therefore, future research should test more hearing aids to answer these questions.

It is acknowledged that there were a few limitations in this current study. First, the residual-speech measurement might overestimate the effect of inversion technique on fidelity of the retrieved-speech signals because the inversion technique was performed twice in the measurement (i.e., first to retrieve speech signals from the speech-plus-noise signals and second to obtain residual speech). This limitation also applies to the test-retest reliability measurement where the inversion technique was performed three times throughout the procedure. Therefore, the residual-speech signals and the residual signal in the retrieved-silence files for test-retest reliability could have additional error introduced by the inversion technique. In other words, the actual error could be smaller than the values reported. Second, only two commercial hearing aids were used and it is difficult to explain if the variability of the results between the two hearing aids was due to the hearing-aid processing algorithm implemented by each manufacturer. For this reason, future research should test more hearing aids with same and different models across manufacturers.

Despite these limitations, overall results of this current study showed that the inversion technique (Hagerman & Olofsson, 2004) is a feasible and reliable tool for extracting speech and noise from speech-plus-noise recordings obtained in a sound field setting under input SNR conditions between 0 to +15 dB SNR. These results supported two of the hypotheses: (i) the amount of attenuation was greater when there were fewer sources of error and (ii) the inversion technique had good test-retest reliability. However, the hypothesis regarding the effect of the inversion technique on speech fidelity was only partially fulfilled because there was some variability in the findings between the two hearing aids tested in this study. Generally, fidelity of the retrieved-speech signals (aided and unaided) was not affected between 0 to +15 dB SNR but was affected when the recordings were made under poor SNR conditions (i.e., -10 and -5 dB SNR). Therefore, it is recommended that verification of speech fidelity to be conducted for aided recordings obtained in free field, particularly the ones recorded under a negative SNR condition, before proceeding with acoustic measurements.

Chapter 3: Acoustic Effects of Single-Microphone Noise Reduction on Mandarin and English Fricatives and Affricates

Single-microphone noise reduction (SMNR), or digital noise reduction, is used in hearing aids to suppress background noise. SMNR has been shown to ease listening comfort and reduce listening effort (Gustafson et al., 2014; Mueller et al., 2006; Palmer et al., 2006; Ricketts & Hornsby, 2005). Studies examining the effect of SMNR on speech intelligibility found no positive or negative effect among adults and school-age children (Alcantara et al., 2003; Boymans & Dreschler, 2000; Nordrum et al., 2006; Sarampalis et al., 2009; Stelmachowicz et al., 2010; Zakis et al., 2009). However, there are still questions about whether SMNR will affect the acoustics of noise-like speech sounds, such as fricatives and affricates when background noise is present (Chong & Jenstad, 2010, 2011).

Very few studies have reported spectrographic analysis on SMNR-processed speech sounds (Chong & Jenstad, 2010, 2011; Stelmachowicz et al., 2010). Spectrographic analysis of SMNR-processed and unprocessed nonsense words (/aba/ and /ava/), carried out by Stelmachowicz et al. (2010), revealed that SMNR processing reduced the magnitude of background noise but the frication noise portion of the consonant /v/ and the silent period prior to the release burst for the consonant /b/ were still obscured by background noise. In addition, identification of the consonant /b/ at 0 dB SNR, by children aged 5 to 7 years old, had an error rate of 87.5% in the SMNR-on condition as opposed to only 37.5% in the SMNR-off condition (Stelmachowicz et al., 2010). In other studies, Chong and Jenstad (2010, 2011) found that some implementations of SMNR have measurable acoustic effects on the fricative consonants /s, z/ and that the effects were greater for fricatives spoken by female talkers. Most of the SMNR studies in the literature used English speech materials and very few used speech materials from other languages, such as Hindi dental and retroflex stops (Marcoux et al., 2006; Turgeon et al., 2009). To date, no study has investigated the effects of SMNR on the acoustics of Mandarin fricative and affricate consonants in noise despite Mandarin being the language with the highest number of talkers in the world (Lewis, Gary, & Charles, 2015).

Due to the possible acoustic effects of SMNR on fricatives, there is a potential that individuals with hearing loss may receive inconsistent exposure to these phonemes. Studies have also shown that early-identified infants with hearing loss have delayed development in consonant production (particularly fricatives and affricates) and vocabulary production despite having access to current hearing aid technologies (Moeller et al., 2007a; 2007b). For example, one study found that Mandarin-speaking children (3 to 5 years old) who used hearing aids had lower consonant recognition scores as compared to age-matched normal-hearing children, particularly for three categories of consonant contrasts: (i) fricative vs. non-fricative; (ii) same place but different manner of articulation (e.g., fricative vs. affricate, fricative vs. stop, affricate vs. stop); and (iii) retroflex vs. non-retroflex (Liu, Zhou, Berger, Huang, & Xu, 2013). Other studies have shown that not having consistent access to English fricatives can lead to poor speech and language performance in English, particularly for children with hearing loss (Stelmachowicz et al., 2004). For these reasons, there is a need to investigate whether SMNR has an acoustic effect on noise-like speech sounds, such as affricates and fricatives in the English and Mandarin languages. Therefore, the aim of this study is to determine the acoustic effects of SMNR on Mandarin fricatives and affricates with the English fricatives and affricates as the control stimuli.

3.1 Acoustic Measurements

Few studies have reported acoustic measurements of SMNR-processed signal (Chong & Jenstad, 2010, 2011; Stelmachowicz et al., 2010) due to the difficulty of separating the mixed signals (e.g., speech in background noise) in the output of hearing aids. The studies that have examined the acoustic effects of SMNR processing focused on measurements such as (i) amount of gain reduction, (ii) amount of signal-to-noise ratio (SNR) change, and (iii) coherence.

To quantify amount of gain reduction, Bentler and Chiou (2006) measured and compared the output of hearing aids when SMNR was turned on and off. The output difference between the SMNR-on and SMNR-off conditions was taken as the amount of gain reduction. Bentler and Chiou (2006) found that the SMNR tested in their study did not reduce gain for speech or speech-like signals (i.e., single-talker modulated ICRA noise) presented in quiet. However, gain was reduced for noise stimuli (e.g., white noise and babble noise) and the amount of reduction varied by hearing aid, input level, and SNR. The amount of gain reduction also increased when SNR for speech-plus-noise signals decreased. In another study, Hoetink et al. (2009) examined the overall gain reduction as a function of frequency across 12 hearing aids with modulation-based noise reduction (MBNR) and showed that the hearing aids exhibited different amounts of maximum gain reduction, varying from 3 to 15 dB. Among the hearing aids tested, three had broadband gain reduction, one had low frequency gain reduction, and two had low to mid frequency gain reduction. Hoetink et al. (2009) also found that three hearing aids had low- and high-frequency gain reduction, while one hearing aid had mid- and high-frequency gain reduction. High-frequency gain reduction is not desirable as this may affect the audibility of consonants with high frequency content, such as fricatives and affricates. Although the gain reduction measurement provides information regarding the strength of noise reduction processing (e.g., amplitude change), it does not provide information on whether SMNR affects SNR and spectral properties of speech signals.

The second type of acoustic measurement is the estimation of SNR post SMNR-processing. This can be examined using the inversion technique to separate speech and noise from the speech-plus-noise output of any hearing aid (Hagerman & Olofsson, 2004). The SNR of the processed signals can be estimated from the isolated speech and noise signals. Hagerman and Olofsson (2004) showed that a small amount of SNR improvement (about 1 to 2 dB) was observed in four of the five hearing aids tested, and one of the hearing aids did not show any SNR change. In addition, the amount of SNR change varied by hearing aid, signal type, and input SNR (Hagerman & Olofsson, 2004). Using the same inversion technique, Pittman (2011a) examined SNR change provided by SMNR in hearing aids worn by 26 school-age children. The output of the hearing aids was recorded in SMNR-on and SMNR-off conditions in a test box. The SNR for each test condition was calculated by comparing the long-term average spectra of the isolated speech and noise signals and was given the term “effective SNR.” Pittman (2011a) reported that the effective SNRs ranged from -2.6 to -1.7 dB for the SMNR-off condition and ranged from -0.3 to 0.6 dB for the SMNR-on condition, an average SNR improvement of 2.3 dB. In another study, Pittman (2011b) reported that the average SNR improvement provided by the SMNR processing was 1.9 dB. Gustafson et al. (2014) also used the inversion technique to measure SNR change of two different makes and models of hearing aids and found that each hearing aid had a different amount of SNR change. One hearing aid had 1.46 and 3.79 dB SNR change at +5 and 0 dB input SNR, respectively; another hearing aid had 6.68 and 7.00 dB SNR change at +5 and 0 dB input SNR, respectively. Although the effective SNR measurement provides information regarding how SMNR changes the overall level of speech and noise, it does not provide information on whether SMNR affects spectral and temporal properties of speech signals.

The third type of acoustic measurement, the Coherence Speech Intelligibility Index (CSII; Kates & Arehart, 2005), has been used to examine the effect of SMNR processing. This coherence-based measurement is proposed by Kates and Arehart (2005) and the basis of CSII calculation is the conventional speech intelligibility index (SII; ANSI S3.5-1997). In the conventional SII calculation, the SNR in each frequency band is calculated and weighted based on the importance function of each frequency band. The frequency importance function is available for critical bands, one-third octave bands, and octave bands (SII; ANSI S3.5-1997). The sum of the weighted SNRs will yield a single number known as the SII. The SII can reliably estimate speech intelligibility under steady-state noise condition, accounting for auditory thresholds and upward spread of masking. However, speech intelligibility in fluctuating noise may not be predicted reliably using the conventional SII (Rhebergen & Versfeld, 2005). Furthermore, various types of nonlinear distortions are not accounted for in the conventional SII (Kates & Arehart, 2005). Thus, Kates and Arehart (2005) proposed to use a coherence function, known as the magnitude-squared coherence, in the CSII calculation. In CSII measurement, the conventional SNR estimation is replaced by the signal-to-distortion ratio estimation in each frequency band. They found that a three-level CSII measurement was more accurate in predicting speech intelligibility than a single-level CSII measurement. In the three level CSII calculation, the envelope of each signal is divided into three amplitude regions and the CSII calculation is performed for each amplitude region. These amplitude regions are (a) low: 10 to 30 dB below overall root-mean-squared (RMS) level, (b) mid: 0 to 10 dB below overall RMS level, and (c) high: at or above overall RMS level. The three-level CSII has been used as an objective estimation of signal intelligibility post hearing-aid processing and is available as a MATLAB function written by Kates (2010). A value of 0 in the CSII measurement indicates that the test signals are less intelligible than the reference signal; a value of 1 indicates that the test signals have similar intelligibility as the reference signal.

Gustafson et al. (2014) used the coherence-based measurement to examine whether SMNR imposed any spectral distortion to speech-plus-noise signals. They compared the spectra of input signals with the spectra of output signals from two hearing aids under two conditions, SMNR-on and SMNR-off. A difference in the coherence values between the two SMNR conditions was taken as an indication of distortion introduced by SMNR processing. Gustafson et al. (2014) showed that the coherence value of speech-plus-noise signals processed with and without SMNR for both hearing aids remained high (e.g., above 0.80). However, the coherence value for the hearing aid with the strongest SMNR effect (e.g., approximately 7 dB SNR improvement) was lower than the value for the hearing aid with the lower SMNR effect (e.g., approximately 3.8 dB SNR improvement). This suggests that stronger SMNR will result in more spectral distortion in the processed signals.

Another measurement that has been used in hearing-aid research is the Hearing Aid Speech Quality Index (HASQI; Kates & Arehart, 2010). The HASQI measurement is developed to estimate the effects of noise, nonlinear processing (e.g., compression and noise reduction algorithm) and linear filtering (e.g., high-pass or low-pass filtering effects of earmolds) on speech quality ratings. In the HASQI measurement, a degraded signal, such as that processed by a hearing aid, is compared to a clean signal in two indices: (a) the noise and nonlinear processing index and (b) the linear filtering index. The noise and nonlinear processing index compares the time-frequency envelope modulation of the clean signal and the distorted signal while ignoring the spectral differences between the two; both signals are time-aligned prior to estimating the envelope differences. The linear filtering index compares the long-term spectra of the clean signal and the degraded signal while ignoring the envelope differences between the two signals. Furthermore, the two signals are scaled to match in overall intensity prior to comparing the long-term spectra of the two signals. The third index, known as the combined index, is derived by multiplying the two aforementioned indices. The score of each index ranges from 0 (two signals being compared are different) to 1 (two signals being compared are similar). Falk et al. (2015) suggested that HASQI could be used as a measurement tool for examining the effect of speech enhancement processing such as SMNR on speech quality estimation.

Amount of gain reduction, effective SNR, and HASQI were selected as the measurement tools in this study to compare the acoustic properties of retrieved-speech signals of SMNR-on and SMNR-off conditions. The HASQI measurement was chosen over the CSII measurement because (i) noise and nonlinear processing index in the HASQI measurement provides temporal information and (ii) the linear filtering index in the HASQI measurement provides spectral information whereas the CSII provides an estimation of signal intelligibility but not the temporal and spectral properties. Temporal information is important for perception of manner of articulation (e.g., stop /b/ vs. fricative /f/) and of the supra-segmental features of speech, whereas spectral information is important for perception of place of articulation (e.g., alveolar fricative /s/ vs. post-alveolar fricative /ʃ/) for speech sounds. Recall that there are three indices for the HASQI measurement: (i) noise and nonlinear index, (ii) linear index, and (iii) combined index. Because the combined index is a multiplication of the nonlinear index and the linear index, it will be biased towards the index with a lower score. Therefore, only the results of the (i) noise and nonlinear index and the (ii) linear filtering index were reported in this study.

3.1.1 Spectrographic Measurements of Fricatives and Affricates

3.1.1.1 Fricatives

Fricatives are produced when air from the lungs goes through a narrow constriction (e.g., by bringing the tongue close to the alveolar ridge) in the oral cavity resulting in turbulent airflow. Sibilant fricatives in English include the alveolar /s, z/ and palato-alveolar /ʃ, ʒ / fricatives. Sibilant fricatives in Mandarin include the alveolar /s/, alveolo-palatal /ɕ/, and retroflex /ʂ/ fricatives. The Mandarin retroflex fricative /ʂ/ and the English palato-alveolar fricatives /ʃ, ʒ/ can be grouped together as post-alveolar fricatives. Although the target surface for these fricatives is the same, the part of the tongue that produces the constriction for the retroflex and palato-alveolar fricatives is different. The retroflex fricative is produced by bringing the tip of the tongue to the back of the alveolar ridge while the palato-alveolar fricatives are produced by bringing the blade of the tongue to the back of the alveolar ridge. Thus, the retroflex fricative is categorized as the apical post-alveolar fricative whereas the palato-alveolar fricatives are categorized as the laminal post-alveolar fricatives (Ladefoged & Johnson, 2011).

Acoustic measurements, such as spectral peak location and spectral moments, have been used to distinguish the place of articulation of sibilant fricatives (Forrest, Weismer, Milenkovic, & Dougall, 1988; Jongman et al., 2000; Lee, Zhang, Li, Tao, & Bond, 2012). In general, a shorter cavity preceding the constriction in the oral cavity is correlated with a higher spectral peak location. For example, the spectral peak location for the alveolar fricatives /s, z/ is generally higher in frequency than the spectral peak location of the palato-alveolar fricatives /ʃ, ʒ/ (Jongman et al., 2000; Lee et al., 2012). Spectral moments are statistical analyses commonly used to examine the spectral properties of stops and fricatives (Forrest et al., 1988; Jongman et al., 2000). There are four spectral moments: (a) spectral mean or centre of gravity, (b) standard deviation, (c) spectral skewness, and (d) spectral kurtosis. According to Jongman et al. (2000), spectral mean reflects the average energy concentration of fricative noise, standard deviation reflects the range of fricative noise concentration, spectral skewness reflects spectral tilt (e.g., positive skewness suggests concentration of energy at lower frequencies whereas negative skewness suggests concentration of energy at higher frequencies), and spectral kurtosis reflects peakedness of the fricative noise spectral distribution (e.g., positive kurtosis suggests a well-defined spectrum whereas negative kurtosis suggests a flat spectrum). Jongman et al. (2000) reported that among English sibilant fricatives, the alveolar fricatives /s, z/ had a higher spectral mean (e.g., 6133 Hz) than the palato-alveolar fricatives /ʃ, ʒ/ (e.g., 4229 Hz); voiceless fricatives (e.g., 5267 Hz) had a higher spectral mean than voiced fricatives (e.g., 5036 Hz); fricatives spoken by female talkers (e.g., 5286 Hz) had a higher spectral mean than the ones spoken by male talkers (e.g., 5018 Hz). For Mandarin sibilant fricatives, Lee et al. (2012) found that spectral peak location, spectral mean, and spectral skewness can distinguish between the alveolar /s/ and retroflex /ʂ/ fricative. Lee (2011) examined the spectral mean of three

Mandarin sibilant fricatives /s, ɕ, ʂ/ and reported that the alveolar fricative /s/ had the highest spectral mean, followed by the alveolo-palatal /ɕ/ and retroflex /ʂ/ fricatives.

It is unknown whether SMNR processing will alter the spectral properties of the fricative consonants. If the spectral mean of a fricative consonant is lowered by a substantial amount, it is possible that this may result in confusion of the place of articulation for the fricative consonant. Research has shown that adults weighted frication noise spectra more than formant transition when discriminating place of articulation for sibilant fricatives, for example the alveolar fricative /s/ vs. the post-alveolar fricative /ʃ/ (Nittrouer, 2002). For this reason, the first spectral moment (i.e., spectral mean) was included in this study to quantify the effect of SMNR-processing on the short-term spectral properties of Mandarin and English fricatives and affricates.

3.1.1.2 Affricates

An affricate consonant is a sequence of a stop consonant followed by a homorganic fricative consonant (Ladefoged & Johnson, 2011). There are two affricates in English (i.e., the palato-alveolar /tʃ, ʤ/ affricates) and six affricates in Mandarin produced at three different place of articulation (i.e., alveolar /ts, ts^h/, alveolo-palatal /tɕ, $tɕ^h$/, and retroflex /tʂ, $tʂ^h$/). There are three distinct differences between Mandarin and English affricates: (i) Mandarin affricates do not have a voicing contrast, but English affricates do, (ii) Mandarin affricates have a place of articulation contrast, but English affricates do not, and (iii) Mandarin affricates produced at the same place of articulation have an aspiration contrast, but English affricates do not.

Phonetic cues such as the frication duration and amplitude rise time (i.e., the time for the frication noise to reach its maximum intensity) are important for the affricate-fricative distinction. However, studies have shown that frication duration cue is more effective in the distinction between fricative and affricate (Howell & Rosen, 1983; Kluender & Walsh, 1992; Mitani, Kitama, & Sato, 2006; Repp, Liberman, Eccardt, & Pesetsky, 1978; Tsao et al., 2006). Howell and Rosen (1983) recorded the voiceless palato-alveolar affricate /tʃ/ and fricative /ʃ/ in running speech, isolated words, and isolated nonsense syllables spoken by four native-English talkers. In each form, these consonants occurred in the word initial, medial, and final positions. Howell and Rosen (1983) found that the rise times and frication durations for affricates were always shorter than for fricatives. Perceptually, consonants with a longer frication duration tend to be perceived as fricatives, whereas consonants with a shorter frication duration tend to be perceived as affricates (Mitani et al., 2006; Repp et al., 1978). Kluender and Walsh (1992) conducted two experiments to determine whether the rise time or the frication duration alone affected the perception of affricate-fricative distinction. In the first experiment, rise time was varied while the frication noise duration was held constant. In the second experiment, rise time was held constant and the duration was varied. Overall, their results showed that fricative noise duration alone is sufficient to distinguish between English affricates and fricatives. In another study, Tsao et al. (2006) examined the effect of duration and amplitude rise time on the perception of Mandarin fricative-affricate contrasts. They found that frication duration is important for the distinction between the aspirated alveolo-palatal affricate /tɕʰ/ vs. unaspirated alveolo-palatal affricate /tɕ/ and the unaspirated alveolo-palatal affricate /tɕ/ vs. alveolo-palatal fricative /ɕ/. Although frication duration is an important cue for affricate-fricative distinction, it is unlikely that SMNR processing will affect the duration of the frication noise. Thus, measurement of frication duration was not included in this study. However, it is unknown whether gain reduction in SMNR processing will alter the amplitude of the release burst and frication noise sections of affricates and fricatives. Thus, measurements of release burst amplitude (for affricates only) and frication noise amplitude (for fricatives and affricates) were included in this study to quantify the acoustic effects of SMNR on fricatives and affricates.

In summary, the aim of Study 2 was to determine the acoustic effects of SMNR on the fricatives and affricates of the English and Mandarin languages. Acoustic measurements were performed on word strings retrieved from the SMNR-on and SMNR-off recordings: (i) the amount of noise reduction measurement was used to quantify the effect of SMNR on noise, (ii) the effective SNR change and HASQI measurements were used to quantify the effect of SMNR on speech and noise, (iii) the frication noise amplitude, release burst amplitude, and spectral-mean measurements were used to quantify the effect of SMNR on Mandarin and English fricatives and affricates. SMNR processing would be expected to cause (i) an increment to the amount of noise reduction when the noise level increased and input SNR decreased, (ii) an improvement on the effective SNR as the input SNR decreased, (iii) changes in the temporal envelope, amplitude, and spectral properties (e.g., long-term spectra and spectral mean).

3.2 Method

3.2.1 Hearing Aids

The Oticon Safari P300 and Phonak Solana M H20 behind-the-ear hearing aids were used because they incorporated manufacturer-specific MBNR. Hereafter, the Oticon Safari P300 and Phonak Solana M H20 are referred to as HA#1 and HA#2, respectively. Electroacoustic measurements were performed on the hearing aids to ensure that they met the manufacturer's specification. Two listening programs were configured in each hearing aid: SMNR was enabled in one of the listening programs and was set to maximum strength according to the manufacturer's programming software; SMNR was disabled in the other program. For each program in both hearing aids, the omnidirectional microphone was activated and all other advanced signal processing features (e.g., feedback cancelation, wind noise reduction, transient noise reduction, frequency lowering, and adaptive directional microphone etc.) were disabled.

Both hearing aids were set to provide gain for a moderate to moderately-severe sensorineural hearing loss (i.e., 40, 45, 50, 55, 60, 60, 65, and 65 dB HL at 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz, respectively) using the Desired Sensation Level v5.0 (Scollie et al., 2005) prescription method for speech at a 65 dB SPL input level. The frequency responses of the SMNR-on and SMNR-off programs in both hearing aids were verified using Speechmap^{TM} in Audioscan Verifit® (Etymotic Design Inc., Dorchester, ON, Canada) to ensure that both programs had similar frequency responses. The frequency response curves for the SMNR-on and SMNR-off programs in HA#1 and HA#2 are shown in Figure 8. Note that the pink curve overlapped with the green curve; the overlapping curves showed that the frequency responses of SMNR-on and SMNR-off programs in both hearing aids were identical for 65 dB speech inputs.

The gain for higher (75 dB SPL) and lower (55 dB SPL) input levels was adjusted to be the same as the gain at 65 dB SPL input level to ensure linearity of the gain provided by the hearing aids. Linear gain was chosen in order to control for nonlinear compression as a confounding factor. The linearity was checked using the input-output test in Audioscan Audioscan Verifit® (Etymotic Design Inc., Dorchester, ON, Canada). Figure 9 and Figure 10 show the input-output curve for HA#1 and HA#2, respectively. Both figures showed that the input-output function for both hearing aids at most frequencies was essentially linear, except at 4000 Hz for HA#1 (Figure 9) and at 250 Hz for HA#2 (Figure 10).

The output level of each program in HA#1 and HA#2 was measured using the Fonix 7000 test system (Frye Electronics, Inc., OR, USA). Composite test signals were presented to each hearing aid at 50 to 90 dB SPL under SMNR-on and SMNR-off conditions, and the values of the hearing aid output were recorded. The compression ratios were then calculated from the 55 to 75 dB SPL input-output function of each hearing aid. Table 3 shows that the compression ratios for both programs in each hearing aid were essentially linear from 500 Hz to 4000 Hz. The maximum power output for each hearing aid was set to the maximum to control for output limiting as a confounding factor.

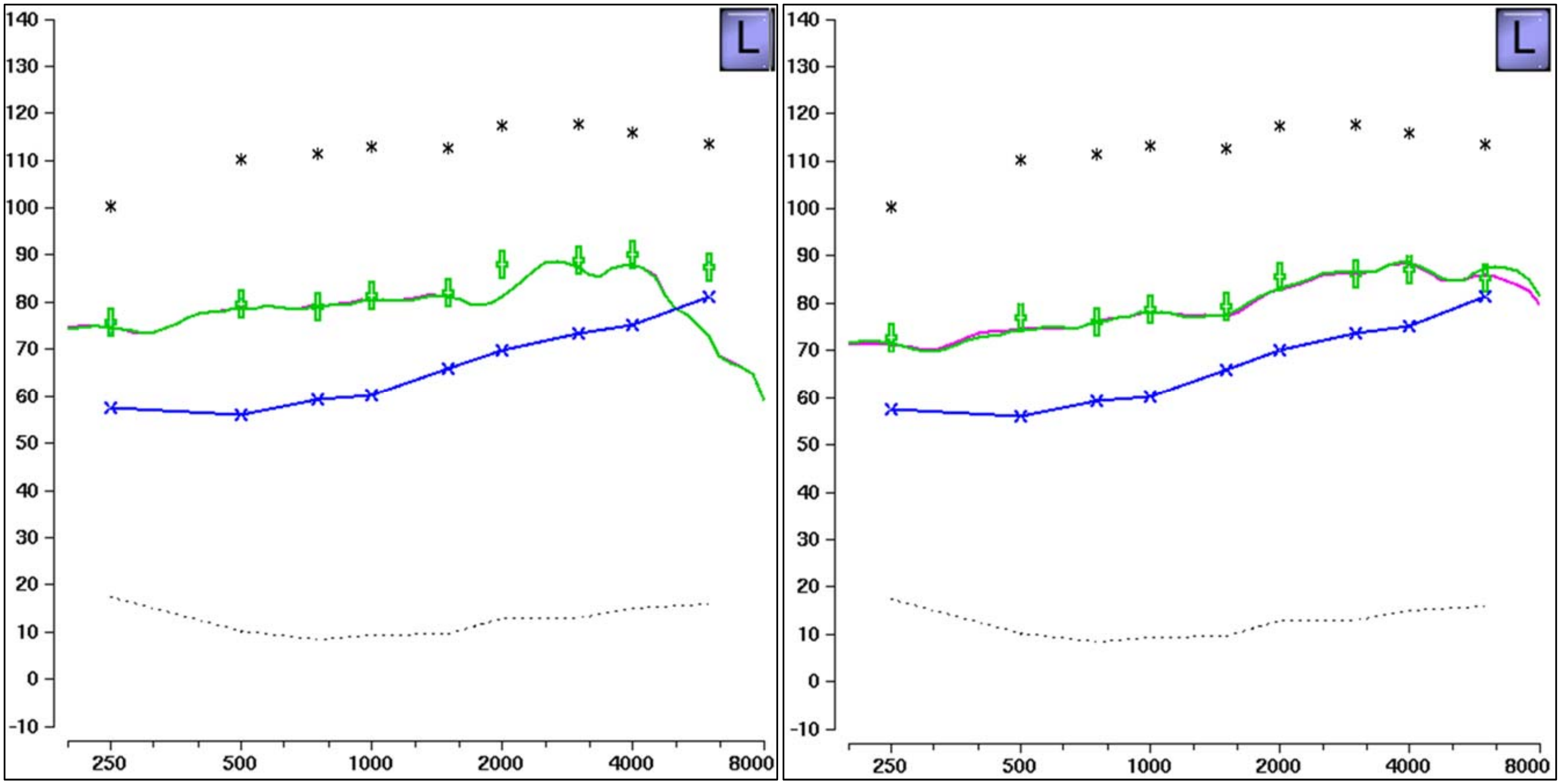


Figure 8. Frequency response curves for HA#1 (left panel) and HA#2 (right panel). The y-axis represents output level (dB SPL). The x-axis represents frequency (Hz). The curves were obtained with speech signal at 65 dB (green and pink curves). The green curves represent the SMNR-on program whereas the pink curves (overlapped by the green curves) represent the SMNR-off program. The blue curves represent the hearing threshold level (the impaired thresholds) used to program the hearing aids. The asterisks indicate the predicted uncomfortable listening level. The dotted lines represent normal hearing threshold levels (i.e., minimum audible pressure).

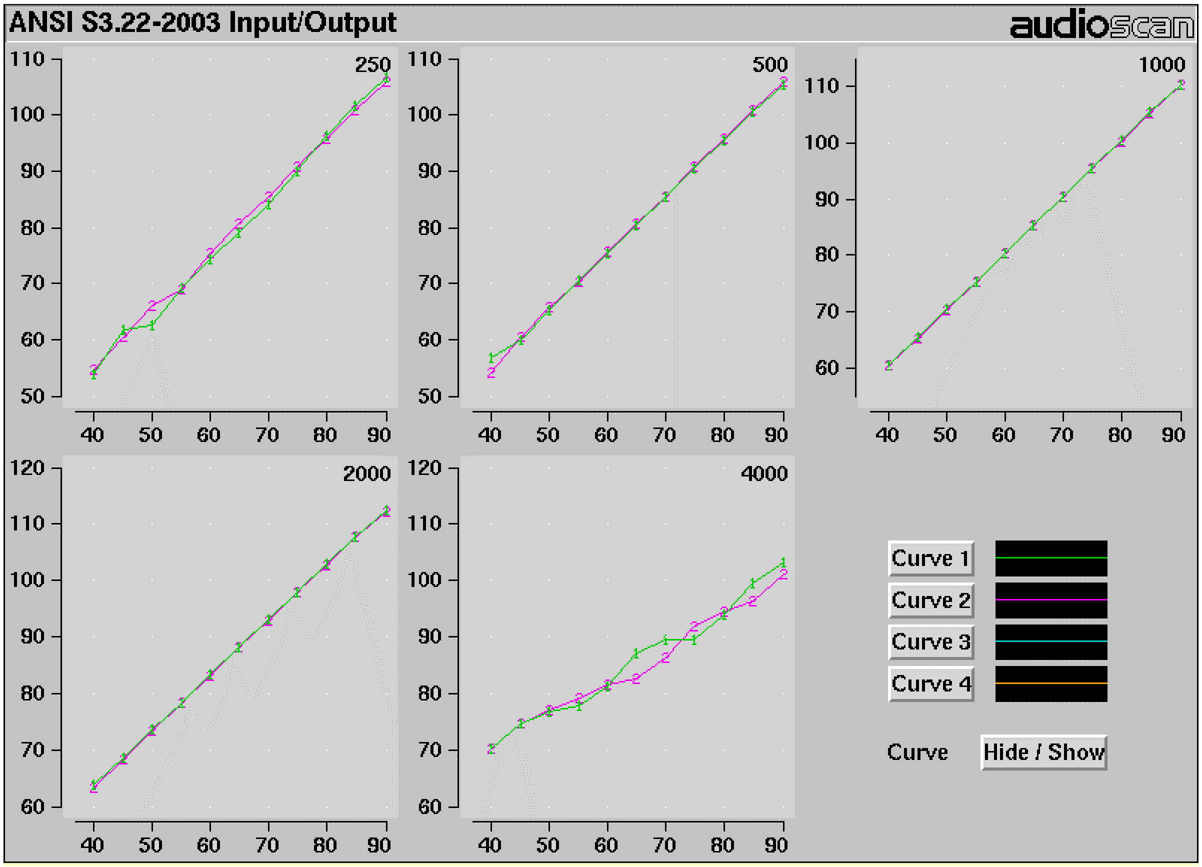


Figure 9. Input-output curves for HA#1. The y-axis represents output level (dB SPL). The x-axis represents input level (dB SPL). The green line represents the SMNR-on program and the pink line represents the SMNR-off program.

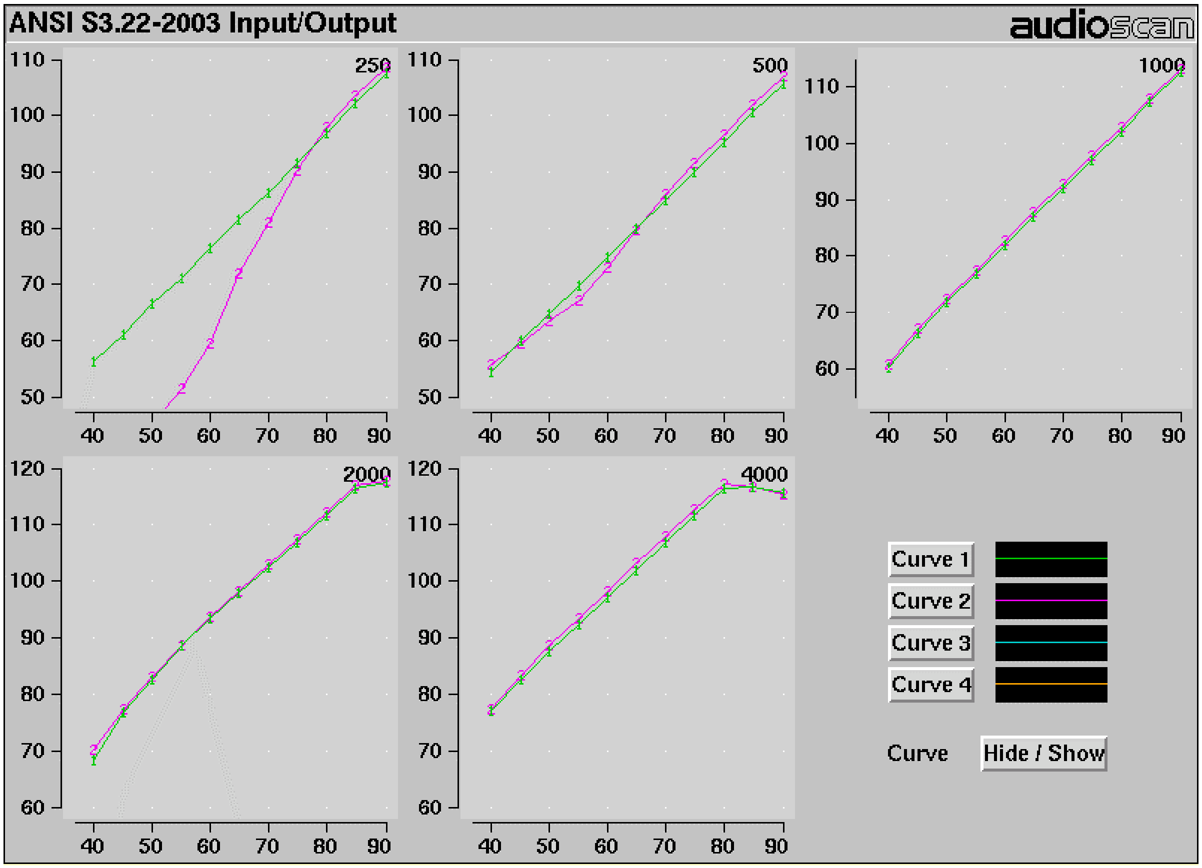


Figure 10. Input-output curves for HA#2. The y-axis represents output level (dB SPL). The x-axis represents input level (dB SPL). The green line represents the SMNR-off program and the pink line represents the SMNR-on program.

Table 3. Compression ratio for HA#1 and HA#2.

	Frequency (Hz)				
Hearing aid	500	1000	2000	3000	4000
HA#1					
SMNR-on	1.0	1.0	1.0	1.0	1.0
SMNR-off	1.0	1.0	1.0	1.0	0.9
HA#2					
SMNR-on	0.9	0.9	0.9	0.9	0.9
SMNR-off	0.9	0.9	0.9	0.9	0.9

3.2.2 Stimuli

The speech-plus-noise stimulus files developed for Study 1 were used as the test stimuli in this study. The speech signals consisted of six Mandarin vowel-consonant-vowel (VCV) syllables and five English VCV syllables, spoken by a female and a male talker of each language. Sets of VCV syllables with the same consonant in three vowel contexts (e.g., /aSa, iSi, uSu/), spoken by a single talker, were repeated five times and concatenated into a word string. In total, there were 15 VCV tokens in each word string. A total of 22 word strings was developed, as listed in Appendix C. All word strings were equalized in RMS voltage prior to adding the noise. The noise consisted of a steady-state pink noise generated from the Audacity program. Steady-state broadband noise was used because its randomness mimics the frication in fricatives and affricates. This noise file was stored as "original noise" and its waveform was 180^{o} phase-inverted and stored as the "inverted noise."

Two types of speech-plus-noise files were developed, as shown in Appendix D: (i) original-speech plus original-noise and (ii) original-speech plus phase-inverted noise. This set of two files allowed the application of the inversion technique for speech and noise extraction post hearing-aid processing (Hagerman & Olofsson, 2004). Each word string and each of the "original" and "phase-inverted" pink noise were mixed at -10 to +10 dB SNR in 5 dB steps to develop sets of speech-plus-noise files. The speech signal was kept at a constant RMS voltage level, while the RMS voltage level of pink noise was varied to achieve the desired SNRs. This range covers the relevant input SNRs in everyday speech-in-noise listening conditions (Olsen, 1998; Pearson, Bennet, & Fidell, 1977). Pink noise was added to the beginning of each speech-plus-noise file to ensure the activation of SMNR prior to the beginning of speech. On average, the duration of the pink noise was 22 seconds prior to the speech-plus-noise signal. All test stimuli were developed using Audacity and Praat (version

5.3.61; Boersma & Weenick, 2014). These stimuli were digitized at 44.1 kHz and 32 bit resolution, and stored in the computer as .wav files.

3.2.3 Procedure

The recording set up was similar to the one used in Study 1, as show in appendix E. The recordings were collected in a double-walled sound-treated booth. Each of the speech-plus-noise files was presented to one of the two hearing aids mounted on the left ear of Knowles Electronic Manikin for Acoustic Research (KEMAR; G.R.A.S. Sound and Vibration, Denmark). The stimuli were presented via a loudspeaker at the levels presented in Table 4. These levels were used to maintain the desired input SNRs and the average speech level at 65 dBA. In addition, these levels were used to minimize the likelihood of reaching the output limiting of the hearing aids. Sound field calibration was conducted for each speech-plus-noise file to ensure consistent presentation level to the microphone of the hearing aid.

Table 4. Sound field calibration level for test stimuli.

Input SNR condition	Level: noise only (dBA)	Level: speech and noise (dBA)
+10 dB SNR	54.7-55.0	63.5-66.0
+5 dB SNR	60.0-60.3	65.5-68.0
0 dB SNR	64.9-65.0	68.0-70.0
-5 dB SNR	69.8-69.9	70.0-71.0
-10 dB SNR	74.6-74.9	75.0-75.4

The speech-plus-noise output from each hearing aid was recorded under the SMNR-on and SMNR-off conditions. In total, 8800 recordings (22 word strings X 2 types of speech-plus-noise files X 5 SNRs X 2 SMNR conditions X 2 hearing aids) were collected. The recordings were digitized at 44.1 kHz and 32-bit resolution and stored in the computer in .wav file format for offline analysis. The inversion technique was used to extract speech (“retrieved-speech”) and noise (“retrieved-noise”) from the speech-plus-noise recordings.

The extraction was done using the Audacity program. Given that the merged files resulted in doubling of intensity level, 6 dB was subtracted from the overall level of the retrieved-speech and retrieved-noise .wav files. The retrieved-speech files for HA#2 in the -5 and -10 dB SNR input conditions were excluded from further analyses because these files did not meet the speech fidelity criteria, as shown in the first study of this dissertation.

Acoustic measurements were performed on the retrieved-speech and retrieved-noise waveforms from the SMNR-on and SMNR-off conditions. All acoustic measurements were performed using Audacity, Praat (version 5.3.61; Boersma & Weenick, 2014), or MATLAB (The MathWorks Inc., USA). The long-term acoustic measurements included (i) amount of noise reduction, (ii) effective SNR improvement, and (iii) HASQI.

For the amount of noise reduction, two 3-second sections of retrieved-noise from the SMNR-on recording condition were selected as shown in Figure 11. The first section selected was at the beginning of the retrieved-noise waveform (i.e., 0 to 3 seconds) where SMNR had not yet been engaged. The second section selected was a steady-state portion of the retrieved-noise waveform (a 3-second section labelled as T1 to T2) after SMNR was fully engaged. The mean energy of each section was computed in Praat. The mean intensity between the first section (0 to 3 seconds) and the second section (T1-T2) was compared and the difference was taken as the amount of noise reduction.

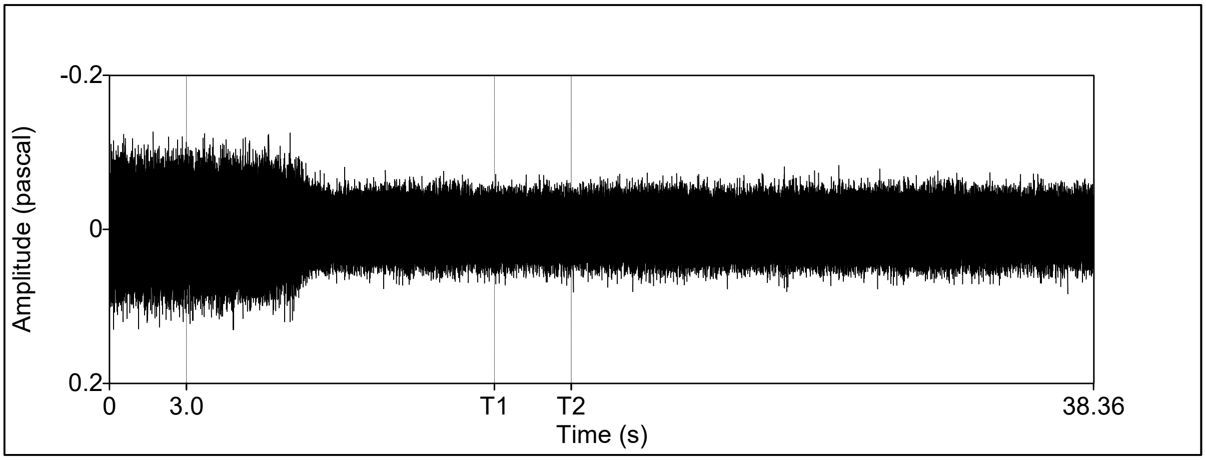


Figure 11. Time selection for measuring amount of noise reduction.

The effective SNR change is defined by Pittman (2011a, 2011b) as the output SNR difference between the SMNR-on and SMNR-off conditions. In order to estimate the output SNR for each SMNR condition (i.e., on vs. off), the long-term average spectra of the retrieved-speech and retrieved-noise signals within each condition were compared. The output SNR for the SMNR-off condition was subtracted from the output SNR of the SMNR-on condition. A positive difference indicates improvement in overall output SNR whereas a negative difference indicates degradation of overall output SNR of the speech-plus-noise signals after SMNR processing.

The HASQI measurement (Kates & Arehart, 2010; Kates & Arehart, 2014) was used to quantify the effect of SMNR processing on speech fidelity. This was carried out using a MATLAB function written by Kates (2010). According to J. M. Kates (personal communication, September 25, 2014), the reference signal in the HASQI measurement should always be a clean, unprocessed signal whereas the test signal should be the processed signal. Hence, the stimulus files (i.e., word string) were used as the reference signals and the retrieved-speech files for SMNR-on and SMNR-off conditions were used as the test signals in all HASQI measurements in this study. Recall that there are two indices in the HASQI measurement: (i) the nonlinear index and (ii) the linear index. The MATLAB function returns a value between 0 and 1 for each index. A value of 0 indicates that the test signals are different from the reference signals and a value of 1 indicates that the test signals are similar to the reference signals. The scores of each index for the SMNR-on and SMNR-off conditions were compared. The difference score between SMNR-on and SMNR-off conditions was taken as an indication of whether SMNR processing affected the speech signals in terms of temporal modulations (as indicated by the HASQI nonlinear index) and long-term spectra (as indicated by the HASQI linear index).

The criterion for a significant difference on the HASQI scale was estimated from the standard deviation (SD) of subjective quality ratings obtained by Arehart et al. (2010), converted into the HASQI scale. In their study, subjective ratings were performed on a 5-point scale that ranged from 1 (*bad quality*) to 5 (*excellent quality*). The scale was implemented as a slide bar with a step size of 0.05 in their experiment. Kates and Arehart (2010) reported that the SD of participants' subjective quality ratings in the Arehart et al. (2010) study was approximately 0.6 for any simulated processing conditions examined in the study. The subjective quality ratings (on a 1 to 5 scale) of each participant were then transformed into the HASQI scale (0 to 1) using a linear interpolation (J.M. Kates, personal communication, November 6, 2014). Based on this conversion, a value of one SD (i.e., 0.6) in the subjective rating scale would be equivalent to a value of 0.15 when it was linearly transformed into the HASQI scale. Therefore, a criterion value of ±0.15 was used to determine whether SMNR processing affected the fidelity of speech signals. For example, a difference score between SMNR-on and SMNR-off conditions that falls within the range of

±0.15 in the HASQI nonlinear index suggests that SMNR processing examined in the current study does not affect the temporal modulations of the speech signals and a difference score within the range of ±0.15 in the HASQI linear index suggests that SMNR does not affect the long-term spectra of speech signals.

The short-term acoustic analyses were performed on the fricative and affricate segments of the retrieved-speech files. The measurements include (i) frication-noise intensity difference, (ii) release-burst intensity difference, and (iii) spectral mean difference for frication noise. The fricative noise and release burst segments were identified using simultaneous auditory and visual inspection of the acoustic waveform and wide-band spectrogram (0.005 second window length) of each VCV syllable within the retrieved-speech files. The wide-band spectrogram was generated using a Gaussian window shape in Praat. Other settings in Praat are stated in Appendix F. For an affricate consonant, the release burst was defined as an impulsive waveform following a period of silence after an initial vowel; frication onset was defined as the point where high frequency energy first appeared on the spectrogram (see Figure 12 for an example). For a fricative consonant, frication onset was defined as the first noticeable turbulence in the sound waveforms following the offset of the voicing bar of the preceding vowels (see Figure 13 for an example). For both consonants, frication offset was defined as the point just before the appearance of the voicing bar of the following vowel. Frication onset is labelled as FS (i.e., frication start) and frication offset is labelled as FE (i.e., frication end) in Figure 12 and 13.

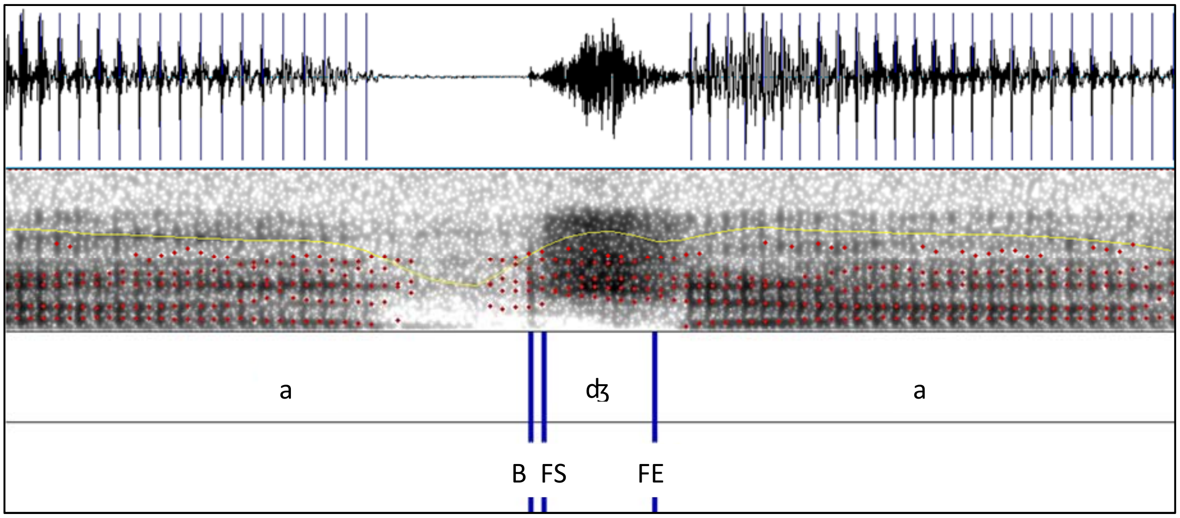


Figure 12. Segmentation of release burst and frication noise onset and offset of a VCV token that contained an affricate. The top panel represents the acoustic waveform of the nonsense word /aJa/; the y-axis represents the amplitude (dB SPL). The second panel represents the spectrogram of that word; the y-axis represents the frequency (0 Hz to 10 kHz). The red dotted lines in the spectrogram represent formant estimates, the yellow lines represents the intensity contour. The level of darkness in the spectrogram shows the strength of the sounds energy (range = 0 dB to 100 dB; darker = higher energy). B = burst onset; FS = frication onset; FE = frication offset.

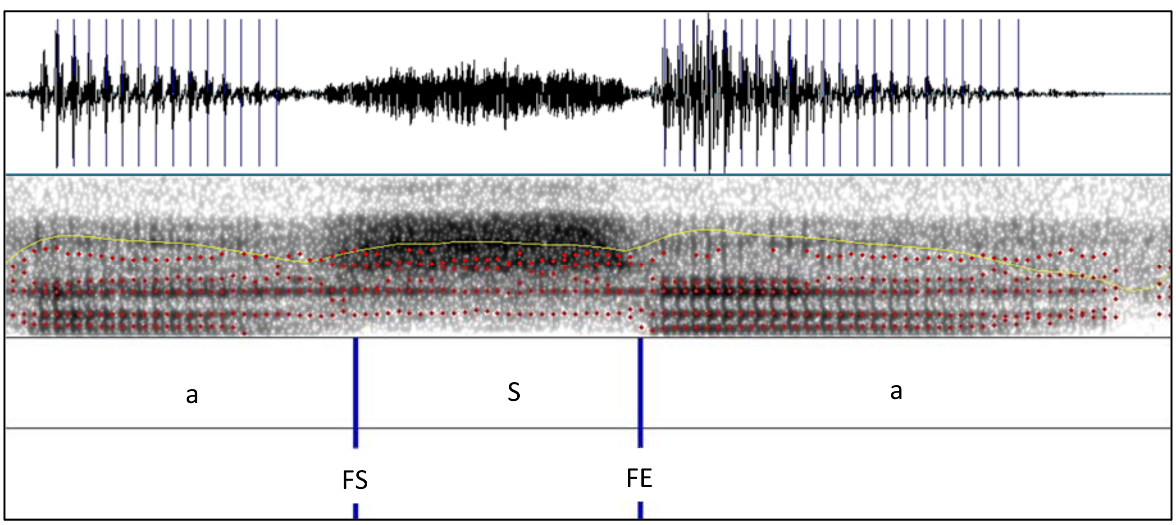


Figure 13. Segmentation of frication noise onset and offset of a VCV token that contained a fricative. The top panel represents the acoustic waveform of the nonsense word /aSa/; the y-axis represents the amplitude (dB SPL). The second panel represents the spectrogram of that word; the y-axis represents the frequency (0 Hz to 10 kHz). The red dotted lines in the spectrogram represent formant estimates, the yellow lines represents the intensity contour. The level of darkness in the spectrogram shows the strength of the sounds energy (range = 0 dB to 100 dB; darker = higher energy). FS = frication onset; FE = frication offset.

Recall that there were 22 speech-in-noise files processed by each of the hearing aids, with SMNR-on and SMNR-off, under five input SNR conditions. Each speech-in-noise file contained a consonant that was paired with one of /a/, /i/, or /u/ in a VCV format (e.g., /aSa, iSi, uSu/). This set of three VCV tokens, spoken by a single talker, was repeated five times in a word string. In total, there were 15 VCV tokens in each speech-in-noise file. These speech-in-noise files were preceded by pink noise during recordings. The effect of SMNR may vary for noise-only and speech-in-noise signal. Therefore, the first set of VCV tokens (that occurred right after the noise-only section) was excluded from the measurement, leaving four tokens per vowel condition for each consonant.

The mean intensities of the frication noise and release burst sections were computed in Praat. First, the selected section was enlarged and the mean intensity of the section was computed using the "Intensity > Get intensity" command in Praat. The averaging method for measuring intensity was "mean energy." The mean intensity of the frication noise and the release burst sections for VCV tokens processed with SMNR-on were compared to the mean intensity of the same sections for the corresponding VCV tokens processed with SMNR-off.

The spectral mean difference was measured with the Praat program (version 5.3.61; Boersma & Weenick, 2014). First, a frication noise section (duration from frication start, FS, to frication end, FE) of a VCV token was selected and a spectral slice was generated. From the spectral slice, the spectral mean was computed by using the "Query > Get centre of gravity…" command. These steps were repeated for VCV tokens obtained under SMNR-on and SMNR-off conditions. The spectral mean of the corresponding VCV tokens for both SMNR conditions were compared and the spectral mean differences were calculated.

3.3 Results

The data are presented descriptively because some of the acoustic analyses such as the intensity difference and spectral mean difference do not have pre-define criterion values.

3.3.1 Amount of Noise Reduction

The amount of noise reduction for HA#1 and HA#2 is displayed in Table 5. The amount of noise reduction for each input SNR condition was an average of 11 measurements. For HA#1, the mean amount of noise reduction ranged from 1.1 to 4.8 dB for noise levels of 55 to 75 dB SPL; the SD was 0.1 dB across all noise input levels. For HA#2, the mean amount of noise reduction ranged from 3.4 to 6.3 dB; the SD ranged from 0.3 to 0.9 dB.

Table 5. Amount of noise reduction for HA#1 and HA#2.

Input noise level (dBA)	Input SNR condition (dB)	Amount of noise reduction (dB)		
		Mean	SD	Range
HA#1				
55	+10	1.1	0.1	1.0 to 1.1
60	+5	1.9	0.1	1.8 to 2.1
65	0	2.8	0.1	2.7 to 3.0
70	-5	3.8	0.1	3.7 to 3.9
75	-10	4.8	0.1	4.6 to 5.0
HA#2				
55	+10	5.4	0.4	4.4 to 5.9
60	+5	6.3	0.8	5.0 to 7.6
65	0	6.1	0.9	5.2 to 7.5
70	-5	5.2	0.4	4.4 to 5.6
75	-10	3.4	0.3	2.9 to 3.8

3.3.2 Effective Signal-to-Noise Ratio Change

Table 6 shows the output SNR and effective SNR change for HA#1 and HA#2, averaged across 22 measurements, in each of the input SNR conditions. The nominal input SNR was not reproduced at the output of the hearing aids under SMNR-on and SMNR-off conditions for either hearing aid. For HA#1, the average effective SNR change ranged from 0.1 to 1.6

dB; SD ranged from 0.2 to 0.6 dB. For HA#2, the average SNR change ranged from 2.1 to 4.8 dB; SD from 1.0 to 1.7 dB.

Table 6. Effective SNR change for HA#1 and HA#2.

	Output SNR (dB)			
Nominal input	SMNR-on	SMNR-off	Effective SNR change (dB)	
SNR (dB)	Mean (SD)	Mean (SD)	Mean (SD)	Range
HA#1				
+10	7.1 (1.2)	6.9 (1.3)	0.2 (0.2)	-0.2 to 0.5
+5	2.5 (1.2)	1.8 (1.3)	0.7 (0.5)	0.2 to 1.7
0	-1.9 (1.1)	-3.2 (1.3)	1.3 (0.6)	0.3 to 2.3
-5	-6.6 (1.1)	-8.2 (1.3)	1.6 (0.6)	0.7 to 2.4
-10	-11.6 (1.2)	-13.1 (1.3)	1.5 (0.5)	0.8 to 2.1
HA#2				
+10	4.7 (1.2)	-0.6 (1.4)	4.6 (1.0)	2.4 to 6.2
+5	0.2 (1.6)	-4.9 (1.5)	5.1 (1.3)	2.7 to 7.1
0	-4.5 (1.6)	-9.3 (1.6)	4.8 (1.7)	1.7 to 6.9
-5	-8.9 (1.6)	-12.4 (1.6)	3.5 (1.1)	1.0 to 5.2
-10	-14.3 (1.8)	-16.3 (1.6)	2.1 (1.2)	0.1 to 6.0

3.3.3 Speech Fidelity

Figure 14 illustrates the mean HASQI nonlinear index and Figure 15 illustrates the mean HASQI linear index for HA#1 in each of the five SNR conditions (+10, +5, 0, -5, and -10 dB SNR). Figure 16 illustrates the mean HASQI nonlinear index and Figure 17 illustrates the mean HASQI linear index for HA#2. For HA#2, only three input SNR conditions were illustrated in both figures. The data points for the -5 and -10 dB SNR conditions were not illustrated because the retrieved-speech signals in these conditions were determined to have the poorest speech fidelity in the first study (Chapter 2). Hence, the retrieved-speech signals of HA#2 in the -5 and -10 dB SNR conditions were excluded from further analysis. The means in Figure 14 to Figure 17 were obtained by averaging six measurements for the Mandarin word strings (i.e., one word string for each of the six consonants) and five measurements for the English word strings (i.e., one word string for each of the five consonants). The error bars represent ±1 SD from the mean. The white bars represent the SMNR-on condition and the grey bars represent the SMNR-off condition.

Table 7 shows the average difference score between SMNR-on and SMNR-off for the HASQI nonlinear index of each hearing aid. Table 7 shows that for HA#1, the difference scores of the HASQI nonlinear index between SMNR-on and SMNR-off conditions was essentially zero in four out of the five SNR conditions (+10, +5, 0, and -5 dB SNR); the difference score was close to -0.1 in the -10 dB SNR condition. For HA#2, the difference scores of the HASQI nonlinear index between SMNR-on and SMNR-off conditions was between -0.04 to -0.09 in the +10, +5 and 0 dB SNR conditions. As noted earlier, a criterion value of ±0.15 was used to determine whether SMNR processing affected the fidelity of speech signals. The difference scores of the HASQI nonlinear index between SMNR-on and SMNR-off conitions for both HA#1 and HA#2 were within the range of the criterion value.

The difference scores of the HASQI linear index between SMNR-on and SMNR-off conditions were essentially zero for both hearing aids (HA#1 and HA#2) under all conditions (e.g., languages, talker gender, and SNRs). A zero difference score of the HASQI linear index indicates SMNR processing has no effect on the long-term spectra of the word strings.

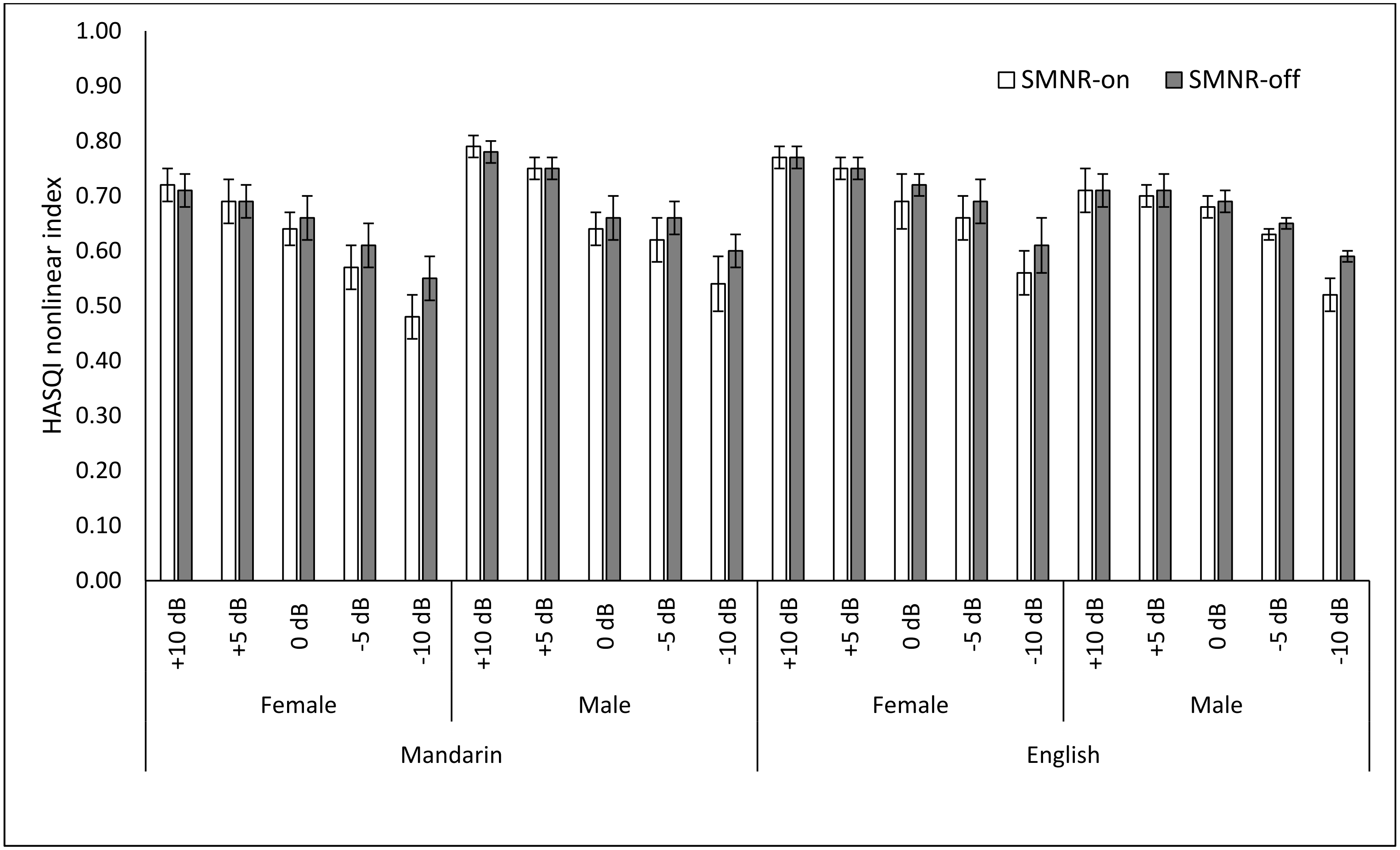


Figure 14. Mean HASQI nonlinear index for speech signals processed by HA#1 under five input SNR conditions. The error bars represent ± one SD from the mean. The white bars represent the SMNR-on condition and the grey bars represent the SMNR-off condition.

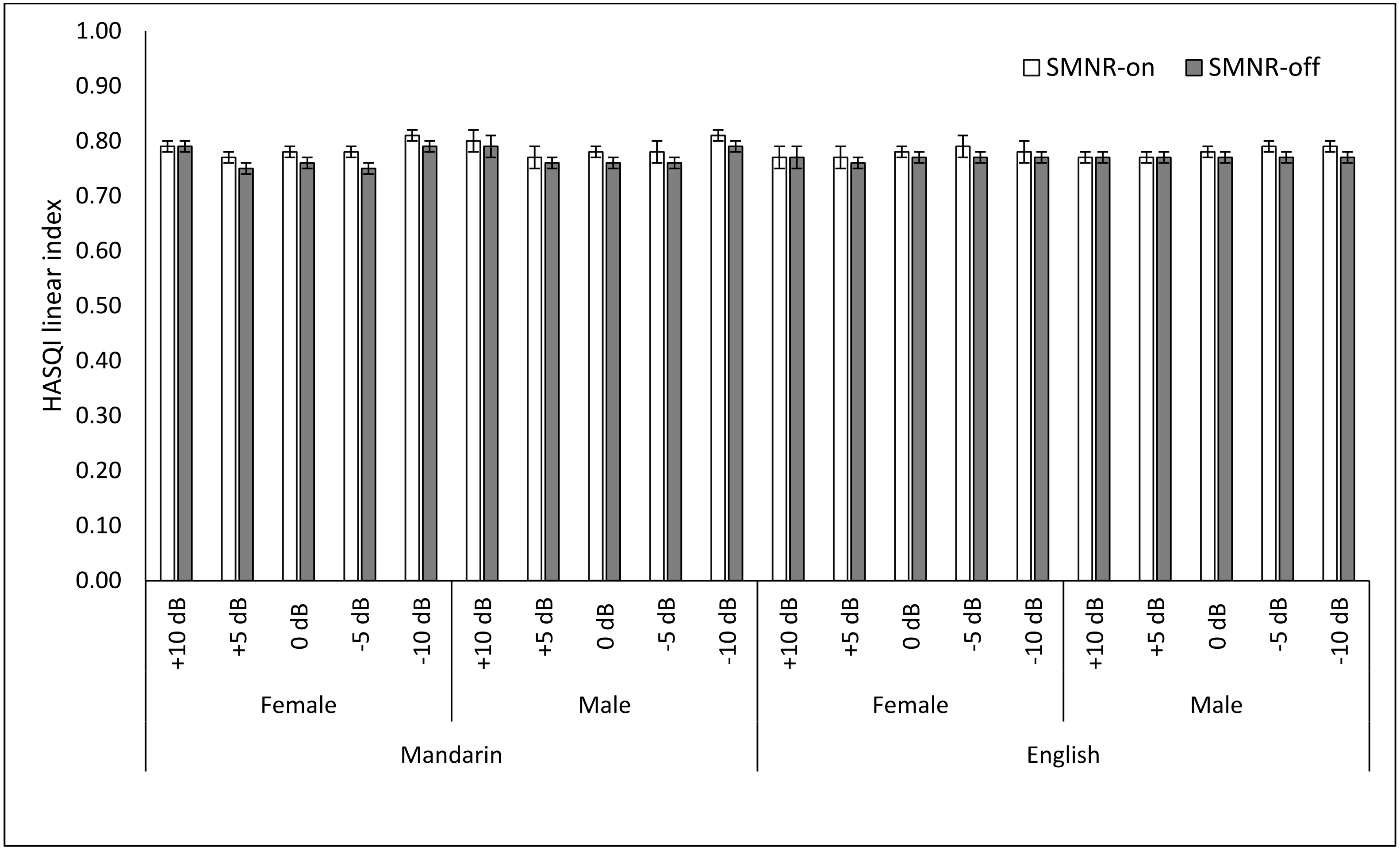


Figure 15. Mean HASQI linear index for speech signals processed by HA#1 under five input SNR conditions. The error bars represent ± one SD from the mean. The white bars represent the SMNR-on condition and the grey bars represent the SMNR-off condition.

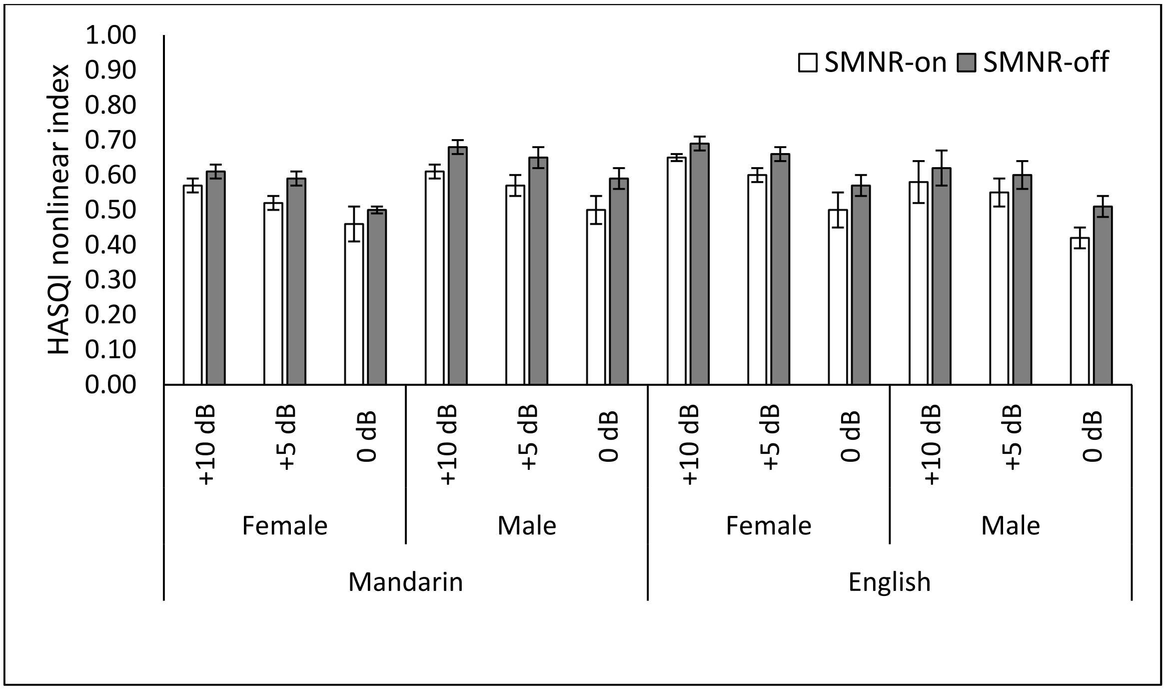


Figure 16. Mean HASQI nonlinear index for speech signals processed by HA#2 under three input SNR conditions. The error bars represent ± one SD from the mean.

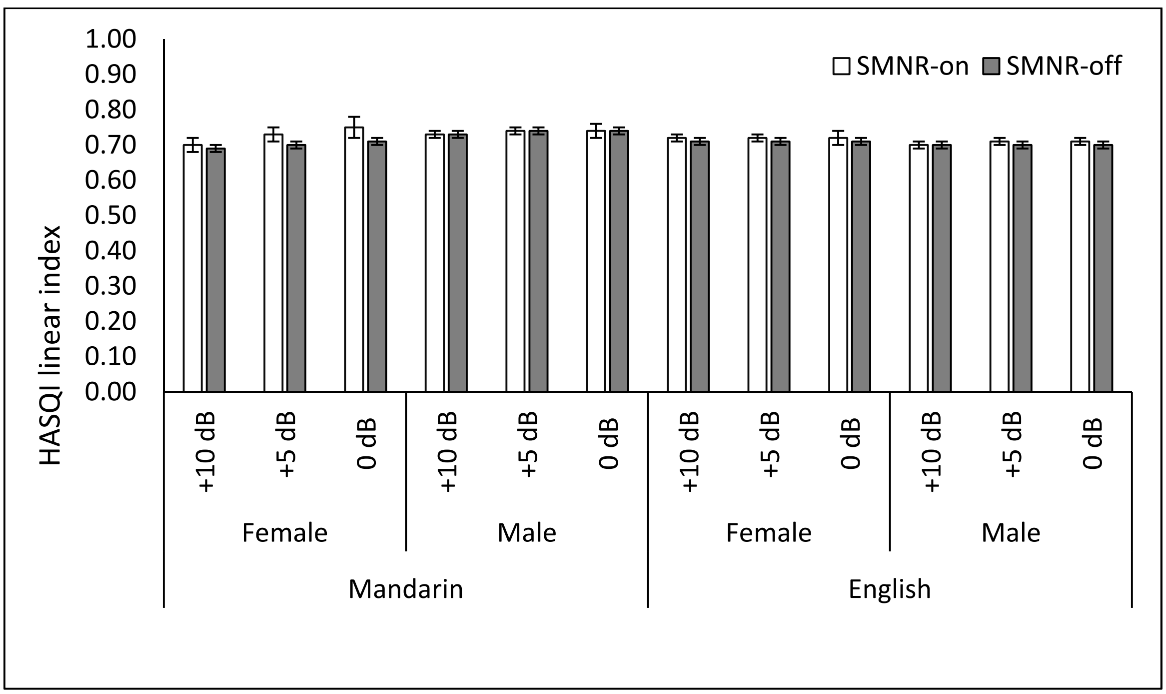


Figure 17. Mean HASQI linear index for speech signals processed by HA#2 under three input SNR conditions. The error bars represent ± one SD from the mean.

Table 7. The mean difference score of HASQI nonlinear index between the SMNR-on and SMNR-off conditions for HA#1 and HA#2.

Input SNR (dB)	Nonlinear index difference score; mean (SD)			
	Mandarin		English	
	Female	Male	Female	Male
HA#1				
+10	0.01 (0.01)	0.01 (0.00)	0.00 (0.00)	0.00 (0.01)
	0.00 (0.01)	0.00 (0.01)	0.00 (0.02)	-0.01 (0.01)
0	-0.02 (0.01)	-0.02 (0.01)	-0.03 (0.02)	-0.01 (0.01)
-5	-0.04 (0.01)	-0.04 (0.01)	-0.03 (0.01)	-0.02 (0.01)
-10	-0.07 (0.02)	-0.06 (0.02)	-0.05 (0.02)	-0.07 (0.03)
HA#2				
+10	-0.04 (0.01)	-0.07 (0.01)	-0.04 (0.02)	-0.04 (0.01)
+5	-0.07 (0.01)	-0.08 (0.01)	-0.06 (0.03)	-0.05 (0.02)
0	-0.04 (0.04)	-0.09 (0.02)	-0.07 (0.04)	-0.09 (0.05)

3.3.4 The Frication-noise Intensity Difference

The median and interquartile range of frication-noise intensity difference for 22 fricatives and affricates (11 consonants X 2 talker genders) processed with HA#1 at five input SNR conditions (+10, +5, 0, -5, and -10 dB SNR) is shown in Table 8 through Table 12, respectively. The frication-noise intensity difference for each consonant was measured under three different vowel contexts (/a, i, u/). Under each vowel context, the measurement was conducted four times (e.g., from four VCV tokens). Due to this small sample size under each vowel context, the values of the frication-noise intensity difference may not be normally distributed. Hence, the median values are reported. The interquartile range of frication-noise intensity difference for consonants processed by HA#1 was between 0.0 to 0.2 dB. The median and interquartile range of frication-noise intensity difference values for HA#1 were also illustrated in Appendix G.

Table 8. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at +10 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	0.5 (0.0)	0.7 (0.0)	0.1 (0.1)
Retroflex unaspirated affricate /tʂ/	0.1 (0.0)	0.2 (0.1)	0.0 (0.0)
Retroflex fricative /ʂ/	0.4 (0.0)	0.4 (0.1)	0.4 (0.0)
Alveolar aspirated affricate /tsʰ/	0.3 (0.0)	-0.5 (0.0)	0.3 (0.1)
Alveolar unaspirated affricate /ts/	-0.2 (0.1)	-0.2 (0.1)	0.3 (0.1)
Alveolar fricative /s/	-0.7 (0.0)	-0.5 (0.1)	-0.1 (0.1)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	0.7 (0.0)	0.7 (0.0)	0.5 (0.1)
Retroflex unaspirated affricate /tʂ/	-0.1 (0.1)	0.0 (0.1)	0.2 (0.0)
Retroflex fricative /ʂ/	0.7 (0.1)	0.6 (0.2)	0.5 (0.0)
Alveolar aspirated affricate /tsʰ/	-1.1 (0.1)	-2.0 (0.2)	-1.2 (0.1)
Alveolar unaspirated affricate /ts/	-0.2 (0.0)	-0.8 (0.1)	-0.3 (0.1)
Alveolar fricative /s/	-0.1 (0.1)	0.0 (0.0)	0.4 (0.1)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-0.2 (0.1)	-0.2 (0.0)	-0.1 (0.0)
Voiced post-alveolar affricate /ʤ/	-0.3 (0.0)	-0.3 (0.1)	-0.3 (0.0)
Voiceless post-alveolar fricative /ʃ/	-0.1 (0.0)	-0.2 (0.0)	-0.1 (0.0)
Voiceless alveolar fricative /s/	-1.0 (0.0)	-1.0 (0.1)	-0.7 (0.0)
Voiced alveolar fricative /z/	-0.4 (0.1)	0.1 (0.0)	-0.8 (0.1)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-0.1 (0.0)	-0.1 (0.0)	-0.1 (0.0)
Voiced post-alveolar affricate /ʤ/	-0.3 (0.1)	-0.2 (0.0)	-0.1 (0.0)
Voiceless post-alveolar fricative /ʃ/	-0.1 (0.0)	-0.1 (0.0)	-0.1 (0.0)
Voiceless alveolar fricative /s/	-0.1 (0.0)	-0.1 (0.0)	0.0 (0.0)
Voiced alveolar fricative /z/	-0.2 (0.3)	-0.2 (0.1)	-0.4 (0.1)

Table 9. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at +5 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-1.2 (0.1)	-0.4 (0.0)	-0.8 (0.1)
Retroflex unaspirated affricate /tʂ/	-1.5 (0.1)	-1.4 (0.1)	-1.3 (0.1)
Retroflex fricative /ʂ/	-1.3 (0.2)	-1.0 (0.0)	-0.9 (0.1)
Alveolar aspirated affricate /tsʰ/	-2.3 (0.1)	-3.5 (0.1)	-2.6 (0.2)
Alveolar unaspirated affricate /ts/	-1.8 (0.1)	-1.9 (0.1)	-1.6 (0.2)
Alveolar fricative /s/	-2.4 (0.1)	-2.4 (0.1)	-1.9 (0.1)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-0.4 (0.3)	-0.2 (0.1)	-0.3 (0.1)
Retroflex unaspirated affricate /tʂ/	-1.7 (0.1)	-1.7 (0.2)	-1.2 (0.1)
Retroflex fricative /ʂ/	-0.4 (0.1)	-0.5 (0.2)	-0.1 (0.0)
Alveolar aspirated affricate /tsʰ/	-1.1 (0.1)	-2.0 (0.2)	-1.2 (0.1)
Alveolar unaspirated affricate /ts/	-1.5 (0.1)	-2.4 (0.1)	-1.9 (0.2)
Alveolar fricative /s/	-1.9 (0.0)	-1.5 (0.0)	-1.3 (0.2)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-0.8 (0.0)	-0.8 (0.0)	-0.4 (0.1)
Voiced post-alveolar affricate /ʤ/	-1.3 (0.1)	-0.9 (0.1)	-0.8 (0.1)
Voiceless post-alveolar fricative /ʃ/	-0.8 (0.1)	-1.1 (0.1)	-0.7 (0.1)
Voiceless alveolar fricative /s/	-2.1 (0.2)	-1.9 (0.0)	-1.7 (0.1)
Voiced alveolar fricative /z/	-0.7 (0.0)	-0.1 (0.0)	-1.6 (0.1)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-0.3 (0.1)	-0.2 (0.0)	-0.4 (0.0)
Voiced post-alveolar affricate /ʤ/	-0.7 (0.2)	-0.5 (0.3)	-0.3 (0.1)
Voiceless post-alveolar fricative /ʃ/	-0.5 (0.1)	-0.4 (0.1)	-0.4 (0.1)
Voiceless alveolar fricative /s/	-0.4 (0.3)	-0.2 (0.2)	-0.3 (0.2)
Voiced alveolar fricative /z/	-1.5 (0.5)	-0.9 (0.2)	-1.2 (0.1)

Table 10. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at 0 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-2.4 (0.2)	-1.7 (0.1)	-2.3 (0.1)
Retroflex unaspirated affricate /tʂ/	-2.7 (0.1)	-2.6 (0.1)	-2.5 (0.0)
Retroflex fricative /ʂ/	-2.9 (0.1)	-2.5 (0.1)	-2.6 (0.1)
Alveolar aspirated affricate /tsʰ/	-2.5 (0.1)	-4.1 (0.1)	-2.7 (0.1)
Alveolar unaspirated affricate /ts/	-3.0 (0.4)	-3.1 (0.1)	-2.8 (0.2)
Alveolar fricative /s/	-3.2 (0.2)	-3.6 (0.2)	-3.3 (0.1)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-2.0 (0.0)	-1.8 (0.1)	-1.5 (0.0)
Retroflex unaspirated affricate /tʂ/	-3.0 (0.2)	-2.7 (0.1)	-2.2 (0.1)
Retroflex fricative /ʂ/	-2.0 (0.1)	-1.5 (0.1)	-0.7 (0.1)
Alveolar aspirated affricate /tsʰ/	-2.1 (0.1)	-3.1 (0.1)	-2.3 (0.1)
Alveolar unaspirated affricate /ts/	-1.0 (0.5)	-3.2 (0.0)	-2.8 (0.1)
Alveolar fricative /s/	-3.7 (0.1)	-2.7 (0.0)	-2.6 (0.3)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-2.5 (0.1)	-2.3 (0.0)	-1.8 (0.1)
Voiced post-alveolar affricate /ʤ/	-2.7 (0.1)	-2.3 (0.1)	-2.4 (0.1)
Voiceless post-alveolar fricative /ʃ/	-2.5 (0.1)	-2.6 (0.1)	-2.5 (0.1)
Voiceless alveolar fricative /s/	-3.2 (0.3)	-3.1 (0.2)	-3.1 (0.1)
Voiced alveolar fricative /z/	-1.5 (0.3)	-0.7 (0.2)	-2.6 (0.2)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-1.7 (0.1)	-1.4 (0.1)	-1.7 (0.1)
Voiced post-alveolar affricate /ʤ/	-2.3 (0.3)	-2.0 (0.3)	-1.8 (0.1)
Voiceless post-alveolar fricative /ʃ/	-1.9 (0.1)	-1.9 (0.1)	-2.0 (0.1)
Voiceless alveolar fricative /s/	-2.4 (0.2)	-1.9 (0.1)	-1.9 (0.3)
Voiced alveolar fricative /z/	-3.4 (0.2)	-2.5 (0.0)	-2.5 (0.2)

Table 11. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at -5 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-3.6 (0.1)	-3.4 (0.1)	-3.6 (0.1)
Retroflex unaspirated affricate /tʂ/	-3.8 (0.2)	-3.8 (0.2)	-3.7 (0.1)
Retroflex fricative /ʂ/	-4.0 (0.0)	-3.6 (0.1)	-4.2 (0.0)
Alveolar aspirated affricate /tsʰ/	-3.3 (0.2)	-4.9 (0.1)	-3.7 (0.2)
Alveolar unaspirated affricate /ts/	-3.9 (0.1)	-4.2 (0.4)	-3.9 (0.1)
Alveolar fricative /s/	-3.7 (0.8)	-4.5 (0.2)	-4.2 (0.3)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-3.8 (0.2)	-3.4 (0.2)	-3.1 (0.1)
Retroflex unaspirated affricate /tʂ/	-4.1 (0.2)	-3.8 (0.4)	-3.4 (0.1)
Retroflex fricative /ʂ/	-3.8 (0.1)	-2.8 (0.2)	-2.7 (0.1)
Alveolar aspirated affricate /tsʰ/	-2.9 (0.5)	-4.1 (0.2)	-3.5 (0.1)
Alveolar unaspirated affricate /ts/	-3.2 (0.9)	-4.4 (0.1)	-3.9 (0.1)
Alveolar fricative /s/	-5.0 (0.1)	-3.8 (0.1)	-4.0 (0.4)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-3.7 (0.1)	-3.8 (0.0)	-3.5 (0.1)
Voiced post-alveolar affricate /ʤ/	-3.8 (0.1)	-3.8 (0.0)	-3.6 (0.3)
Voiceless post-alveolar fricative /ʃ/	-4.0 (0.2)	-3.9 (0.3)	-4.1 (0.1)
Voiceless alveolar fricative /s/	-3.5 (0.7)	-3.5 (0.6)	-4.3 (0.1)
Voiced alveolar fricative /z/	-2.2 (0.1)	-1.7 (0.1)	-3.5 (0.3)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-3.5 (0.0)	-3.3 (0.4)	-3.3 (0.2)
Voiced post-alveolar affricate /ʤ/	-3.8 (0.2)	-3.6 (0.2)	-3.5 (0.1)
Voiceless post-alveolar fricative /ʃ/	-3.4 (0.1)	-3.4 (0.1)	-3.5 (0.1)
Voiceless alveolar fricative /s/	-4.5 (0.0)	-3.7 (0.2)	-3.9 (0.3)
Voiced alveolar fricative /z/	-5.0 (0.2)	-3.8 (0.3)	-3.8 (0.2)

Table 12. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at -10 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-4.6 (0.1)	-4.9 (0.1)	-4.7 (0.2)
Retroflex unaspirated affricate /tʂ/	-4.7 (0.3)	-4.7 (0.2)	-4.7 (0.2)
Retroflex fricative /ʂ/	-4.9 (0.1)	-4.6 (0.2)	-5.0 (0.1)
Alveolar aspirated affricate /tsʰ/	-4.2 (0.2)	-5.4 (0.3)	-4.7 (0.3)
Alveolar unaspirated affricate /ts/	-3.8 (0.8)	-4.7 (0.5)	-4.6 (0.3)
Alveolar fricative /s/	-3.9 (0.8)	-4.1 (1.2)	-4.1 (0.6)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-4.8 (0.2)	-4.7 (0.4)	-4.5 (0.2)
Retroflex unaspirated affricate /tʂ/	-4.7 (0.4)	-4.7 (0.5)	-4.5 (0.3)
Retroflex fricative /ʂ/	-4.8 (0.0)	-3.9 (0.2)	-4.6 (0.1)
Alveolar aspirated affricate /tsʰ/	-3.1 (0.4)	-4.8 (0.3)	-4.3 (0.4)
Alveolar unaspirated affricate /ts/	-2.2 (0.9)	-4.7 (1.1)	-4.6 (0.3)
Alveolar fricative /s/	-5.5 (0.5)	-4.6 (0.1)	-5.0 (0.4)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-4.8 (0.3)	-5.0 (0.1)	-4.7 (0.2)
Voiced post-alveolar affricate /ʤ/	-4.7 (0.3)	-4.9 (0.2)	-4.9 (0.2)
Voiceless post-alveolar fricative /ʃ/	-5.0 (0.1)	-4.9 (0.4)	-5.2 (0.2)
Voiceless alveolar fricative /s/	-2.3 (0.6)	-4.4 (0.4)	-5.4 (0.1)
Voiced alveolar fricative /z/	-3.8 (0.9)	-3.3 (0.5)	-4.6 (0.3)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-4.4 (0.0)	-4.5 (0.4)	-4.5 (0.1)
Voiced post-alveolar affricate /ʤ/	-4.7 (0.4)	-4.5 (0.3)	-4.6 (0.2)
Voiceless post-alveolar fricative /ʃ/	-4.6 (0.4)	-4.8 (0.2)	-4.6 (0.1)
Voiceless alveolar fricative /s/	-5.9 (0.1)	-5.1 (0.1)	-5.3 (0.1)
Voiced alveolar fricative /z/	-6.1 (0.1)	-5.3 (0.2)	-5.0 (0.3)

The median and interquartile range of frication-noise intensity difference for 22 fricatives and affricates (11 consonants X 2 talker genders) processed with HA#2 at three input SNR conditions (+10, +5, and 0 dB SNR) is shown in Table 13 through Table 15, respectively. The frication-noise intensity difference for each consonant was measured under three different vowel contexts (/a, i, u/). Under each vowel context, the measurement was conducted four times (e.g. from four VCV tokens). For HA#2, only three input SNR conditions were included. The data points for the -5 and -10 dB SNR conditions were excluded from the analysis because the retrieved-speech signals in these conditions were determined to have the poorest speech fidelity in the first study (Chapter 2).. The interquartile range of frication-noise intensity difference for consonants processed by HA#1 was between 0.1 to 0.4 dB. The median and interquartile range of frication-noise intensity difference values for HA#2 were also illustrated in Appendix H.

Table 13. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#2 at +10 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-0.9 (0.1)	0.2 (0.1)	0.0 (0.6)
Retroflex unaspirated affricate /tʂ/	-3.3 (0.2)	-1.3 (0.2)	-3.0 (0.2)
Retroflex fricative /ʂ/	-1.9 (0.1)	0.3 (0.3)	-0.1 (0.2)
Alveolar aspirated affricate /tsʰ/	-4.1 (0.2)	-1.8 (0.4)	-2.7 (0.1)
Alveolar unaspirated affricate /ts/	-4.8 (0.2)	-2.8 (0.3)	-3.5 (0.4)
Alveolar fricative /s/	-5.7 (0.1)	-1.9 (0.2)	-4.8 (0.2)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	0.9 (0.2)	0.4 (0.1)	0.0 (0.1)
Retroflex unaspirated affricate /tʂ/	-8.5 (0.6)	-5.7 (0.2)	-6.1 (0.1)
Retroflex fricative /ʂ/	5.5 (0.2)	6.4 (0.1)	6.1 (0.2)
Alveolar aspirated affricate /tsʰ/	-3.4 (0.3)	-0.5 (0.1)	-1.0 (0.1)
Alveolar unaspirated affricate /ts/	-4.8 (0.3)	-1.9 (0.1)	-3.4 (0.3)
Alveolar fricative /s/	0.0 (0.1)	0.6 (0.3)	-0.5 (0.1)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-1.2 (0.3)	0.0 (0.2)	-0.8 (0.4)
Voiced post-alveolar affricate /ʤ/	-5.4 (0.6)	-1.3 (0.7)	-1.8 (0.3)
Voiceless post-alveolar fricative /ʃ/	-0.1 (0.2)	0.6 (0.1)	-0.1 (0.2)
Voiceless alveolar fricative /s/	3.2 (0.5)	-1.0 (0.8)	0.3 (0.3)
Voiced alveolar fricative /z/	-5.7 (1.0)	-2.3 (0.2)	-5.3 (0.8)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-0.8 (0.3)	-0.1 (0.2)	-0.3 (0.1)
Voiced post-alveolar affricate /ʤ/	-2.1 (0.6)	-2.2 (0.2)	-1.5 (0.1)
Voiceless post-alveolar fricative /ʃ/	-1.0 (0.1)	-0.6 (0.2)	-0.6 (0.1)
Voiceless alveolar fricative /s/	0.2 (0.1)	0.1 (0.5)	0.3 (0.2)
Voiced alveolar fricative /z/	0.2 (0.2)	-0.9 (0.2)	-0.3 (0.1)

Table 14. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#2 at +5 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-3.5 (0.1)	-1.0 (0.5)	-1.6 (0.1)
Retroflex unaspirated affricate /tʂ/	-5.1 (0.0)	-2.7 (0.1)	-5.6 (0.2)
Retroflex fricative /ʂ/	-4.0 (0.3)	-1.1 (0.3)	-1.4 (0.3)
Alveolar aspirated affricate /tsʰ/	-5.9 (0.3)	-2.8 (0.1)	-5.1 (0.3)
Alveolar unaspirated affricate /ts/	-6.0 (0.1)	-4.5 (0.8)	-5.2 (0.7)
Alveolar fricative /s/	-8.1 (0.4)	-3.5 (0.2)	-7.3 (0.5)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	0.6 (0.5)	0.0 (0.3)	-0.5 (0.6)
Retroflex unaspirated affricate /tʂ/	-4.5 (0.3)	-0.5 (0.6)	-0.2 (0.1)
Retroflex fricative /ʂ/	-1.8 (0.4)	0.1 (0.1)	-0.1 (0.0)
Alveolar aspirated affricate /tsʰ/	-5.0 (0.3)	-0.6 (0.3)	-2.7 (0.2)
Alveolar unaspirated affricate /ts/	-6.3 (0.4)	-2.7 (0.2)	-4.6 (1.1)
Alveolar fricative /s/	-1.0 (0.4)	0.3 (0.4)	-1.3 (0.1)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-3.0 (0.4)	-0.5 (0.2)	-0.4 (0.3)
Voiced post-alveolar affricate /ʤ/	-7.6 (0.9)	-2.3 (0.2)	-3.3 (0.3)
Voiceless post-alveolar fricative /ʃ/	-0.6 (0.6)	0.1 (0.4)	0.0 (0.2)
Voiceless alveolar fricative /s/	2.1 (0.2)	1.9 (0.0)	1.7 (0.1)
Voiced alveolar fricative /z/	-7.8 (0.6)	-4.7 (0.5)	-7.9 (0.8)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-1.9 (0.7)	-0.2 (0.3)	-0.8 (0.3)
Voiced post-alveolar affricate /ʤ/	-3.3 (0.3)	-3.6 (0.7)	-2.9 (0.4)
Voiceless post-alveolar fricative /ʃ/	-1.9 (0.2)	-0.3 (0.3)	-1.2 (0.2)
Voiceless alveolar fricative /s/	0.1 (0.3)	-0.7 (0.2)	0.0 (0.3)
Voiced alveolar fricative /z/	0.1 (0.4)	-1.3 (0.5)	-0.8 (0.2)

Table 15. The median and interquartile range of frication-noise intensity difference (dB) between SMNR-on and SMNR-off for HA#2 at 0 dB SNR.

Consonant	Median (interquartile range)		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-5.8 (1.0)	-2.6 (0.2)	-3.2 (0.9)
Retroflex unaspirated affricate /tʂ/	-6.6 (0.2)	-4.1 (0.4)	-7.0 (0.6)
Retroflex fricative /ʂ/	-6.2 (0.6)	-1.9 (0.2)	-1.6 (0.7)
Alveolar aspirated affricate /tsʰ/	-7.2 (0.5)	-4.9 (0.2)	-6.4 (0.2)
Alveolar unaspirated affricate /ts/	-6.5 (1.4)	-5.9 (1.1)	-6.0 (0.2)
Alveolar fricative /s/	-8.3 (1.1)	-5.4 (0.1)	-7.9 (0.8)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	0.6 (0.1)	0.0 (0.2)	0.2 (0.4)
Retroflex unaspirated affricate /tʂ/	-7.1 (0.8)	-1.5 (0.8)	-1.1 (0.6)
Retroflex fricative /ʂ/	-2.3 (0.4)	0.0 (0.4)	-0.1 (0.5)
Alveolar aspirated affricate /tsʰ/	-7.0 (0.3)	-0.9 (0.7)	-3.4 (0.5)
Alveolar unaspirated affricate /ts/	-6.6 (0.6)	-3.1 (1.2)	-5.3 (0.5)
Alveolar fricative /s/	-1.0 (0.3)	0.1 (0.2)	-2.5 (0.2)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-3.9 (0.6)	-1.2 (0.3)	-0.7 (0.3)
Voiced post-alveolar affricate /ʤ/	-5.5 (0.8)	-2.2 (0.4)	-3.6 (0.3)
Voiceless post-alveolar fricative /ʃ/	1.0 (0.3)	1.5 (0.3)	1.4 (0.2)
Voiceless alveolar fricative /s/	-5.3 (1.4)	-3.1 (0.3)	-3.6 (0.9)
Voiced alveolar fricative /z/	-5.1 (0.7)	-3.9 (0.4)	-6.9 (0.6)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-1.5 (0.8)	0.2 (0.5)	-0.7 (0.8)
Voiced post-alveolar affricate /ʤ/	-2.7 (0.6)	-2.6 (1.1)	-2.4 (0.5)
Voiceless post-alveolar fricative /ʃ/	1.3 (0.8)	1.6 (0.3)	0.4 (0.3)
Voiceless alveolar fricative /s/	0.2 (0.1)	0.2 (0.6)	1.3 (0.2)
Voiced alveolar fricative /z/	0.3 (0.9)	-0.8 (0.4)	-0.8 (0.4)

3.3.5 Release-burst Intensity Difference

The median and interquartile range of release-burst intensity difference for 12 affricates (6 consonants X 2 talker genders) processed with HA#1 at five input SNR conditions (+10, +5, 0, -5 , and -10 dB SNR) is shown in Table 16 through Table 18 respectively. Similar to the frication-noise intensity difference, the release-burst intensity difference for each consonant was measured under three different vowel contexts (/a, i, u/). Under each vowel context, the measurement was conducted four times (e.g., from four VCV tokens). Release bursts could not be identified for the (i) voiceless post-alveolar affricate, spoken by the English female talker, in the vowel /u/ context at all input SNR conditions; (ii) aspirated alveolar affricate, spoken by the Mandarin male talker, in the vowel /a, i/ contexts at -10 dB input SNR condition; and (iii) unaspirated alveolar affricate, spoken by the Mandarin male talker, in the vowel /a/ context at -10 dB input SNR condition. The typical interquartile range of release-burst intensity difference for consonants processed by HA#1 was between 0.0 to 0.3 dB. The median and interquartile range of release-burst intensity difference values for HA#1 were also illustrated in Appendix I.

Table 16. The median and interquartile range of release-burst intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at +10 and +5 dB SNR.

	Median (interquartile range)		
Consonant	Vowel /a/	Vowel /i/	Vowel /u/
+10 dB SNR			
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	0.4 (0.0)	0.5 (0.0)	0.1 (0.0)
Retroflex unaspirated affricate /tʂ/	0.1 (0.1)	0.1 (0.0)	0.2 (0.0)
Alveolar aspirated affricate /tsʰ/	0.1 (0.1)	0.3 (0.1)	0.4 (0.1)
Alveolar unaspirated affricate /ts/	0.0 (0.1)	0.0 (0.1)	0.4 (0.1)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	0.7 (0.0)	0.6 (0.0)	0.4 (0.1)
Retroflex unaspirated affricate /tʂ/	-0.1 (0.1)	0.0 (0.1)	0.2 (0.1)
Alveolar aspirated affricate /tsʰ/	-1.3 (0.3)	-2.1 (0.1)	-1.9 (0.3)
Alveolar unaspirated affricate /ts/	-0.2 (0.2)	-0.4 (0.2)	-0.3 (0.0)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-0.2 (0.1)	-0.4 (0.0)	n/a
Voiced post-alveolar affricate /ʤ/	-0.4 (0.0)	-0.4 (0.1)	-0.1 (0.0)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-0.1 (0.0)	-0.1 (0.0)	-0.1 (0.0)
Voiced post-alveolar affricate /ʤ/	-0.4 (0.1)	-0.1 (0.0)	-0.2 (0.0)
+5 dB SNR			
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-1.3 (0.1)	-0.5 (0.1)	-0.8 (0.1)
Retroflex unaspirated affricate /tʂ/	-1.4 (0.1)	-1.4 (0.1)	-0.8 (0.1)
Alveolar aspirated affricate /tsʰ/	-2.2 (0.3)	-2.6 (0.6)	-2.7 (0.3)
Alveolar unaspirated affricate /ts/	-1.4 (0.3)	-1.3 (0.2)	-1.5 (0.2)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-0.6 (0.4)	-0.4 (0.2)	-0.6 (0.0)
Retroflex unaspirated affricate /tʂ/	-1.8 (0.0)	-1.7 (0.2)	-1.2 (0.2)
Alveolar aspirated affricate /tsʰ/	-1.3 (0.3)	-2.1 (0.1)	-1.9 (0.3)
Alveolar unaspirated affricate /ts/	-1.0 (0.2)	-2.0 (0.2)	-1.8 (0.1)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-1.0 (0.2)	-1.3 (0.0)	n/a
Voiced post-alveolar affricate /ʤ/	-1.3 (0.1)	-1.0 (0.1)	-0.4 (0.1)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-0.2 (0.1)	-0.2 (0.0)	-0.4 (0.0)
Voiced post-alveolar affricate /ʤ/	-1.0 (0.2)	-0.3 (0.3)	-0.5 (0.1)

Note: n/a = release burst could not be identified.

Table 17. The median and interquartile range of release-burst intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at 0 and -5 dB SNR.

Consonant	Median (interquartile range)		
	Vowel /a/	Vowel /i/	Vowel /u/
0 dB SNR			
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-2.5 (0.1)	-1.8 (0.1)	-2.1 (0.1)
Retroflex unaspirated affricate /tʂ/	-2.4 (0.1)	-2.4 (0.2)	-1.8 (0.1)
Alveolar aspirated affricate /tsʰ/	-2.5 (0.4)	-2.7 (1.5)	-3.0 (0.2)
Alveolar unaspirated affricate /ts/	-3.0 (1.0)	-2.2 (0.3)	-2.6 (0.1)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-2.3 (0.1)	-2.2 (0.1)	-1.7 (0.1)
Retroflex unaspirated affricate /tʂ/	-2.4 (0.1)	-2.4 (0.2)	-1.8 (0.1)
Alveolar aspirated affricate /tsʰ/	-2.1 (0.5)	-3.1 (0.4)	-2.9 (0.2)
Alveolar unaspirated affricate /ts/	-0.1 (0.7)	-2.7 (0.1)	-2.7 (0.4)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-2.8 (0.3)	-2.7 (0.1)	n/a
Voiced post-alveolar affricate /ʤ/	-2.9 (0.2)	-2.5 (0.0)	-2.0 (0.1)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-1.6 (0.1)	-1.7 (0.1)	-1.9 (0.1)
Voiced post-alveolar affricate /ʤ/	-2.6 (0.3)	-1.9 (0.2)	-1.9 (0.1)
-5 dB SNR			
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-3.6 (0.2)	-3.0 (0.0)	-3.2 (0.2)
Retroflex unaspirated affricate /tʂ/	-3.4 (0.2)	-3.5 (0.3)	-2.9 (0.0)
Alveolar aspirated affricate /tsʰ/	-2.4 (0.5)	-2.5 (0.6)	-3.8 (0.2)
Alveolar unaspirated affricate /ts/	-3.3 (0.8)	-2.6 (0.2)	-3.6 (0.1)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-4.2 (0.1)	-4.1 (0.3)	-3.3 (0.1)
Retroflex unaspirated affricate /tʂ/	-4.0 (0.2)	-3.7 (0.3)	-3.2 (0.1)
Alveolar aspirated affricate /tsʰ/	-5.1 (0.7)	-4.3 (0.9)	-4.0 (0.1)
Alveolar unaspirated affricate /ts/	-2.8 (1.2)	-3.9 (0.1)	-3.5 (0.5)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-4.5 (0.6)	-3.9 (0.1)	n/a
Voiced post-alveolar affricate /ʤ/	-4.0 (0.2)	-3.4 (0.5)	-3.5 (0.3)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-3.5 (0.1)	-3.6 (0.4)	-3.5 (0.2)
Voiced post-alveolar affricate /ʤ/	-3.9 (0.1)	-3.6 (0.2)	-3.5 (0.1)

Note: n/a = release burst could not be identified.

Table 18. The median and interquartile range of release-burst intensity difference (dB) between SMNR-on and SMNR-off for HA#1 at -10 dB SNR.

Consonant	Median (interquartile range)		
	Vowel /a/	Vowel /i/	Vowel /u/
-10 dB SNR			
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-4.4 (0.1)	-4.3 (0.7)	-4.6 (1.0)
Retroflex unaspirated affricate /tʂ/	-4.2 (0.3)	-4.3 (0.2)	-3.4 (1.0)
Alveolar aspirated affricate /tsʰ/	-3.2 (1.6)	-3.0 (0.7)	-4.6 (0.7)
Alveolar unaspirated affricate /ts/	-2.9 (0.3)	-2.8 (2.1)	-4.6 (0.2)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-5.1(0.3)	-5.0 (0.3)	-4.4 (0.3)
Retroflex unaspirated affricate /tʂ/	-4.1 (0.8)	-4.7 (0.5)	-4.2 (0.1)
Alveolar aspirated affricate /tsʰ/	n/a	n/a	-4.7 (0.4)
Alveolar unaspirated affricate /ts/	n/a	-3.9 (1.9)	-4.2 (0.5)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-5.1 (0.3)	-5.1 (0.2)	n/a
Voiced post-alveolar affricate /ʤ/	-3.7 (0.6)	-2.5 (0.5)	-4.8 (0.3)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-4.8 (0.4)	-4.7 (0.5)	-4.6 (0.1)
Voiced post-alveolar affricate /ʤ/	-4.8 (0.1)	-4.5 (0.3)	-4.6 (0.3)

Note: n/a = release burst could not be identified.

The median and interquartile range of release-burst intensity difference for 12 affricates (6 consonants X 2 talker genders) processed with HA#2 at three input SNR conditions (+10, +5, and 0 dB) is shown in Table 19 through Table 20, respectively. Release burst could not be identified for the (i) alveolar unaspirated affricate /ts/, spoken by the Mandarin male talker, in the vowel /a/ context at 0 dB SNR condition; (ii) voiceless post-alveolar affricate /tʃ/, spoken by the English female talker, in the vowel /u/ context at +5 and 0 dB SNR conditions; and (iii) voiced post-alveolar affricate /ʤ/, spoken by the English male talker, in the vowel /a/ context at 0 dB SNR condition. The typical interquartile range of release-burst intensity difference for consonants processed by HA#2 was between 0.2 to 0.4 dB at positive SNR conditions and between 1.0 to 1.4 dB at the 0 dB input SNR condition.

The median and interquartile range of release-burst intensity difference values for HA#2 were also illustrated in Appendix J.

Table 19. The median and interquartile range of release-burst intensity difference (dB) between SMNR-on and SMNR-off for HA#2 at +10 and +5 dB SNR.

	Median (interquartile range)		
Consonant	Vowel /a/	Vowel /i/	Vowel /u/
+10 dB SNR			
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-4.9 (0.2)	-4.5 (0.2)	-3.5 (0.2)
Retroflex unaspirated affricate /tʂ/	-5.6 (0.2)	-5.4 (0.3)	-6.2 (0.2)
Alveolar aspirated affricate /tsʰ/	-6.5 (0.3)	-7.1 (0.3)	-6.1 (0.6)
Alveolar unaspirated affricate /ts/	-6.9 (0.3)	-5.5 (0.7)	-5.2 (0.5)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-1.6 (0.3)	-3.6 (0.1)	-3.0 (0.1)
Retroflex unaspirated affricate /tʂ/	-10.3 (0.5)	-8.3 (0.3)	-8.7 (0.2)
Alveolar aspirated affricate /tsʰ/	-5.4 (0.3)	-4.6 (0.4)	-4.5 (0.3)
Alveolar unaspirated affricate /ts/	-5.1 (0.4)	-7.5 (0.1)	-6.8 (0.6)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-5.1(0.4)	-4.4 (0.5)	-5.7 (0.6)
Voiced post-alveolar affricate /ʤ/	-6.4 (0.6)	-7.0 (0.5)	-4.8 (0.5)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-3.7 (0.6)	-3.3 (0.9)	-3.6 (0.7)
Voiced post-alveolar affricate /ʤ/	-4.5 (0.4)	-3.5 (0.4)	-3.5 (0.3)
+5 dB SNR			
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-5.8 (0.4)	-6.1 (0.4)	-5.7 (0.3)
Retroflex unaspirated affricate /tʂ/	-6.5 (0.4)	-7.0 (0.2)	-7.5 (0.4)
Alveolar aspirated affricate /tsʰ/	-7.9 (0.4)	-8.4 (1.3)	-7.0 (0.5)
Alveolar unaspirated affricate /ts/	-6.9 (1.2)	-7.4 (1.6)	-5.8 (0.2)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-3.0 (0.3)	-4.5 (0.2)	-4.1 (0.2)
Retroflex unaspirated affricate /tʂ/	-5.3 (0.2)	-4.0 (0.1)	-3.8 (0.4)
Alveolar aspirated affricate /tsʰ/	-5.3 (1.3)	-6.0 (0.8)	-6.0 (0.4)
Alveolar unaspirated affricate /ts/	-6.6 (1.7)	-8.5 (0.2)	-6.7 (0.7)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-6.0 (0.9)	-6.9 (0.2)	n/a
Voiced post-alveolar affricate /ʤ/	-7.5 (0.3)	-8.1 (0.3)	-5.7 (0.4)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-3.7 (0.3)	-4.0 (0.7)	-3.1 (0.6)
Voiced post-alveolar affricate /ʤ/	-5.2 (0.4)	-4.1 (1.0)	-4.6 (0.5)

Note: n/a = release burst could not be identified.

Table 20. The median and interquartile range of release-burst intensity difference (dB) between SMNR-on and SMNR-off for HA#2 at 0 dB SNR.

	Median (interquartile range)		
Consonant	Vowel /a/	Vowel /i/	Vowel /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-7.0 (0.6)	-7.7 (0.3)	-7.3 (0.2)
Retroflex unaspirated affricate /tʂ/	-6.8 (1.4)	-7.5 (0.7)	-7.7 (1.3)
Alveolar aspirated affricate /tsʰ/	-7.6 (0.6)	-6.2 (2.7)	-8.2 (1.3)
Alveolar unaspirated affricate /ts/	-3.4 (2.1)	-6.9 (1.5)	-5.5 (0.6)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-3.3 (0.5)	-4.9 (0.2)	-3.8 (0.6)
Retroflex unaspirated affricate /tʂ/	-6.9 (0.6)	-5.8 (0.5)	-2.5 (0.2)
Alveolar aspirated affricate /tsʰ/	-6.6 (1.0)	-7.9 (0.5)	-6.4 (0.8)
Alveolar unaspirated affricate /ts/	n/a	-7.8 (0.6)	-6.6 (0.6)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-5.1 (0.4)	-6.1 (0.7)	n/a
Voiced post-alveolar affricate /ʤ/	-5.5 (1.1)	-5.0 (0.8)	-4.5 (1.1)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-2.7 (0.4)	-1.0 (1.2)	-2.5 (1.3)
Voiced post-alveolar affricate /ʤ/	n/a	-2.8 (0.8)	-2.9 (1.4)

Note: n/a = release burst could not be identified.

3.3.6 Spectral Mean Difference

Preliminary spectral mean measurements for 22 fricatives and affricates processed with HA#1 at five input SNR conditions (+10, +5, 0, -5, and -10 dB SNR) showed that some of the spectral mean values for the English female voiced and voiceless alveolar fricative (/s, z/) were less than 1000 Hz. Because frication noise energy is expected to be at higher frequency regions above 1000 Hz, the frication noise sections of these consonants were high-pass filtered at a cut-off frequency of 1000 Hz and the spectral mean of these consonants were re-measured.; this procedure was comparable to the procedure used by Kortekaas and Stelmachowicz (2000)where the spectral moments of the fricative /s/ were calculated for frequency bins above 1000. If the spectral mean values changed after being high-pass filtered, the new spectral mean values were taken. In total, 72 out of 1320 tokens (22 consonants X 3 vowel contexts X 5 SNRs X 4 tokens) were re-measured.

For HA#2, only three input SNR conditions (+10, +5, and 0 dB SNR) were included in the analysis. Preliminary spectral mean measurements showed that some of the spectral mean values for the Mandarin alveolar affricates and consonants (/tsʰ, ts, s/; written as /c, z, s/ in *PinYin*) and English alveolar fricatives (/s, z/) were over 7000 Hz. These values were considered to be measurement errors because the hearing aid response began to roll-off beyond 7000 Hz, as shown in Figure 18. Thus, the frication noise sections of these consonants were low-pass filtered at a cut-off frequency of 7000 Hz and the spectral mean values were re-measured. If the spectral mean values changed after being low-pass filtered, the new spectral mean values were taken. In total, 276 out of 792 tokens (22 consonants X 3 vowel contexts X 3 SNRs X 4 tokens) were re-measured.

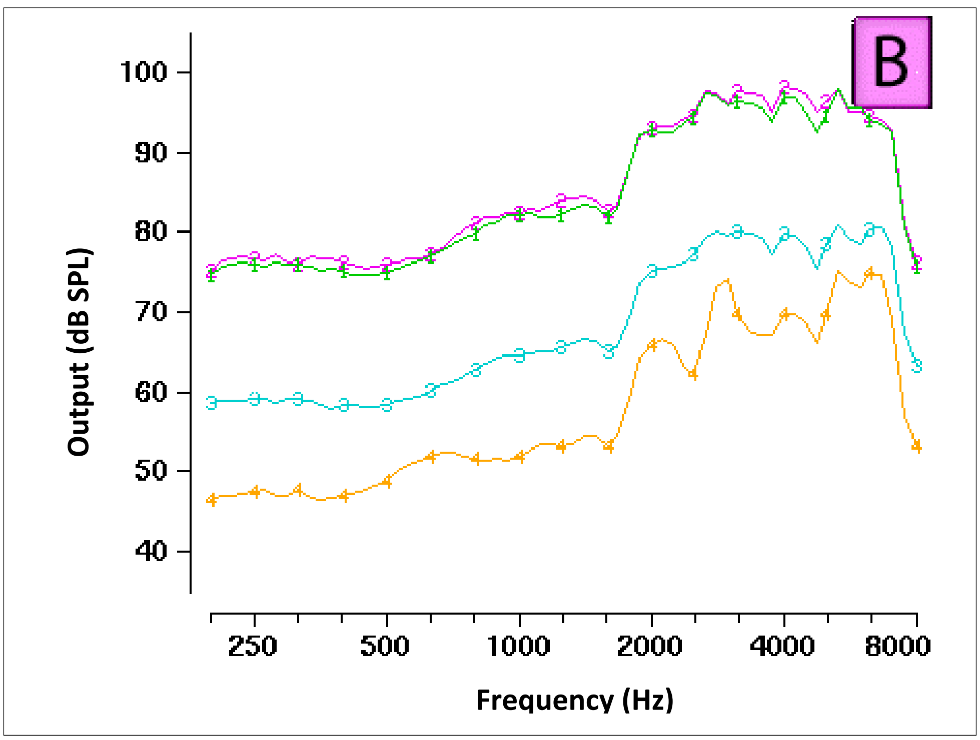


Figure 18. Frequency response of HA#2. The green (SMNR-off) and pink (SMNR-on) curves were obtained with a pure-tone sweep at 60 dB SPL. The blue (SMNR-off) and orange (SMNR-on) curves were obtained with a pink-noise at 60 dB SPL.

The median and interquartile range of spectral mean difference for 22 fricatives and affricates (11 consonants X 2 talker genders) processed with HA#1 at five input SNR conditions (+10 , +5 , 0 , -5 , and -10 dB SNR) is shown in Table 21 to Table 25. The median and interquartile range of spectral mean difference for 22 fricatives and affricates processed with HA#2 at three input SNR conditions (+10, +5, and 0 dB) is shown in Table 26 to Table 28. The spectral mean difference for each consonant was measured under three different vowel contexts (/a, i, u/). Under each vowel context, the measurement was conducted four times. The typical interquartile range of spectral mean difference for consonants processed by HA#1 was between 0 to 20 Hz for the positive SNR conditions, 0 to 30 Hz for the 0 and -5 dB input SNR conditions, and 0 to 40 Hz for the -10 dB SNR condition. The typical interquartile range of spectral mean difference for consonants processed by HA#2 was between 0 to 20 Hz for the positive input SNR conditions.

The values of these tables were also illustrated in the figures of Appendix K and Appendix L. The diamond shape marker represents the median value and the bar represents the interquartile range of these four measurements. The y-axes of the figures were scaled to the same range for comparisons. Positive values on the y-axis indicate spectral mean increment whereas negative values indicate spectral mean reduction when SMNR was turned on. The spectral mean values for each consonant (averaged across vowel contexts) measured in the SMNR-on and SMNR-off conditions are tabulated in Appendix M.

Table 21. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#1 at +10 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	35 (2)	-15 (3)	-200 (25)
Retroflex unaspirated affricate /tʂ/	-5 (4)	-19 (1)	-258 (8)
Retroflex fricative /ʂ/	30 (5)	-32 (10)	-24 (2)
Alveolar aspirated affricate /tsʰ/	-77 (7)	-296 (15)	-52 (8)
Alveolar unaspirated affricate /ts/	-123 (21)	-225 (18)	-20 (3)
Alveolar fricative /s/	-203 (5)	-279 (11)	-210 (26)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-268 (59)	631 (15)	-403 (3)
Retroflex unaspirated affricate /tʂ/	-24 (6)	-61 (6)	-55 (21)
Retroflex fricative /ʂ/	7 (3)	-5 (17)	-36 (4)
Alveolar aspirated affricate /tsʰ/	-25 (14)	-74 (21)	-7 (5)
Alveolar unaspirated affricate /ts/	-153 (27)	-330 (14)	-9 (3)
Alveolar fricative /s/	-49 (16)	-3 (5)	-7 (1)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-18 (11)	-46 (9)	-5 (9)
Voiced post-alveolar affricate /ʤ/	-73 (6)	-118 (11)	-91 (2)
Voiceless alveolar fricative /s/	-261 (68)	-292 (7)	-60 (6)
Voiced alveolar fricative /z/	-281 (28)	-134 (6)	-232 (11)
Voiceless post-alveolar fricative /ʃ/	8 (5)	9 (5)	7 (1)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-3 (10)	-1 (2)	3 (4)
Voiced post-alveolar affricate /ʤ/	-76 (13)	-42 (6)	-13 (1)
Voiceless alveolar fricative /s/	4 (3)	0 (1)	4 (3)
Voiced alveolar fricative /z/	19 (3)	2 (34)	7 (12)
Voiceless post-alveolar fricative /ʃ/	-21 (5)	-32 (5)	-16 (6)

Table 22. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#1 at +5 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-88 (17)	-68 (1)	-543 (50)
Retroflex unaspirated affricate /tʂ/	-29 (6)	-39 (3)	-434 (13)
Retroflex fricative /ʂ/	-19 (9)	-127 (12)	-160 (32)
Alveolar aspirated affricate /tsʰ/	-209 (12)	-219 (8)	-78 (6)
Alveolar unaspirated affricate /ts/	-186 (18)	-290 (53)	-29 (8)
Alveolar fricative /s/	-210 (61)	-327 (29)	-208 (18)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-17 (4)	-91 (10)	7 (13)
Retroflex unaspirated affricate /tʂ/	-37 (1)	-93 (7)	-136 (31)
Retroflex fricative /ʂ/	27 (9)	17 (31)	11 (16)
Alveolar aspirated affricate /tsʰ/	-172 (40)	-296 (34)	-133 (6)
Alveolar unaspirated affricate /ts/	-231 (63)	-384 (23)	-31 (2)
Alveolar fricative /s/	-66 (25)	-24 (8)	-17 (8)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-163 (12)	-131 (15)	-94 (11)
Voiced post-alveolar affricate /ʤ/	-92 (11)	-215 (8)	-167 (13)
Voiceless alveolar fricative /s/	-381 (38)	-411 (47)	-129 (9)
Voiced alveolar fricative /z/	-510 (82)	-175 (10)	-326 (53)
Voiceless post-alveolar fricative /ʃ/	32 (16)	-14 (16)	33 (11)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-74 (3)	-39 (2)	-41 (6)
Voiced post-alveolar affricate /ʤ/	-166 (26)	-96 (7)	-29 (0)
Voiceless alveolar fricative /s/	-13 (6)	4 (7)	22 (10)
Voiced alveolar fricative /z/	13 (26)	-128 (44)	-45 (18)
Voiceless post-alveolar fricative /ʃ/	-84 (11)	-117 (16)	-83 (13)

Table 23. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#1 at 0 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-120 (10)	-98 (5)	-854 (16)
Retroflex unaspirated affricate /tʂ/	-32 (3)	-49 (4)	-483 (23)
Retroflex fricative /ʂ/	-37 (9)	-220 (5)	-370 (25)
Alveolar aspirated affricate /tsʰ/	-291 (13)	-235 (48)	-116 (9)
Alveolar unaspirated affricate /ts/	-130 (56)	-269 (51)	-26 (5)
Alveolar fricative /s/	-235 (30)	-399 (56)	-186 (30)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-54 (14)	-179 (49)	-189 (39)
Retroflex unaspirated affricate /tʂ/	-36 (4)	-104 (7)	-188 (45)
Retroflex fricative /ʂ/	-26 (4)	-55 (13)	-76 (61)
Alveolar aspirated affricate /tsʰ/	-307 (14)	-384 (46)	-181 (3)
Alveolar unaspirated affricate /ts/	-849 (25)	-473 (61)	-51 (14)
Alveolar fricative /s/	-247 (28)	-133 (12)	-33 (8)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-229 (22)	-132 (9)	-164 (17)
Voiced post-alveolar affricate /ʤ/	-85 (23)	-217 (9)	-197 (18)
Voiceless alveolar fricative /s/	-321 (104)	-414 (38)	-118 (15)
Voiced alveolar fricative /z/	-679 (147)	-164 (32)	-420 (60)
Voiceless post-alveolar fricative /ʃ/	-36 (28)	-90 (18)	-110 (18)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-122 (14)	-73 (7)	-131 (2)
Voiced post-alveolar affricate /ʤ/	-168 (22)	-111 (5)	-33 (2)
Voiceless alveolar fricative /s/	11 (17)	-5 (16)	23 (29)
Voiced alveolar fricative /z/	-18 (36)	-262 (27)	-84 (20)
Voiceless post-alveolar fricative /ʃ/	-121 (9)	-157 (11)	-138 (27)

Table 24. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#1 at -5 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-135 (41)	-87 (6)	-834 (25)
Retroflex unaspirated affricate /tʂ/	-33 (4)	-48 (7)	-418 (18)
Retroflex fricative /ʂ/	-43 (5)	-241 (19)	-428 (24)
Alveolar aspirated affricate /tsʰ/	-264 (20)	-261 (35)	-99 (25)
Alveolar unaspirated affricate /ts/	-229 (25)	-318 (78)	-31 (11)
Alveolar fricative /s/	-470 (227)	-473 (65)	-199 (33)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-82 (21)	-173 (17)	-263 (33)
Retroflex unaspirated affricate /tʂ/	-44 (8)	-95 (21)	-249 (33)
Retroflex fricative /ʂ/	-48 (19)	-126 (27)	-277 (62)
Alveolar aspirated affricate /tsʰ/	-326 (57)	-460 (25)	-209 (25)
Alveolar unaspirated affricate /ts/	-529 (227)	-456 (41)	-48 (4)
Alveolar fricative /s/	-243 (53)	-226 (7)	-40 (8)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-189 (39)	-114 (11)	-164 (34)
Voiced post-alveolar affricate /ʤ/	-71 (21)	-162 (10)	-184 (36)
Voiceless alveolar fricative /s/	-1130 (393)	-473 (65)	-199 (33)
Voiced alveolar fricative /z/	-726 (52)	-182 (58)	-356 (49)
Voiceless post-alveolar fricative /ʃ/	-86 (53)	-120 (12)	-177 (13)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-110 (17)	-56 (16)	-156 (12)
Voiced post-alveolar affricate /ʤ/	-124 (26)	-83 (6)	-26 (2)
Voiceless alveolar fricative /s/	16 (14)	-1 (19)	48 (43)
Voiced alveolar fricative /z/	-58 (24)	-537 (56)	-193 (25)
Voiceless post-alveolar fricative /ʃ/	-107 (10)	-129 (16)	-128 (21)

Table 25. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#1 at -10 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-103 (19)	-66 (7)	-670 (81)
Retroflex unaspirated affricate /tʂ/	-19 (11)	-40 (2)	-290 (66)
Retroflex fricative /ʂ/	-19 (9)	-181 (20)	-345 (16)
Alveolar aspirated affricate /tsʰ/	-200 (33)	-339 (69)	-72 (17)
Alveolar unaspirated affricate /ts/	-212 (138)	-186 (69)	-13 (10)
Alveolar fricative /s/	-179 (72)	-358 (120)	-114 (100)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	-75 (12)	-143 (36)	-293 (30)
Retroflex unaspirated affricate /tʂ/	-46 (26)	-81 (19)	-202 (44)
Retroflex fricative /ʂ/	-72 (8)	-146 (10)	-402 (108)
Alveolar aspirated affricate /tsʰ/	-271 (23)	-432 (33)	-175 (13)
Alveolar unaspirated affricate /ts/	-849 (609)	-403 (80)	-37 (23)
Alveolar fricative /s/	-476 (267)	-202 (13)	-19 (10)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	-159 (13)	-114 (23)	-173 (16)
Voiced post-alveolar affricate /ʤ/	-21 (65)	-147 (38)	-160 (39)
Voiceless alveolar fricative /s/	-956 (252)	-524 (298)	-109 (57)
Voiced alveolar fricative /z/	-473 (331)	-133 (112)	-265 (182)
Voiceless post-alveolar fricative /ʃ/	-109 (28)	-125 (42)	-221 (35)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	-87 (13)	-40 (27)	-138 (35)
Voiced post-alveolar affricate /ʤ/	-139 (38)	-106 (23)	-26 (10)
Voiceless alveolar fricative /s/	-101 (74)	-64 (42)	-219 (48)
Voiced alveolar fricative /z/	-89 (48)	-314 (36)	-143 (45)
Voiceless post-alveolar fricative /ʃ/	-90 (11)	-106 (21)	-132 (4)

Table 26. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#2 at +10 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-395 (80)	-319 (9)	55 (4)
Retroflex unaspirated affricate /tʂ/	-347 (26)	-400 (96)	-230 (17)
Retroflex fricative /ʂ/	-531 (76)	-200 (19)	6 (5)
Alveolar aspirated affricate /tsʰ/	-245 (58)	-44 (3)	-314 (29)
Alveolar unaspirated affricate /ts/	-322 (100)	-137 (14)	-1095 (76)
Alveolar fricative /s/	-32 (24)	16 (19)	-154 (44)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	82 (21)	-55 (29)	28 (19)
Retroflex unaspirated affricate /tʂ/	-696 (119)	34 (17)	90 (30)
Retroflex fricative /ʂ/	-28 (20)	-52 (8)	56 (8)
Alveolar aspirated affricate /tsʰ/	36 (20)	162 (11)	-147 (19)
Alveolar unaspirated affricate /ts/	17 (19)	108 (3)	-427 (65)
Alveolar fricative /s/	87 (5)	250 (13)	-952 (105)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	231 (44)	-31 (4)	214 (5)
Voiced post-alveolar affricate /ʤ/	-178 (40)	221 (19)	-45 (31)
Voiceless alveolar fricative /s/	-44 (11)	71 (9)	-301 (10)
Voiced alveolar fricative /z/	-220 (67)	-278 (44)	-381 (31)
Voiceless post-alveolar fricative /ʃ/	-117 (16)	-35 (6)	-1 (17)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	137 (77)	7 (24)	-260 (18)
Voiced post-alveolar affricate /ʤ/	-139 (127)	-738 (45)	-101 (40)
Voiceless alveolar fricative /s/	2 (13)	-186 (18)	-28 (3)
Voiced alveolar fricative /z/	-16 (7)	-157 (24)	86 (41)
Voiceless post-alveolar fricative /ʃ/	-430 (30)	-360 (9)	-260 (13)

Table 27. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#2 at +5 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-445 (94)	-492 (40)	57 (7)
Retroflex unaspirated affricate /tʂ/	-380 (76)	-359 (200)	-334 (19)
Retroflex fricative /ʂ/	-406 (67)	-395 (34)	-7 (27)
Alveolar aspirated affricate /tsʰ/	-344 (8)	32 (32)	-309 (97)
Alveolar unaspirated affricate /ts/	-285 (149)	-253 (32)	-773 (59)
Alveolar fricative /s/	-159 (40)	-7 (37)	-255 (87)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	110 (15)	-39 (21)	143 (3)
Retroflex unaspirated affricate /tʂ/	-482 (249)	26 (28)	207 (20)
Retroflex fricative /ʂ/	-245 (36)	-38 (4)	65 (5)
Alveolar aspirated affricate /tsʰ/	-14 (18)	209 (12)	-438 (18)
Alveolar unaspirated affricate /ts/	-91 (120)	126 (4)	-148 (115)
Alveolar fricative /s/	102 (14)	445 (39)	-1040 (80)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	314 (69)	3 (8)	441 (8)
Voiced post-alveolar affricate /ʤ/	-262 (63)	195 (66)	-72 (105)
Voiceless alveolar fricative /s/	-178 (42)	14 (18)	-353 (171)
Voiced alveolar fricative /z/	-245 (39)	-334 (64)	-277 (54)
Voiceless post-alveolar fricative /ʃ/	169 (13)	28 (27)	110 (15)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	40 (74)	-33 (64)	-202 (29)
Voiced post-alveolar affricate /ʤ/	-223 (57)	-720 (185)	-249 (92)
Voiceless alveolar fricative /s/	11 (12)	-118 (11)	94 (20)
Voiced alveolar fricative /z/	26 (2)	-166 (24)	192 (44)
Voiceless post-alveolar fricative /ʃ/	-612 (99)	-362 (102)	-198 (23)

Table 28. The median and interquartile range of spectral mean difference (Hz) between SMNR-on and SMNR-off for HA#2 at 0 dB SNR.

Consonant	Median (interquartile range); Hz		
	Vowel context /a/	Vowel context /i/	Vowel context /u/
Mandarin Female Talker			
Retroflex aspirated affricate /tʂʰ/	-462 (290)	-596 (65)	112 (16)
Retroflex unaspirated affricate /tʂ/	-343 (109)	-190 (167)	-323 (69)
Retroflex fricative /ʂ/	-306 (236)	-518 (334)	138 (45)
Alveolar aspirated affricate /tsʰ/	-378 (21)	-37 (22)	-364 (30)
Alveolar unaspirated affricate /ts/	-524 (78)	-242 (76)	-438 (172)
Alveolar fricative /s/	-95 (76)	-123 (83)	-195 (121)
Mandarin Male Talker			
Retroflex aspirated affricate /tʂʰ/	125 (71)	-29 (10)	247 (33)
Retroflex unaspirated affricate /tʂ/	-575 (327)	-73 (85)	321 (59)
Retroflex fricative /ʂ/	-70 (110)	-49 (15)	96 (38)
Alveolar aspirated affricate /tsʰ/	-136 (73)	274 (32)	-211 (109)
Alveolar unaspirated affricate /ts/	-130 (83)	112 (24)	97 (116)
Alveolar fricative /s/	114 (16)	639 (51)	-1139 (201)
English Female Talker			
Voiceless post-alveolar affricate /ʧ/	261 (45)	135 (15)	603 (27)
Voiced post-alveolar affricate /ʤ/	-21 (120)	-61 (93)	50 (123)
Voiceless alveolar fricative /s/	535 (354)	384 (73)	202 (16)
Voiced alveolar fricative /z/	-261 (446)	-194 (159)	81 (99)
Voiceless post-alveolar fricative /ʃ/	460 (48)	357 (95)	357 (14)
English Male Talker			
Voiceless post-alveolar affricate /ʧ/	119 (59)	-233 (113)	51 (67)
Voiced post-alveolar affricate /ʤ/	68 (269)	-271 (300)	-174 (249)
Voiceless alveolar fricative /s/	52 (34)	-82 (38)	285 (107)
Voiced alveolar fricative /z/	82 (49)	-34 (50)	192 (17)
Voiceless post-alveolar fricative /ʃ/	-232 (118)	-102 (50)	173 (59)

3.4 Discussion and Conclusions

The aim of Study 2 was to document the acoustic effects of SMNR on the fricatives and affricates of the English and Mandarin languages. Two SMNR systems were tested in the study and a few crucial steps were taken in order to ensure that any acoustic differences found in the study could be attributed to the SMNR processing tested. First, each hearing aid was verified prior to the recording of speech-plus-noise signals to ensure that both programs in each hearing aid had the same frequency response and input/output functions. The only difference between the two programs was that SMNR was turned on and set to the maximum strength in one of the listening programs whereas SMNR was turned off in the other program. Second, the test stimuli, test procedures (e.g., stimulus presentation and recording), and analysis procedures (e.g., speech-signal retrieval using the inversion technique and parameters in the Praat program) were kept the same for both SMNR-on and SMNR-off conditions. Third, sound field calibration was conducted daily to ensure consistent presentation levels for all stimulus files within an input SNR condition. Fourth, all stimulus files contained a noise-only portion to ensure that the noise reduction system was activated prior to the beginning of speech signals.

The first acoustic measurement, amount of noise reduction, showed that SMNR processing in both hearing aids reduced the magnitude of steady-state pink noise. However, the amount of noise reduction varied between the two hearing aids even though the strength of SMNR in both hearing aids was set to maximum. HA#1 showed level-dependent gain reduction: the amount of noise reduction for HA#1 increased gradually from 1.1 to 4.8 dB when the level of steady-state pink noise was increased from 55 to 75 dBA. For each 5 dB of noise increment, the amount of noise reduction increased by approximately 1.0 dB. The amount of noise reduction for HA#2 increased from 5.0 to 6.0 dB when pink noise was increased from 55 to 65 dBA. However, further increment in the noise level resulted in less noise reduction: the amount of noise reduction for HA#2 decreased to 3.0 dB when pink noise increased to 75 dBA. The difference between the two SMNR tested in the current study can be attributed to manufacturer's specific SMNR algorithms. Previous research has shown that MBNR of different hearing aids will execute different amounts of gain reduction based on the input level, SNR, and type of signal (Bentler & Chiou, 2006; Hoetink et al., 2009).

For example, Brons et al. (2012) examined SMNR in hearing aids by four different manufacturers and found that gain reduction by these SMNR processing ranged from 0 to 12 dB. Bentler & Chiou (2006) showed that that MBNR in two hearing aids had different amounts of gain reduction when steady-state white noise was detected: (i) one reduced gain by about 5 dB and another reduced gain up to 20 dB. Bentler et al. (2008) reported that SMNR examined in their study had higher amount of gain reduction when the input level of white noise increased: approximately 9 and 18 dB of gain reduction for 65 and 81 dB SPL of white noise, respectively. Another study reported that the amount of attenuation was between 8 to 10 dB for a mixture of speech-in-noise signals (steady-state ICRA noise plus single-talker modulated ICRA noise) presented at 0 dB SNR (Alcantara et al., 2003). The amount of gain reduction reported in the current study was within the range reported in the literature. In addition, a variable amount of gain reduction was found for MBNR in different hearing aids, consistent with the findings in previous research. Reduction of gain when noise is detected may improve listening comfort, reduce aversiveness and annoyance of noise; however, too much gain reduction may affect audibility. The exact amount of gain reduction that will result in significant improvement in listening comfort may vary for each individual because of different listening environment. In addition, listening comfort, aversiveness or annoyance of noise are generally assessed through subjective scale ratings or self-report questionnaires (Alcantara et al, 2003; Bentler et al, 2008; Boymans & Dreschler, 2000; Palmer et al., 2006; Walden et al., 2000; Zakis et al., 2009), which may vary at the individual level due to different listening needs and expectations. However, knowing the amount of noise reduction is helpful to a clinician, because the clinician can use this information to decide which parameter of SMNR to choose (e.g., higher strength means more gain reduction) during hearing aid programming and to counsel a hearing aid user regarding the efficacy of the SMNR feature during the hearing-aid fitting session. This information is not readily available to the clinicians.

The second acoustic measurement, effective SNR change, was used to determine if the SMNR tested in this study modified the output SNR of the speech-plus-noise signals. A general observation was that SMNR processing in both hearing aids improved the output SNR of speech-plus-noise signals. However, HA#1 exhibited less SNR change (i.e., less than 2 dB improvement) as compared to HA#2 (i.e., 2 to 5 dB improvement). At high SNR conditions (+10 and +5 dB SNR), HA#1 showed minimal SNR improvement (less than 1 dB) and the amount of SNR improvement increased gradually as the input SNR conditions became poorer (i.e., from the 0 to -10 dB SNR). In contrast, HA#2 showed a maximum amount of SNR improvement (4 to 5 dB improvement) at positive input SNR conditions; further decrease in input SNR resulted in smaller effective SNR improvement (e.g., 2 to 3 dB improvement). To date, a few studies in the literature incorporated the effective SNR measurement and found that the effective SNR improvement ranged from 1.5 to 7.0 dB for input level at 0 to +5 dB SNR (Gustafson et al., 2014; Hagerman & Olofsson, 2004; Pittman, 2011a; 2011b). The results of the current study showed that the effective SNR improvement measured in HA#1 and HA#2 was well within the range of effective SNR improvement reported in the literature. Similar to the Gustafson et al. (2014) study, the current study tested one hearing aid from each of two hearing aid manufacturers: Oticon and Phonak. Although different models of hearing aids were tested as compared to the study, a similar trend of effective SNR change was observed in the current study where HA#1 (Oticon Safari P300) provided less SNR improvement as compared to HA#2 (Phonak Solana M H20). Another observation found in the current study was that the output SNR was lower than the nominal input SNR for both hearing aids. Under the SMNR-off condition, the output SNR for HA#1 was approximately 3 dB poorer than the nominal input SNR (+10, +5, 0, -5, and -10 dB SNR). For HA#2, the output SNR was approximately 6 to 10 dB poorer than the nominal input SNR. This finding was consistent with Gustafson et al. (2014) where they reported that the output SNR of two hearing aids was 3 to 4 dB poorer than the nominal input SNR (i.e., 0 dB SNR) when there was no SMNR processing. SMNR improved the output SNR but not back to the unaided SNR. A few studies suggested that SMNR processing did not have an effect on speech perception when effective SNR improvement was 2.0 dB or less (Marcoux et al., 2006; Pittman, 2011a, 2011b).

The third measurement, the HAQSI nonlinear index and the linear index, was used to quantify the effect of SMNR processing on speech fidelity. For both devices (HA#1 and HA#2), all of the nonlinear index difference scores were below the criterion value (i.e., 0.15). For HA#1, the nonlinear index difference scores between SMNR-on and SMNR-off were essentially zero, except at the poorest input SNR condition (i.e., -10 dB SNR) with a difference value of 0.10. For HA#2, the nonlinear index difference scores between SMNR-on and SMNR-off were approximately 0.10 for most of the test conditions. This finding indicated that SMNR tested in this study did not affect the modulation envelope of speech signals to a significant level based on the criterion level adopted in this study. Arehart et al. (2010) examined the subjective quality ratings by two groups of listeners (hearing impaired and normal hearing) on several different signals. Among those signals, one was speech-in-babble noise at +5 dB SNR and another was speech-in-babble noise at +5 dB SNR with moderate degree of noise reduction processing (e.g., spectral subtraction with maximum attenuation of 12 dB). Subjects were required to rate the quality of those signals on a 5-point scale that ranged from 1 (corresponded to bad quality) to 5 (corresponded to excellent quality). The quality ratings for speech-in-babble with and without noise reduction processing were very similar. The normal-hearing group gave a score of 2.30 to speech-in-babble and a score of 2.34 for speech-in-babble with noise reduction processing; the hearing-impaired group gave a score of 1.74 and 1.76 to speech-in-babble with and without noise reduction processing, respectively. Their findings suggested that noise reduction processing might not have a great impact on the sound quality rating of speech-in-babble signal. The subjective data obtained by Arehart et al. (2010) was then used to validate the HASQI nonlinear index for predicting sound quality (Kates & Arehart, 2010). Based on the findings in Arehart et al. (2010) and the relation of their subjective data to the HASQI nonlinear index, it is likely that SMNR processing examined in the current study would not have substantial effect on subjective quality ratings. However, further research is needed to test this assumption.

The linear index difference score between SMNR-on and SMNR-off was essentially zero for both hearing aids (HA#1 and HA#2) across all talker and SNR conditions. This suggested that SMNR processing did not affect the long-term spectra of speech signals to a significant level based on the criterion level (i.e., 0.15 dB) adopted in this study. Gustafson et al. (2014) also reported that the spectra of speech-plus-noise signals with and without SMNR processing were similar. The findings of the current study were consistent with the Gustafson et al. (2014) study despite a few differences: (i) the coherence measurement rather than HASQI was used to compare the spectrum of an input signal with the output signal of hearing aids and (ii) speech-plus-noise rather than retrieved-speech signals were used in the

Gustafson et al. (2014) study. One advantage of using the HASQI measurement was that additional information regarding the modulation envelope of the speech signal was obtained (i.e., as indicated by the nonlinear index score). In addition, comparing the spectra of retrieved-speech signals instead of speech-plus-noise signals in SMNR-on and SMNR-off conditions provided information regarding whether SMNR processing affected the spectra of speech signals. Nonetheless, it is important to acknowledge that the HASQI measurement was not used to predict subjective quality ratings in this study. Instead, it was used as a measurement tool for objective quantification of the effect of SMNR processing on temporal modulation and long-term spectra of word string. One limitation of the HASQI measurement was that the scores could only suggest whether a test signal (retrieved-speech) was different from a reference signal (input speech) in terms of temporal modulation and long-term spectra. However, it did not provide information such as at which frequency region or at which time-point the two sounds being compared were different. Therefore, three other spectrographic measurements were carried out to provide more detail information on whether SMNR processing affects the acoustic of fricatives and affricates.

The fourth acoustic measurement, frication-noise intensity difference, showed that SMNR processing in HA#1 and HA#2 reduced the intensity of frication noise with the exception that HA#1 showed minimal intensity change in the +10 dB SNR condition. The amount of intensity reduction was also different between the two hearing aids. The results showed that SMNR processing in HA#1 resulted in a consistent pattern of frication-noise intensity reduction; when the input SNR became poorer and noise level became higher, the frication-noise intensity reduction became higher (median values became more negative). For HA#1, the median values for frication-noise intensity difference were between -0.5 to 0.5

dB in the +10 dB SNR condition, suggesting that SMNR processing in HA#1 had little effect on the frication-noise intensity. As the input SNR decreased in 5 dB steps from +10 to -10 dB SNR, the amount of frication-noise intensity reduction became higher. For example, the majority of the median values ranged from 0 dB to -2 dB in the +5 dB SNR condition, -1 to -3 dB at the 0 dB SNR condition, -3 to -5 dB at the -5 dB SNR condition, and -4 to -6 dB at the -10 dB SNR condition. The maximum amount of frication-noise intensity reduction for HA#1 was approximately -6.0 dB in the poorest input SNR condition (-10 dB SNR). The distribution of the median values across all consonants was also consistent within each of the input SNR conditions. For HA#2, the distribution of median values across all consonants in a particular input SNR condition was more varied compared to HA#1. There was also no consistent pattern of distribution: the majority of the fricative and affricate consonants had negative values (intensity reduction) but some had positive values (intensity enhancement). For example, in the +10 dB SNR condition, the Mandarin retroflex fricative /ʂ/ spoken by a male talker had an intensity increment of 6 dB while the Mandarin unaspirated retroflex affricate /tʂ/ spoken by the same talker had an intensity reduction of 6 dB after SMNR processing. Overall the results indicated that the amount of frication-noise intensity difference varied between SMNR in hearing aids tested and SMNR in HA#1 exhibited a more consistent effect across consonants.

The fifth acoustic measurement, release-burst intensity difference, showed that SMNR processing in HA#1 and HA#2 reduced the intensity of the release burst in 12 Mandarin and English affricate consonants tested in this study with the exception that SMNR processing in HA#1 resulted in minimal release-burst intensity difference (i.e., -0.5 to 0.5 dB difference) in the +10 dB SNR condition. The amount of release-burst intensity reduction for

HA#1 became higher (median values became more negative) with decreasing input SNR from +10 to -10 dB SNR in 5 dB steps. The maximum amount of release-burst intensity reduction for HA#1 was approximately -5.0 dB in the poorest input SNR condition (-10 dB SNR). For HA#2, results showed that the median values of release-burst intensity difference varied across consonants under each input SNR condition and the variation became larger with poorer input SNR condition. The maximum amount of release-burst intensity reduction for HA#2 was approximately -8.0 dB in the poorest input SNR condition (0 dB SNR). Overall, the results of the release-burst intensity measurement followed the same pattern found in the frication-noise intensity measurement: the amount of release-burst intensity difference varied between hearing aids tested in the current study and was more consistent across consonants for HA#1.

The sixth acoustic measurement, spectral mean difference, showed that SMNR processing in HA#1 reduced the spectral-mean of frication noise of fricatives and affricates tested in this study. For HA#1, the median values of spectral mean difference ranged between +30 to -500 Hz for the zero and positive input SNR conditions. The spectral mean reduction became larger (between 0 Hz to -600 Hz) with negative SNR input conditions. For HA#2, the spectral mean difference had positive values (spectral mean increment) and negative values (spectral mean reduction). As the input SNR decreased, the number of positive spectral mean difference increased. Overall results of spectral mean measurement indicated that SMNR processing might decrease or increase the spectral mean of the Mandarin and English fricatives and affricates but the effects varied between hearing aids. This finding was in apparent contrast with the results of the HASQI linear index. However, the HASQI linear index examined the overall spectrum of the word strings, which included a mixture of vowels and consonants, whereas the spectral mean measurement examined the spectrum of the frication noise. Thus, any subtle changes in the frication noise might be overlooked in the long-term measurement.

It is difficult to compare the results of the three spectrographic measurements with previous studies of SMNR because acoustic analysis on SMNR-processed speech such as frication-noise intensity change, release-burst intensity change, and spectral mean change is rarely reported. One comparable study was conducted by Parikh and Loizou (2005) where they measured the release-burst frequency difference for stop consonants in quiet and in noise. They defined the release-burst frequency as the most prominent spectral peak in the spectrum of a release-burst. They found that at -5 dB SNR, the release burst frequency was reduced as much as 2500 Hz in the presence of babble noise or speech-shaped noise. However, the identification score for the alveolar stops /t, d/ remained high, suggesting that release-burst frequency might not be relevant for place of articulation perception of stop consonants in the presence of noise. The results of Parikh and Loizou (2005) may not be generalized to the findings of the current study because release burst frequency difference was not included in the current study. However, one might predict that reduction in frication-noise and release-burst amplitude in the presence of noise might reduce audibility of fricatives and affricates. It is acknowledged that there are many other acoustic cues related to perception of fricatives and affricates. However, only the acoustic cues that were most likely to be affected by SMNR such as intensity and spectral mean were examined in this study. Other types of measurement such as formant transition and frication noise duration were not included in the study because these acoustic cues are unlikely to be altered by SMNR processing and Jongman et al. (2000) found that second formant transition and frication noise duration do not reliably distinguish place of articulation for fricative consonants acoustically.

Spectral peak location and spectral moments have been used to distinguish the place of articulation of sibilant fricatives (Forrest et al., 1988; Jongman et al., 2000; Lee et al., 2012). For English, alveolar fricatives /s, z/ had a higher spectral mean than the palato-alveolar fricatives /ʃ, ʒ/; for Mandarin, alveolar fricative /s/ had the highest spectral mean, followed by the alveolo-palatal /ɕ/ and retroflex /ʂ/ fricatives (Li, Edwards, & Beckman, 2007). Hence, one might predict that reduction the spectral mean might affect place of articulation perception for fricatives. One study found that lowering the spectral mean of /s/ by about 1430 Hz affected the production and perception of /s/ (Shiller, Sato, Gracco, & Baum, 2009). In this study, Shiller et al. (2009) conducted speech recognition testing on two groups of female subjects and recorded their production of /sa, si, and su/, before and after they received speech training. During the speech training session, the experimental group listened to stimuli containing word-initial /s/ with lowered spectral mean and another group listened to unaltered stimuli. According to the researchers, this amount of spectral mean change caused the fricative /s/ to have a spectral mean closer to the fricative /ʃ/ but listeners would still categorize these sounds as the fricative /s/. The subjects continued to listen to the training stimuli when their production of /sa, si, su/ was being recorded. Shiller et al. (2009) found that the experimental group produced /s/ with higher spectral mean and the subjects' /s-ʃ/ boundary was shifted toward a lower spectral mean frequency. Based on the findings of Shiller et al. (2009) study, it is likely that if the spectral mean change is 1430 Hz or more, there will be perceptual and speech production consequences. The results of the current study showed that most of the median values of spectral mean change were approximately 600 Hz.

The highest average spectral mean difference was 1139 Hz for fricative /s/ spoken by the Mandarin male talk, recorded from HA#2 at 0 dB SNR condition. This amount of spectral mean change is less than 1430 Hz reported by Shiller et al. (2009). Thus, it is likely that the amount of spectral mean change documented in the current study would not have any perceptual effect. However, it should be noted that all experimental conditions in the Shiller et al. (2009) study were conducted in quiet and it should be cautiously interpreted when comparing to test conditions that include background noise.

The different findings between HA#1 and HA#2 for the three spectrographic measurements could be attributed to the proprietary SMNR algorithm of each hearing aid manufacturer. However, with only two hearing aids tested in the study, the results can only be generalized to the SMNR in these hearing aids. Interaction of SMNR with other advanced processing features was not examined in the current study. Future studies should sample more hearing aids with different types of SMNR and should test interaction of SMNR with other advanced features such as frequency lowering. It should also be noted that the results in the current study were obtained under the worst case scenario where (i) noise reduction was set to the maximum setting (ii) steady-state pink noise was used to activate the SMNR. The setting may not be the usual recommended or default setting in the manufacturer's software and the background noise may not represent the real-world listening situations encountered by each hearing aid users. The use of noise-like speech sounds (fricative and affricates) and poor SNR conditions might have increased the likelihood that acoustical effects of SMNR would be present.

In summary, this study systematically documented the acoustic effects of SMNR on Mandarin and English fricatives and affricates. There are a few findings in the current study.

First, SMNR systems that use signal modulations and/or SNR to detect the presence of noise were effective in reducing steady-state broadband background noise. The amount of gain reduction was level dependent and differed between manufacturers. Second, SMNR processing in both hearing aids improved the output SNR of speech-plus-noise signals, with HA#1 exhibiting less SNR improvement than HA#2. The effective SNR improvement found in this study was consistent with the SNR improvement reported in the literature. Third, the HASQI measurement showed that SMNR processing had minimal effect on the temporal envelope and the long-term spectra of word-strings containing Mandarin and English fricative and affricate consonants. Fourth, spectrographic analyses showed that SMNR tested in this study had acoustic effects on Mandarin and English fricatives and affricates in the amplitude and spectral domain. Nonetheless, it is uncertain how these acoustic changes will affect the perception of fricatives and affricates.

Chapter 4 Effects of Single-Microphone Noise Reduction on Novel Speech Sound Identification in Noise

Single-microphone noise reduction (SMNR) is implemented in hearing aids to suppress background noise. Although clinicians are advised to deactivate SMNR when fitting hearing aids to young children and infants due to the unknown effects of SMNR on this population (American Academy of Audiology, 2013; Bagatto, Scollie, Hyde, & Seewald, 2010; King, 2010), studies support the use of SMNR in adults and school-age children based on positive outcome in listening comfort and lack of negative effect on speech intelligibility (Gustafson et al., 2014; Mueller et al., 2006; Nordrum et al., 2006; Palmer et al., 2006; Stelmachowicz et al., 2010). Nevertheless, most of the SMNR studies tested English speech materials and only the two studies used a cross-language paradigm to examine the effect of SMNR on novel speech sound discrimination (Marcoux et al., 2006; Turgeon et al., 2009). To date, very few studies have examined the effects of noise reduction algorithms on Mandarin (Chen et al., 2015; Liu et al., 2012). The noise reduction algorithms examined in Chen et al. (2015) were used in cochlear implants whereas Liu et al. (2012) examined transient noise reduction algorithm that target transient signals such as door slamming which do not apply to SMNR used in hearing aids. No study has examined the effects of SMNR in hearing aids on the identification of Mandarin speech sound despite Mandarin being the language with the highest number of users in the world. (Lewis et al., 2015).

There are 22 Mandarin consonants: 21 of them can occur at the word-initial position and one of them can only occur at the word-final position (/ŋ/). Among the 21 word-initial consonants, half of them are comprised of fricatives and affricates: five fricatives /f, s, ʂ, ɕ, x/

and six affricates /ts, tsʰ, tʂ, tʂʰ, tɕ, tɕʰ/. Thus, fricatives and affricates play an important role in Mandarin. The remaining half of the Mandarin consonants are six plosives /p, pʰ, t, tʰ, k, kʰ/, two nasals /m, n/ (note that /n/ that can occur at the word-final position as well), one approximant /ɹ/, and one lateral /l/ (Hua & Dodd, 2000; Zhao & Li, 2009). The place of articulation for Mandarin fricatives includes labio-dental /f/, alveolar /s/, retroflex /ʂ/, alveolo-palatal /ɕ/, and velar /x/. These sounds are written as "f", "s", "sh", "x", and "h", respectively in the *Pinyin* system (the Romanization system of Mandarin). The place of articulation for Mandarin affricates includes alveolar /ts, tsʰ/, retroflex /tʂ, tʂʰ/, and alveolo-palatal /tɕ, tɕʰ/ and are written as "z", "c", "zh", and "ch", "j", and "q" respectively in the *Pinyin* system. Notice that the fricatives and affricates at the alveolar, retroflex, and alveolo-palatal place of articulation have a three-way contrast (e.g., /s-ts-tsʰ/): two pairs of fricative-affricate contrasts (e.g., /s-ts/ and /s-tsʰ/) and one pair of aspiration contrast (e.g., /ts-tsʰ/). Note that there is no voicing contrast in Mandarin consonants because all Mandarin consonants are voiceless. In comparison, English has a pair of voiceless fricative-affricate contrast only (i.e., palato-alveolar fricative /ʃ/ and affricate /tʃ/) and there is no three-way contrast between these two sounds. The Mandarin retroflex and alveolo-palatal fricative and affricates are not in the phonemic inventory of the English language. However, the Mandarin alveolo-palatal fricative /ɕ/ and affricates /tɕ, tɕʰ/ are considered to be allophonic with the alveolar or retroflex fricatives and affricates (Munro, 2008).

Studies have shown that acquiring retroflex consonants can be challenging for native Mandarin-speaking children who are developing speech and language. It is even more challenging for children who have hearing loss to acquire retroflex consonants. Hua and Dodd (2000) recorded speech samples from 129 native Mandarin-speaking children with normal hearing (aged 1 year 6 months to 4 years 6 months) and found that these children acquired fricatives after nasals and plosives. Hua and Dodd (2000) also reported that the retroflex sounds /ʂ, tʂ, tʂʰ/ and the alveolar affricates /ts, tsʰ/ were the last consonants to be acquired at the age of four years six months and above. Research also showed that Mandarin-speaking children with hearing loss who used cochlear implants had lower production scores for alveolar and retroflex fricative and affricates (mean score: 28.33% to 60.83 %) as compared to other consonants such as plosives and nasals (mean score: 60.42 % to 97.50%) (Peng, Weiss, Cheung, & Lin, 2004). Liu et al. (2013) examined Mandarin consonant contrast recognition by native Mandarin-speaking children with normal hearing, cochlear implant, or hearing aids. All children were matched in terms of chronological age (3 to 5 years old) and the groups with hearing loss were matched in terms of mean age at of fitting devices (at 24 months of age). The test stimuli were 87 Mandarin consonant contrasts that were grouped into six categories (e.g., aspirated–non aspirated contrasts, same–place different–manner contrasts, and retroflex–non retroflex contrasts). Their study showed that children who used a cochlear implant or hearing aids had significantly lower recognition scores than the normal hearing group, particularly for two categories of speech contrasts: (i) retroflex–non retroflex and (ii) same–place different–manner contrasts. These studies suggested that fricatives and affricates, particularly the retroflex sounds, were difficult to perceive and articulate by native Mandarin children with hearing loss.

In two studies, native English-speaking adults and children were tested with Hindi speech contrasts in a series of discrimination tests and the results showed that SMNR did not affect the speech discrimination ability for novel speech sounds by naïve listeners (Marcoux et al., 2006; Turgeon et al., 2009). In one study, two groups of native English-speaking adults with normal hearing were tested: one group of subjects (n=10) was presented with the SMNR-processed stimuli and the other group of subjects (n=10) was presented with the unaided stimuli (Marcoux et al., 2006). The stimuli used to test the subjects included the voiced and voiceless Hindi dental-retroflex contrasts (i.e., /d̪ - ɖ/ and /t̪ - ʈ/, respectively). All subjects participated in four sessions of discrimination testing within a two-week period. During the discrimination testing, the subjects did not receive any kind of training or feedback in the discrimination task. Using a subset of stimuli (9 out of 90 pairs of stimuli) from the Marcoux et al. (2006) study, Turgeon et al. (2009) tested 19 monolingual English-speaking children, aged 4 to 5 years old. Although Marcoux et al. (2006) and Turgeon et al. (2009) found that SMNR did not affect the discrimination of novel speech sounds (i.e., Hindi dental and retroflex stops) by naïve listeners, their findings may not be generalized to the Mandarin retroflex sounds for two reasons. First, there is no voicing contrast in Mandarin retroflex sounds as opposed to the Hindi dental and retroflex stops. Second, there is a three-way contrast among the Mandarin retroflex sounds (i.e., /ʂ-tʂ-tʂʰ/), while the Hindi dental and retroflex stops do not have this three-way contrast. This three-way contrast involves manner of articulation contrast and aspiration.

In terms of manner of articulation, a plosive sound is characterized by a silent period prior to the presence of a transient burst in a speech waveform. The silent period signifies the obstruction of the oral track and the transient burst signifies the release of pressure built-up behind the point of obstruction during the production of a plosive sound. A fricative sound is characterized by random waveforms with high frequency energy in a spectrogram. The random waveforms signify turbulence created when airstream from the lungs is forced through a narrow constriction during speech production. The concentration of energy is determined by the place of the constriction along the vocal tract. An affricate sound is a sequence of a plosive sound followed by a homorganic fricative sound (Ladefoged & Johnson, 2011). The randomness of the frication noise within a fricative or an affricate consonant mimics the randomness of wideband noise (e.g., white noise or pink noise). For this reason, SMNR processing may affect fricative and affricate sounds when broadband background noise is present.

In summary, Mandarin has the highest number of users around the world but few studies have tested the effect of SMNR in hearing aids on Mandarin speech sounds. Mandarin speech sounds such as fricatives and affricates are important because they comprise half of the Mandarin consonants. Mandarin fricatives and affricates, particularly the retroflex sounds, can be challenging for native Mandarin-speaking children with hearing loss. Given that fricatives and affricates have noise-like features, they are more likely to be affected by SMNR processing when background noise is present as compared to the voicing contrast and the place of articulation contrast of Hindi stop consonants. In order to extend the studies conducted by Marcoux et al. (2006) and Turgeon et al. (2009), Mandarin retroflex fricatives and affricates may provide an interesting test case to evaluate the effect of SMNR processing on the identification of novel speech sounds in noise by naïve listeners, such as the native English talkers because Mandarin retroflex fricatives and affricates are not within the phonemic inventory of English talkers.

4.1 Mandarin Fricatives and Affricates: Novel Speech Sounds for Non-Native Adult Listeners

Mandarin affricates and fricatives are difficult to perceive by non-native listeners, either in quiet or in noise (Lai, 2009; Lee et al., 2012; Tsao et al., 2006). Tsao et al. (2006) examined the discrimination performance of native English-speaking adults on Mandarin alveolo-palatal contrasts and found that native English-speaking adults performed poorer than the Mandarin-speaking adults in a discrimination task. The researchers tested 18 native English-speaking and 18 native Mandarin-speaking young adults. The test stimuli included nine synthesized alveolo-palatal affricate-fricative contrasts (/tɕʰi-ɕi/, /tɕi-ɕi/, and /tɕʰi-tɕi/) with various combinations of frication duration and rise time. The contrastive pairs were presented at 65 dBA through earphones to the subjects in a sound-attenuating booth. The results indicated that the Mandarin talkers had higher discrimination scores than the English talkers. This finding indicated that the native English-speaking adults had more difficulty in discriminating Mandarin alveolo-palatal affricate-fricative contrasts in quiet, as compared to the native Mandarin-speaking adults.

Testing two different populations, Lai (2009) found that non-native talkers had lower discrimination scores than the Mandarin talkers when tested with Mandarin affricates. In the study, Lai (2009) tested two groups of non-native talkers (Malay and Burmese) who were learning Mandarin as a second language. The Malay group had five to six years of experience whereas the Burmese group had only one to two years of experience in learning Mandarin. Each of the non-native groups had 10 subjects and their discrimination performance was compared to a group of 20 native Mandarin talkers from Taiwan. The test stimuli were 18 minimal pairs of disyllable phrases spoken by two native Mandarin-speaking male talkers. Each pair comprised two disyllable phrases that contained an affricate consonant either in the first or second syllable position. The affricate consonants within each pair differed by place of articulation (e.g., retroflex vs. palatal) or manner of articulation (e.g., aspirated vs. unaspirated). Subjects were presented with 72 trials (12 pairs x 2 talkers x 2 trials) and their task was to judge whether the phrases within each pair sounded the same or different. Lai (2009) found that the non-native groups had lower discrimination scores than the native-Mandarin group, particularly for the alveolar-retroflex contrasts (e.g., /ts-tʂ/ and /ts^h-tʂ^h/). The retroflex sounds were heard as the alveolar sounds approximately 70% to 75% of the time for both non-native groups. In addition to the discrimination test, the subjects were asked to rate how similar the six Mandarin affricates were to the speech sounds in their first language. The subjects were also asked to indicate which sounds in their first language were similar to each of the Mandarin affricates. The native-Malay group reported that all six Mandarin affricates were new to them whereas the native-Burmese group reported that four of the six Mandarin affricates were new to them (Lai, 2009). These results indicated that the Mandarin alveolar and retroflex affricates /ts, ts^h, tʂ, tʂ^h/ are difficult to perceive by non-native adult listeners whose native language inventory does not contain these speech sounds.

Unlike the two aforementioned studies that tested non-native listeners in quiet listening condition, Lee et al. (2012) conducted Mandarin fricative identification tests for non-native listeners in a noisy listening condition. Lee et al. (2012) examined the effects of noise and talker variability on Mandarin fricative identification by native-Mandarin and non-native adult listeners. The test stimuli comprised four consonant-vowel syllables (/fa, sa, ɕa, ʂa/) spoken by three adult Mandarin talkers of each gender. The syllables were presented in quiet and in speech-spectrum noise at four signal-to-noise ratios (SNR; -15, -10, -5, and 0 dB). These syllables were presented in blocks according to two formats: (a) talker-specific blocks and (b) mixed-talker blocks. The subjects' task was to identify the syllables they heard by selecting one of the four labels (written in *Pinyin*) displayed on a computer screen. The percentage of correct identification was recorded and arcsine-transformed before statistical analysis was conducted. A significant difference in identification scores between the non-native and native groups was observed at three SNR conditions (i.e., 0, -5, and -10 dB SNR) in the mixed-talker block presentation and at the -5 dB SNR condition in the talker-specific block presentation. The results showed that the fricative-identification performance of non-native listeners was always lower than that of the native listeners. Furthermore, the addition of background noise impeded the identification performance of non-native listeners more than the native-listeners.

To summarize, the results of the above mentioned studies suggested that Mandarin fricatives and affricates were novel and difficult to perceive for non-native adult listeners. The presence of noise degraded the identification performance of non-native listeners even more than the native listeners (Lee et al., 2012). For these reasons, the Mandarin retroflex fricative and affricate consonants were chosen as the test stimuli in the present study.

4.2 Effects of Training on Learning Novel Speech Contrasts

Studies have shown that training with feedback may facilitate adult listeners to learn novel speech contrasts (Flege, 1989; Flege, 1995; Jamieson & Morosan, 1986; Pisoni, Aslin, Perey, & Hennessy, 1982), despite reduced sensitivity to non-native speech contrasts after the first year of life due to linguistic experience (Werker & Tees, 1984). Pisoni et al. (1982) examined the effect of training with immediate feedback on the identification performance of voicing contrasts for stop consonants by native-English speaking adults. A stimulus with -70 ms voice onset time (VOT) is considered to be novel because the VOT does not exist in English. Stimuli with -70, 0, and +70 ms VOT were used in the training sessions. For the test sessions, 15 synthetic stimuli from a full VOT continuum that ranged from -70 ms through 70 ms VOT in 10 ms steps were used. First, Pisoni et al. (1982) tested one group of subjects (i.e., the control group; $n = 15$) who did not receive any training prior to the testing sessions.

The identification test session included two blocks of 150 stimuli from the full VOT continuum (15 stimuli that ranged from -70 to +70 ms VOT in 10 ms steps). A subject's task was to categorize the 15 stimuli into three consonant categories /b, p, p^h/ and no feedback was given during the testing. In a separate experiment, Pisoni et al. (1982) tested another group of subjects (i.e., the experimental group; n = 6) who received two training sessions prior to the identification testing. In the first training session, the experimental group was trained to categorize 240 stimuli (80 repetitions for each stimulus with -70, 0, and +70 ms VOT) into three consonant categories /b, p, p^h/; each trial was followed by immediate feedback indicating the correct response. In the second training session, the experimental group received similar training but with fewer trials (75 stimuli). Then, the subjects underwent the same identification testing as in the control group. The results showed that the experimental group that received identification training had higher identification scores for the stimuli with negative VOT (a novel sound category for the English talkers) as compared to the control group that did not receive any training. This indicated that training with immediate feedback helped naïve listeners in improving their identification performance for non-native speech category.

In another study, Jamieson and Morosan (1986) examined the effect of identification training on the performance of 20 native French-speaking adults on the identification of a voiced and voiceless dental fricatives contrast (/ð-Θ/) that does not exist in French and found that identification training resulted in increased performance in the identification of a non-native speech contrast by native French-speaking adults. In the study, subjects were divided into two groups: one group received identification training (n = 10) and another group did not receive any training (n = 10). In the first session, all subjects underwent identification tests. Twenty four consonant-vowel stimuli (16 natural speech tokens and 8 synthetic tokens) were used in the identification tests. Each stimulus was presented 12 times throughout four blocks of testing (72 trials in each block). After each presentation, the subjects indicated if they heard the voiced or voiceless fricative by pressing either response buttons labeled “the” or “teeth”. In the second through fourth sessions, the experimental group received identification training. All subjects in this group underwent 12 levels of training, starting from the easiest level to the most difficult levels. Within each level, subjects were trained with at least three blocks of 20 trials; subjects moved on to the next level of training when they have completed three consecutive blocks with no more than one error in each block (no more than 5% error in each block). During the training sessions, subjects listened to the synthetic stimuli and indicated if they heard the voiced or voiceless fricative. A light above the response buttons was illuminated when an incorrect answer was chosen. In the fifth session, identification tests were administrated to both groups of subjects. Jamieson and Morosan (1986) found that the post-test identification scores were significantly higher than the pretest identification scores in the experimental group. However, there was no significant difference in the control group.

With the similar aim of examining whether training with feedback will help to improve the identification performance among non-native listeners, Flege (1989) tested native-Chinese subjects on identifying a word-final /t-d/ contrast and found that training with feedback resulted in minor but significant improvement on the performance of Chinese subjects to identify non-native speech contrasts. In the Mandarin language, only the consonant “n” and consonant cluster “ng” are allowed to occur in the word-final position. The word-final /t-d/ contrast are non-native to the Chinese subjects. Flege (1989) hypothesized that the native Mandarin talkers may rely on the acoustic cues (e.g., release burst) that distinguish stops in the word-initial position in the identification task. If these cues were removed, the Chinese subjects may have difficulty identifying this contrast. In the first experiment, Flege (1989) examined the identification performance on English voiceless and voiced dental stop contrasts (/t/ vs. /d/) in word-final position by native-Chinese subjects and native English subjects. Flege (1989) found that the both English and Chinese subjects performed equally well when identifying word-final /t/ and /d/ in natural speech tokens but the Chinese subjects had significantly poorer identification score when the release bursts for the word-final /t/ and /d/ consonants were removed. In the second experiment, two groups of seven subjects (one control and one experimental group) underwent a baseline test, a training session, and a post-training test. Stimuli were consisted of 14 words (seven words with /d/ and another seven words with /t/ in the word-final position). In the baseline and post-training test sessions, 168 trials (14 words X 4 repetitions X 3 blocks) were presented to both groups and feedback was not provided in those sessions. In the training session, 224 trials (14 words X 4 repetitions X 4 blocks) were presented to both group and feedback was provided to the experimental group only. In the post-test session, the experimental group had an increment of 9% in the identification performance whereas the control group had little improvement (1% only). However, the amount of improvement was not significantly different between the two groups. In the third experiment, Flege (1989) decided to include more trials with fewer words in the training session to examine if training could improve the identification score. Subjects (n = 16) were trained with four words (240 trials; 4 words X 4 repetitions X 15 blocks) but tested with eight words (96 trials). Four untrained words with a /t/ or /d/ in the word-final position were also included as test stimuli to examine whether training could be generalized to other untrained words. Flege (1989) found that (i) the identification score was significantly higher in the post-test session than the baseline session, (ii) the post-test identification score (75% correct) was slightly higher than the post-test scores in the second experiment (65% correct), and (iii) the training effect was generalized to the untrained words.

In another study, Flege (1995) examined the effects of two types of training (i.e., identification training vs. discrimination training) on native Mandarin talkers' ability to identify word-final /t-d/ contrasts in English and found that the identification performance was not significantly different between the group that received identification training and the group that received the same-different discrimination training. In the study, all subjects underwent one baseline test, seven training sessions, and seven post-tests. Retention tests were conducted at two weeks and also at two months after the last post-test. Twenty Mandarin talkers were randomly assigned into two groups: one group (n = 10) received the identification training and the other (n = 10) received the same-different discrimination training. Stimuli were 96 naturally spoken English words (8 words X 6 talkers X 2 conditions) in a consonant-vowel-consonant format. Each word started with the consonant /b/ and contained a word-final /t/ or /d/. During each identification training session, 240 trials (24 tokens X 10 repetitions) were presented to the subject. A classification procedure with two alternatives was used where the subjects had to indicate whether they heard /t/ or /d/. For the same-different discrimination training, pairs of words were presented to the subjects and they are required to indicate if the words within a pair sounded the "same" or "different." A total of 120 stimuli pairs (60 same pairs and 60 different pairs) were presented to the subjects. Feedback was given to both group of subjects during the training sessions where the button for the correct response lit up after the subjects have responded after each trial. For the identification tests (baseline and post-training sessions), each subject's percent correct scores were recorded and transformed into Rationalized Arcsine Unit (RAU) prior to statistical analysis (Studebaker, 1985). Flege (1995) found that the RAU scores for the baseline test, post-training tests and the retention tests were not significantly different between the group that received identification training and the group that received the same-different discrimination training. However, generally the group that received identification-training had higher RAU scores in all post-training tests than the other group. Self-reports from this group also revealed that they enjoyed the training process more and were more willing to receive further training than the subjects from the same-different discrimination training group. In addition, Flege (1995) also found that the first post-training identification score (averaged across groups) was not significantly higher than the baseline identification score. The second to seventh post-test identification scores were significantly higher than the baseline identification score but there was no significant difference among the second to the sixth post-test identification scores. These results indicated for the training to be effective, at least two sessions of training are required. Flege (1995) also found that the identification scores obtained during the retention tests (i.e., two weeks and then two months after the last training session) showed that training effect could last up to two months post training.

The findings from the above mentioned studies had a few implications that were relevant to this current study. First, identification training with immediate feedback could improve the performance of adult listeners in identifying novel speech categories (Jamieson & Morosan, 1986; Flege, 1989; Pisoni et al., 1982). Second, Flege (1995) showed that (i) identification training had the same effect as the same-different discrimination training, (ii) at least two training sessions were required to measure significant improvement from the baseline identification scores, (iii) increasing training sessions from two to six training sessions did not significantly improve the identification scores, and (iv) the training effect could be retained for up to two months post-training suggesting that there was a carryover effect. Therefore, some methodological considerations in the current study included (i) identification training was used to train adult listeners to learn novel speech contrasts, (ii) immediate feedback was provided to the subjects during the training sessions to facilitate novel speech contrast learning, (iii) at least two training sessions were included in order to measure significant improvement from the baseline identification scores, and (iv) a between-group design was implemented due to crossover effects of training.

In summary, the purpose of this study was to examine the effect of SMNR on the identification of novel speech sounds in background noise by naive listeners. A cross-language paradigm was chosen for several reasons. First, the identification task can be more challenging for English talkers when Mandarin retroflex fricative and affricates are used as the test stimuli as opposed to the English post-alveolar fricative and affricate contrast, because the Mandarin retroflex sounds are not within the phonemic inventory of English. In English, there is a one-way contrast for the voiceless post-alveolar fricative and affricate (i.e., /ʃ – tʃ/) but in Mandarin, there is a three-way contrast for the retroflex fricative and affricates (i.e., /tʂ –tʂʰ/, /tʂ – ʂ/, and /tʂʰ – ʂ/). Naïve listeners whose main language is English may assimilate the Mandarin affricates /tʂ, tʂʰ/ with the English affricate /tʃ/ and the Mandarin fricative /ʂ/ with the English fricative /ʃ/. According to Best and Tyler (2007), when two non-native phonemes are perceived as a native phoneme by a naïve listener, this may be regarded as the single-category assimilation or the category-goodness assimilation. These two types of assimilation may yield a moderate or poor performance in a discrimination or an identification task that involves non-native speech stimuli. Therefore, it is assumed that the identification task will be more challenging for English talkers when Mandarin retroflex sounds are used as the test stimuli. Most of the SMNR studies that have tested English listeners with English materials did not find any benefit of SMNR on speech intelligibility. Perhaps by using a more challenging speech stimuli to test the English talkers, the identification test can be more sensitive in detecting the benefit of SMNR. Second, by using the non-native speech contrasts, the linguistic cues can be minimized and the naïve listeners may be forced to rely on bottom up processing (where the acoustical properties of a stimulus plays a role) in the identification task. Third, previous studies have shown that listeners are disrupted more by background noise in recognition of a non-native language (Lecumberri et al., 2010). Consistent with the previous studies, Lee et al. (2012) also showed that adult English talkers had lower scores than the adult Mandarin talkers in the identification of four Mandarin fricatives in background noise. Lee et al. (2012) concluded that the addition of background noise impeded the identification performance of non-native listeners more than the native-listeners in the identification of Mandarin fricatives. Therefore, by testing English talkers with Mandarin fricatives and affricates in background noise, the question of whether SMNR processing will have any effect on novel speech sound identification by naïve listeners in background noise can be examined.

There were at least three possible outcomes. First, the identification performance could have been significantly higher in the SMNR-on condition than the SMNR-off condition. This result would imply that SMNR is beneficial in the identification of Mandarin retroflex sounds in noise by naïve listeners. Second, the identification performance could have been significantly lower in the SMNR-on condition than the SMNR-off condition. This result would imply that SMNR has a detrimental effect on the identification of Mandarin retroflex sounds in noise. Third, there could be no difference in the identification performance between the SMNR-on and SMNR-off conditions. This result would imply that SMNR has no detrimental effect on the identification of Mandarin speech sounds in noise by naïve listeners. Identification of these speech contrasts is important because under the same vowel and tonal contexts, different Mandarin retroflex consonants will carry different meanings in a consonant-vowel syllable. For example the syllable /tʂʰɻ/ ("*chi*" in *PinYin* with tone 1) means "eat" (吃), the syllable /ʂɻ/ ("*shi*" in *PinYin* with tone 1) means "poem" (诗), and the syllable /tʂɻ/ ("*zhi*" in *PinYin* with tone 1) means juice ("汁"). Furthermore, it is assumed that if the naïve listeners can identify the Mandarin retroflex fricative and affricates under the SMNR-on condition as accurate or better than in the SMNR-off condition, it is unlikely for the native Mandarin listeners to have difficulties in identifying the retroflex fricative and affricates under the SMNR-on condition. This is because it is well documented in the literature that Mandarin listeners have better discrimination and identification performance for Mandarin speech sounds as compared to the English listeners (Lee et al., 2012; Tsao et al., 2006). Thus, the results can provide indirect evidence to support the use of SMNR in Mandarin listeners.

In spite of the reasons explained above, there are a few limitations of using this cross-language paradigm in this study. First, the perception of non-native speech contrasts by infants and adults are different where language experience plays a role (Kuhl et al., 2006; Tsao et al., 2006; Werker & Tees, 1984). The general findings in the literature suggest that infants at the age of 6-8 months have the ability to discriminate native and non-native speech contrasts but the ability to discriminate the non-native speech contrasts declines while the ability to discriminate native speech contrasts improves at around 10 to 12 months old (Kuhl et al., 2006; Werker & Tees, 1984). Therefore, the findings of the current study may not directly apply to the pediatric population because this population was not tested in the current study. Second, studies on adults have shown that adult listeners can learn non-native speech sounds with training (Flege, 1989; Flege, 1995; Jamieson & Morosan, 1986; Pisoni, Aslin, Perey, & Hennessy, 1982). However, a training paradigm that requires multiple sessions may induce fatigue and increase the likelihood of subject drop out. Precaution steps such as providing break sessions and motivation to the listeners can be implemented to minimize fatigue and subject drop out. Flege (1995) showed that (i) at least two training sessions were required to see a significant improvement in the non-native identification performance and (ii) the non-native identification scores of the second and sixth post-training sessions did not have a significant difference. In order to balance the risk of subject drop-out and the chance of obtaining a significant improvement in the identification score, four sessions of training were implemented in the current study instead of more extensive training.

In the current study, two groups of self-reported native English talkers with normal hearing were tested. These subjects had no previous knowledge or experience with Mandarin speech sounds and they were not learning Mandarin as a second language. Therefore, they are considered as naïve listeners (Best & Tyler, 2007). One group listened to stimuli processed with SMNR turned on and the other group listened to stimuli processed with SMNR turned off. The identification scores and reaction times were documented. I hypothesized that the identification performance would be different for SMNR-on and SMNR-off groups and the direction of the difference was uncertain. I predicted that the identification score would be lower in the baseline session than the post-test sessions. I hypothesized that reaction time would be (i) shorter with increasing number of training sessions and (ii) shorter for the SMNR-on group than the SMNR-off group. The underlying assumptions were (i) less processing time would be required when the sound categories were learned and (ii) a reduction in reaction time would reflect reduced listening effort (Sarampalis et al., 2009) and therefore subjects in the SMNR-on group were expected to have a shorter reaction time.

4.3 Method

4.3.1 Subjects

A total of 32 subjects (23 females and 9 males) participated in this study. The inclusion criteria were (i) self-reported English-speaking adults, (ii) passed hearing screening at 20 dB HL from 250 to 8000 Hz bilaterally, (iii) no reported history of speech and language disorder, (iv) no reported cognitive impairment, and (v) no phonetic knowledge or training with Mandarin. Subjects were recruited through posters displayed in public areas such as universities, public transit stations, and coffee shops.

All subjects reported that English was their dominant language. Some of the subjects had studied other languages at school; however, they were not fluent nor using these languages regularly on a daily basis. Approximately 62.50% (n = 20) of the subjects reported that they learned French, 28.12% (n = 9) learned Spanish, and 21.88% (n = 7) learned German. In addition, some subjects reported that they were exposed to other languages when they were very young but were not fluent nor using these languages on daily basis since school age; these languages included Korean (n = 4), Japanese (n = 3), Russian (n = 3), Cantonese (n = 2), Swedish (n = 1), Vietnamese (n = 1), Portuguese (n = 1), and Hindi (n = 1).

All subjects were post-secondary school students, with mean age of 22 years old and 1 month (standard deviation = 1 year 8 months; maximum age = 25 years 7 months; minimum age = 19 years 2 months). All subjects passed hearing screening at 20 dB HL across octave frequencies. The use of normal-hearing subjects allowed the effect of SMNR to be tested without confounding factors of audibility or acclimatization to personal hearing aids. Subjects were randomly assigned into one of the two groups: one control group tested with stimuli processed without SMNR (n = 16) and one experimental group tested with stimuli processed with SMNR (n = 16). Each subject received $40 at the end of experiment as a token of appreciation.

4.3.2 Equipment

The testing equipment included a PC with two monitors (one with touch screen function); the SykofizX software, PA5 attenuator, and HB7 headphone driver of the Tucker Davis Technology system (Tucker-Davis Technologies Inc., USA); and a pair of Sennheiser HD 265 linear headphones (Sennheiser Canada Inc.). In all instances, stimuli were routed from the PC soundcard to the Tucker Davis Technology system using the SykofizX software and presented to each subject via the Sennheiser headphones. One PC monitor was set outside the test booth for the tester to monitor the subject's progress. Another touch screen monitor was located inside the sound booth. The response choices were displayed on the touch screen monitor and the SykofizX software (Tucker-Davis Technologies Inc., USA) was used to record the subject's response and reaction time.

4.3.3 Stimuli

Vowel-consonant-vowel (VCV) nonsense words containing one of the three Mandarin retroflex consonants (retroflex fricative /ʂ/, aspirated retroflex affricate /tʂʰ/, and unaspirated retroflex affricate /tʂ/) in three vowel contexts /a, I, u/ were used as the test stimuli. Each

VCV token had the same initial and final vowel. The Mandarin retroflex consonants are not in the English phonemic inventory, therefore, these consonants were considered novel to the participants (Munro, 2008). The VCV tokens were spoken by a native Mandarin female and a male talker. These tokens were concatenated into word strings and were presented to and recorded from a commercial behind-the-ear hearing aid (Oticon Safari P300; labeled as HA#1 in Chapter 2 and Chapter 3) in the presence of pink noise at -5 dB SNR. According to the manufacturer's specifications, this hearing aid has a modulation-based noise reduction (MBNR) and a six-channel processing. This hearing aid was selected for two reasons. First, the extracted speech from the recordings of this hearing aid had better speech fidelity (as indicated by higher amount of attenuation) as compared to speech signals processed with another hearing aid (see Figure 6; Chapter 2). Second, acoustic measurements (see Chapter 3) showed that the acoustic effects from this hearing aid were more consistent across consonants, vowel contexts, and between the two languages examined in the second study. Thus the results obtained with this study would be easier to generalize to other consonants and languages tested in the second study.

The recordings were collected in a double-walled sound-treated booth. Each of the speech-plus-noise files was presented to the hearing aid mounted on the left ear of Knowles Electronic Manikin for Acoustic Research (G.R.A.S. Sound and Vibration, Denmark). The stimuli were presented via a Behringer Truth B2303A loudspeaker (MUSIC group IP Ltd., Philippine) at 70 dBA. Sound field calibration was conducted for each speech-plus-noise file to ensure consistent presentation level to the microphone of the hearing aid. The recordings were obtained with the hearing aid under two conditions: (i) SMNR-on and (ii) SMNR-off. The recordings were digitized at 44.1 kHz and 32-bit resolution and stored in the computer in wav file format. The VCV tokens embedded in pink noise were then excised from the speech-plus-noise recordings and used as the test stimuli. All test stimuli were down sampled to 25000 Hz due to the requirement of the SykofizX software.

4.3.4 Procedure

All subjects participated in five sessions with at least 24 hour separation between sessions (Jamieson & Morosan, 1986). On Day 1, written consent was obtained from each subject. Subsequently, all testing was conducted in a double-walled sound-treated booth. Each subject's hearing was screened at 20 dB HL from 250 Hz to 8000 Hz in octave and inter-octave frequencies using the Grason Stadler GSI 61 clinical audiometer with an ER3A insert earphone transducer. Next, each subject underwent the identification baseline test. On Day 2 through Day 5, subjects underwent a training session followed by a post-test each day. Table 29 shows the summary of procedure for all five days. Throughout all sessions, subjects were given a break any time they needed to avoid fatigue. They were informed about their daily identification score at the end of each session in order to maintain the participants' interest and motivation for subsequent sessions. In addition, participants also received stickers when their daily performance improved. Stimuli were presented to each subject monaurally via the headphones at a comfortable listening level (approximately 64-65 dBA measured with fast sound level meter setting).

Table 29. Activities conducted on each day of testing.

Day		Activities
Day 1	i.	Obtain written consent
	ii.	Hearing screening
	iii.	Baseline identification test
Day 2	i.	Training session #1
	ii.	Identification post-test #1
Day 3	i.	Training session #2
	ii.	Identification post-test #2
Day 4	i.	Training session #3
	ii.	Identification post-training test #3
Day 5	i.	Training session #4
	ii.	Identification post-test #4

4.3.4.1 Identification Tests: Baseline and Post-training Tests

The control group was tested with the SMNR-off stimuli and the experimental group was tested with the SMNR-on stimuli. All participants were tested using VCV tokens spoken by both talkers (a male and a female). A classification procedure with three alternatives was used to test the subjects. In this paradigm, a subject's task was to indicate what they heard by selecting one of the three response buttons on a touch screen monitor after each VCV token was presented. The response buttons were labeled with three pictures as shown in Figure 19. The picture of a pig was used to represent the unaspirated retroflex affricate /tʂ/, the picture of a tea cup was used to represent the aspirated retroflex affricate /tʂʰ/, and the picture of a book was used to represent the retroflex fricative /ʂ/. These pictures were used because the first consonant of the objects represented in each picture corresponds to the retroflex sounds being tested in the present study. For example, the Mandarin word for "pig" is written as "Zhū", "tea cup" is written as "Chá Bēi", and "book" is written as "Shū" in *Pinyin*. Before the baseline identification test began, the subjects were familiarized with the task; they were informed that three pictures would be displayed on a touch screen and their task was to select the picture corresponding to the consonant. Participants were also told to concentrate on the consonants and ignore the vowels.

Figure 19. Participant screen for the baseline identification test and the post-training identification tests.

Subjects were presented with two blocks of 45 trials (3 consonants X 3 vowel contexts X 5 repetitions). Each block represented a talker gender (either female or male). The sequence of consonants and vowel contexts was randomized in each block. The sequence of talker-gender block was counterbalanced on each day of testing. No feedback was provided during the baseline identification test and post-tests. The stimulus interval was controlled by each participant according to their speed of responding. The time to complete the identification test was approximately 10 to 15 minutes for each subject. Identification scores were documented for the baseline (Day 1) and the post-test sessions (Day 2, Day 3, Day 4, and Day5). The identification scores were transformed into RAU in order to meet assumptions for statistical analysis (Studebaker, 1985). Reaction time was also recorded.

4.3.4.2 Identification Training

The control group was trained with the SMNR-off stimuli and the experimental group was trained with the SMNR-on stimuli. Half of the participants from each group were trained with VCV tokens spoken by the female talker, the other half were trained with VCV tokens spoken by the male talker. The purpose for training the participants with one talker gender but testing them with both talker genders was to determine whether training with one would generalize to the other talker. The classification procedure with three alternatives, similar to the one in the identification testing but with the addition of visual feedback, was used in each training session. During each training session, subjects were presented with two blocks of 90 trials (3 Mandarin consonants X 3 vowel contexts X 10 repetitions). In total, subjects were presented with 180 trials in each training session.

Before the training began, participants were informed that they would see three pictures on a touch screen and these pictures were the same as the ones they saw during the baseline test. Their task was to select the picture corresponding to the consonant. Participants were informed that when their answer was correct, the response button would flash once; if their answer was incorrect, the correct response button would flash three to four times. The typical training time was approximately 20 to 30 minutes for each subject.

4.4 Results

4.4.1 Identification Performance

A mixed-model analysis of variance (ANOVA) was used to analyze the RAU score as the dependent variable. The between group variables included SMNR (2 levels: SMNR-on vs. SMNR-off) and training voice (2 levels: female vs. male). The within group variables included test voice (2 levels: female vs. male) and test session (5 levels: baseline, post-test 1, post-test 2, post-test 3, and post-test 4). For the within group variables, Mauchly's test of sphericity revealed that the test session main factor did not meet the sphericity assumption and the Greenhouse-Geisser epsilon correction was used to adjust the degrees of freedom. The Test Voice x Test Session interaction factor met the sphericity assumption. The Levene's test revealed that all but one condition met the assumption of equality of error variance (i.e., female-voice post-test 1 and female-voice post-test 2). Therefore, a more stringent alpha level was used as the significance criterion (i.e., .025 instead of .05) for all main effects and interaction effects (Gamst, Meyers, & Guarino, 2008).

Table 30 shows the main effects and interaction effects with their significance level and effect size. The mixed-model ANOVA revealed that the three-way interaction effect of Training Voice x Test Voice x Test Session was significant. The two-way interaction effects of Training Voice x Test Voice, Training Voice x Test Session, and Test Voice x Test Session were significant. The training voice, test voice, and test session main effects were significant. The SMNR main effect was not significant. The four-way interaction, other three-way interactions, and two-way interaction effects were not significant. Because the Test Voice x Training Voice x Test Session three-way interaction effect was significant, simple main effect testing was conducted and other lower level of interaction factors and main effects need not be interpreted.

Table 30. Summary of the mixed-model ANOVA results for identification performance.

Effects	df; error	F value	*p* value	η_p^2
Main				
SMNR	1; 28	1.05	.32	.04
Training voice	1; 28	9.21	.01*	.25
Test voice	1; 28	65.80	.00**	.70
Test session	2.93; 82	57.86	.00**	.67
Two-way interaction				
SMNR x Training voice	1; 28	0.03	.86	.00
SMNR x Test Voice	1; 28	1.13	.30	.04
SMNR x Test Session	2.93; 82	0.57	.63	.02
Training Voice x Test Voice		72.96	.00**	.72
Training Voice x Test Session	2.93; 82	4.11	.01*	.13
Test Voice x Test Session	4; 112	11.83	.00**	.30
Three-way interaction				
SMNR x Training Voice x Test Voice	1; 28	3.01	.09	.10
SMNR x Training Voice x Test Session	2.93; 82	0.82	.48	.03
SMNR x Test Voice x Test Session	4; 112	0.41	.80	.02
Training Voice x Test Voice x Test Session	4; 112	9.33	.00**	.25
Four-way interaction				
SMNR x Training Voice x Test Voice x Test Session	4; 112	1.33	0.26	.05

Note. * $p < .025$; ** $p < .001$

Figure 20 illustrates the mean RAU for the female voice test (top panel) and male voice test (bottom panel) plotted as a function of test session and training voice. The solid line represents the group of subjects who received training in female voice and the dotted line represents the group of subjects who received training in male voice.

For the female-voice test, simple main effects testing showed that there was no significant difference in RAU between the group trained in female and male voice across all test sessions. For the male-voice test, simple main effect testing showed that there was no significant difference in RAU between the groups trained with female and male voice in the baseline test. However, the subjects trained in male voice had significantly higher mean

RAU than the subjects trained in female voice across all post-tests. Table 31 shows the results of pair-wise comparisons (Bonferroni adjustment) with significant difference.

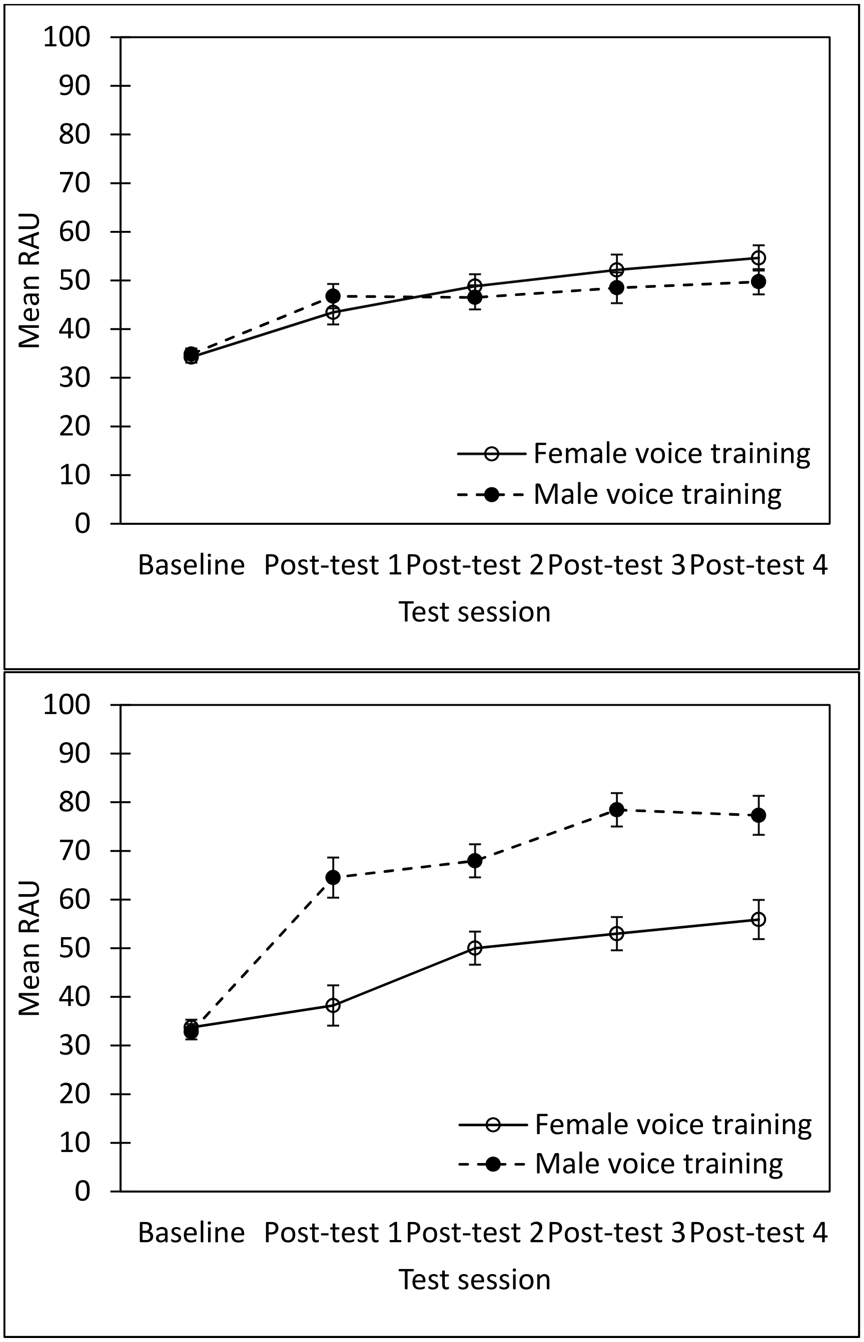


Figure 20. Mean RAU for female-voice test (top panel) and male-voice test (bottom panel) as a function of test sessions and training voice. Error bars represent standard errors of the mean.

Table 31. Pairwise comparison between groups for each test session under the male-voice test condition.

Test session	Mean RAU (SE) Female-voice training	Male-voice training	Mean difference (SE), *p* value
Baseline	33.74 (1.57)	32.84 (1.57)	0.90 (2.22), $p = .69$
Post-test 1*	38.23 (4.14)	64.51 (4.14)	26.28 (5.85), $p < .001$
Post-test 2*	50.01 (3.40)	67.97 (3.40)	17.97 (4.81), $p = .001$
Post-test 3*	52.98 (3.42)	78.45 (3.42)	25.47 (4.84), $p < .001$
Post-test 4*	55.89 (4.02)	77.32 (4.02)	21.43 (5.69), $p = .001$

Note. SE = standard error; * significant

Under each test voice condition, simple main effects testing also showed that post-test 3 and post-test 4 were significantly higher than the baseline RAU scores for all groups. Table 32 shows the results of pair-wise comparisons with significant difference.

Table 32. Pairwise comparison between test sessions for each group of subjects under the female-voice and male-voice test conditions.

Test Voice	Baseline vs. Post-test 3			Baseline vs. Post-test 4		
	Mean difference	SE	*p* value	Mean difference	SE	*p* value
Female						
Group trained with female voice	17.9	3.5	$p < .001$	20.4	2.9	$p < .001$
Group trained with male voice	13.6	3.5	$p = .005$	14.9	2.9	$p < .001$
Male						
Group trained with female voice	19.2	4.2	$p = .001$	22.2	4.5	$p < .001$
Group trained with male voice	45.6	4.2	$p = .001$	44.5	4.5	$p < .001$

Note: SE = standard error.

Figure 21 shows the mean RAU plotted as a function of test session with separate lines representing the SMNR x Test Voice conditions. The solid lines represent female voice test and the dotted line represents male voice test. The filled symbols and the open symbols represent the SMNR-on condition and SMNR-off conditions, respectively.

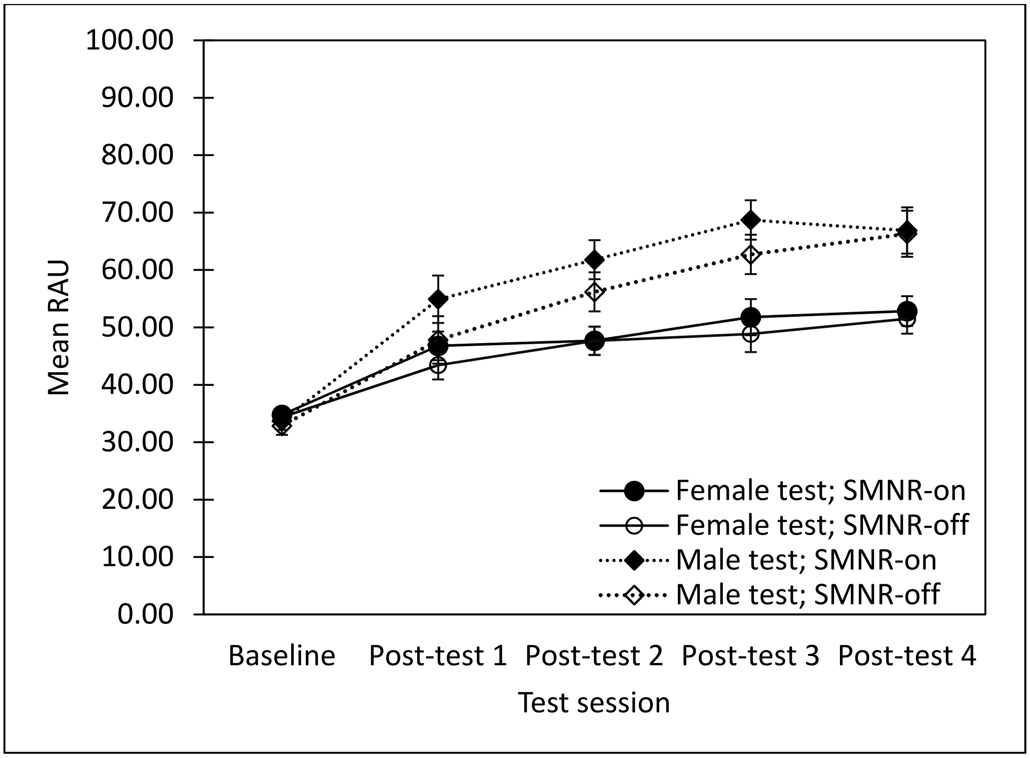


Figure 21. Mean RAU as a function of test sessions and SMNR conditions. The dotted lines represent mean RAU for male voice tests. The solid lines represent the mean RAU for the female voice tests. The filled markers represent the SMNR-on condition and the open markers represent the SMNR-off condition. Error bars represent standard errors of the mean.

4.4.2 Reaction Time in the Identification Tests

The mean reaction time for the correct responses of each test session was calculated. Reaction times exceeding two standard deviations above or below the mean for each individual were excluded (Lee et al., 2012). The reaction times were log transformed in order to meet normality assumption for conducting a parametric test. A mixed-model ANOVA was used to analyze the log-transformed reaction time as the dependent variable. The between group variables included SMNR (2 levels: SMNR-on vs. SMNR-off) and training voice (2 levels: female vs. male). The within group variables included test voice (2 levels: female vs. male) and test session (4 levels: post-test 1, post-test 2, post-test 3, and post-test 4). Mauchly's test of sphericity indicated that the sphericity assumption for the test session main factor was violated. Therefore, the Greenhouse-Geisser epsilon correction was used to adjust the degrees of freedom. Levene's test of homogeneity of variance revealed that most test conditions met the homogeneity of variance assumption except for the female-talker post-test 3 condition. Therefore, a more stringent alpha level (i.e., .025 instead of .05) was used as the significance criterion (Gamst, Meyers, & Guarino, 2008).

Table 33 shows the main and interaction effects with their significance levels and effect sizes. The Test Voice x Training Voice two-way interaction effect was significant. The test voice and test session main factors were significant. The SMNR and training voice main factors were not significant. All other two-way, three-way, and four-way interaction effects were not significant. Because the Training Voice x Test Voice two-way interaction effect was significant, simple effects testing was conducted (with Bonferroni adjustment to keep Type I error rate at 2.5%).

Table 34 shows that the mean reaction time (log transformed) for the group trained in male voice had significantly shorter reaction time when tested in male voice than when tested in female voice. The group trained in female voice had no difference when tested in female or male voice. These findings are reflected in Figure 22.

Table 33. Summary of the mixed-model ANOVA results for reaction time in the identification tests.

Effects	F value	p value	η_p^2
Main			
SMNR	0.30	.59	.01
Training Voice	1.62	.21	.06
Test Voice	33.75	.00*	.55
Test Session	13.31	.00*	.32
Two-way interaction			
SMNR x Training Voice	1.11	.30	.04
SMNR x Test Voice	0.08	.78	.00
SMNR x Test Session	1.64	.21	.06
Training Voice x Test Voice	12.13	.00*	.30
Training Voice x Test Session	0.13	.87	.01
Test Voice x Test Session	0.70	.55	.02
Three-way interaction			
SMNR x Training Voice x Test Voice	0.78	.39	.03
SMNR x Training Voice x Test session	0.62	.53	.02
SMNR x Test Voice x Test Session	0.72	.55	.03
Training Voice x Test Voice x Test Session	0.90	.45	.03
Four-way interaction			
SMNR x Training Voice x Test Voice x Test Session	0.25	.86	.01

Note. * $p < .001$

Table 34. Pairwise comparison of the log-transformed reaction time for different groups under the female-voice and male-voice test conditions.

	Log-transformed reaction time Mean (SE)		
Group	FV	MV	Mean difference (SE), *p* value
Female-voice training	2.99 (.04)	2.95 (.05)	0.04 (.02), $p = .111$
Male-voice training*	2.98 (.04)	2.82 (.05)	0.16 (.02), $p < .001$

Note: FV = female-voice test; MV = male-voice test; SE = standard errors of mean; * significant

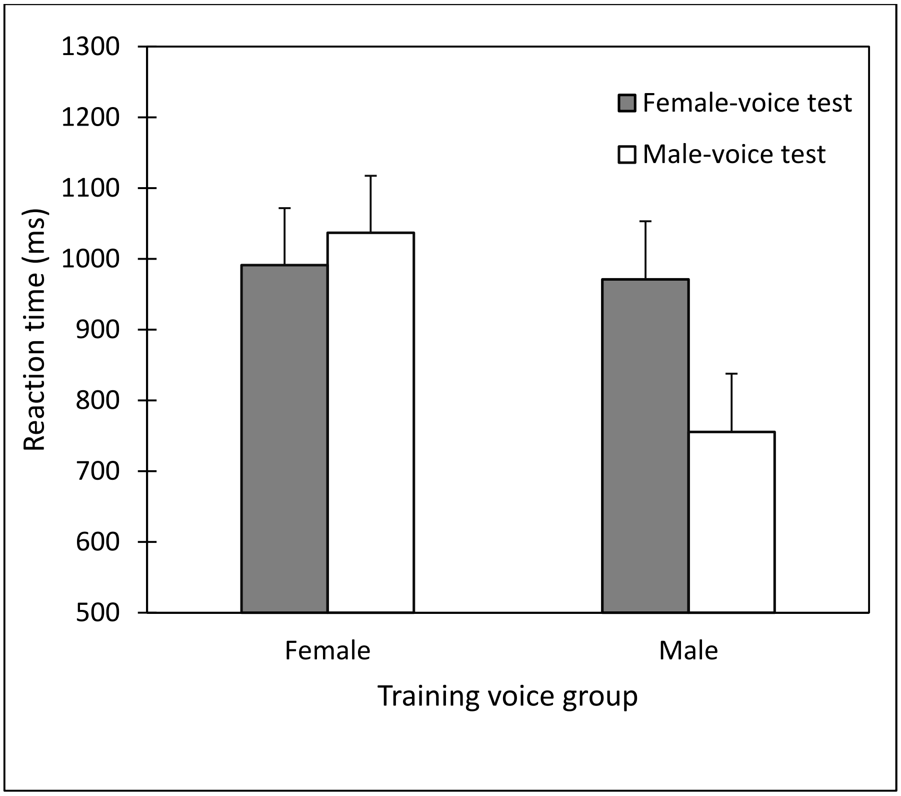


Figure 22. Mean reaction time (non-transformed) as a function of training voices. The grey bars represent the mean reaction time in the female-voice test condition. The white bars represent mean reaction time in the male-voice test condition. Error bars represent standard errors of the mean.

Post-hoc testing (Bonferroni) was conducted for the test session main effect. These findings are shown in Figure 23 where the non-transformed mean reaction time was plotted as a function of test session. Table 35 shows that the mean reaction time in post-test 1 and post-test 2 was significantly higher than post-test 3 and post-test 4.The mean reaction time in post-test 3 and post-test 4 was not significantly different from each other.

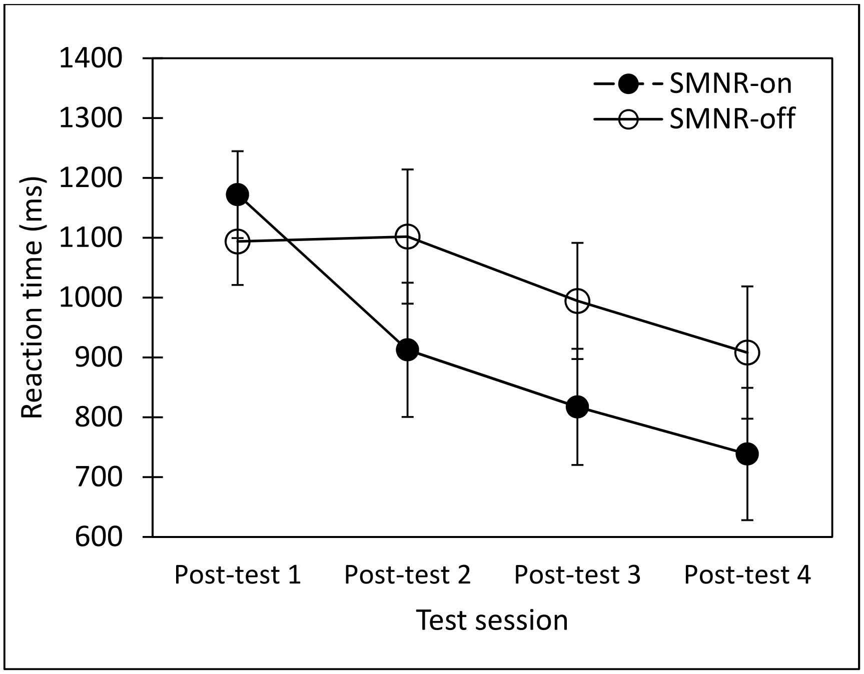


Figure 23. The interaction effect of Test Session X SMNR. Although the interaction effect was not significant, this figure is shown because the main objective of the current study was to examine the SMNR effect. The y-axis in each panel represents the non-transformed mean reaction time. The error bars denote standard error of the mean.

Table 35. Pairwise comparison of the log-transformed reaction time for different test sessions.

Test session	**Mean difference (SE), *p* value**
Post-test 1 compared with	
post-test 2	0.05 (.02), *p* = .231
post-test 3	0.12 (.03), *p* = .007*
post-test 4	0.14 (.03), *p* = .001*
Post-test 2 compared with	
post-test 3	0.06 (.02), *p* = .004*
post-test 4	0.09 (.02), *p* = .001*
Post-test 3 compared with	
post-test 4	0.03 (.02), *p* = .100

Note: SE = standard errors of mean; * significant

4.5 Discussion and Conclusions

4.5.1 Identification Performance

This purpose of the present study was to examine the effect of SMNR on novel speech sound identification in background noise by naïve listeners. The results of the identification tests revealed that the SMNR main effect and other two-way, three-way, and four-way interaction effects that involved SMNR were not significant. These findings suggested that SMNR had no significant effect on the learning of novel speech sound identification by naïve listeners. This finding did not support the hypothesis. Figure 21 showed that there was no significant difference in RAU scores between the SMNR-on and SMNR-off groups in the baseline identification test where all subjects performed at chance level (e.g., RAU score of 34.56). This suggested that all subjects had the same level of performance prior to any intervention. Results also showed that there was no significant difference in the mean RAU score between the SMNR-on and SMNR-off groups in all post-tests: the mean RAU between SMNR-on and SMNR-off groups were similar in the female-voice test condition. Although the SMNR-on group obtained slightly higher scores than the SMNR-off group in the male-voice test condition (e.g., post-test 1 through post-test 3), the difference in RAU was not statistically significant (see Figure 21) and the effect size of SMNR was small. These results indicated that the activation of SMNR did not improve or degrade the listeners' performance in identifying novel speech sounds in background noise. Despite a few minor methodological distinctions between this study and the Marcoux et al (2006) study, the results of this study were consistent with Marcoux et al. (2006) and provided further evidence to support that SMNR does not improve or degrade identification of novel speech sounds by naïve listeners. The results were also consistent with previous studies that found that SMNR has no effect on speech intelligibility (Alcantara et al., 2003; Boymans & Dreschler, 2000; Ricketts & Hornsby, 2005; Stelmachowicz et al., 2010).

One possible explanation regarding the lack of positive effect is that although SMNR processing provided some improvement to the output SNR, the SNR improvement was small (i.e., mean = 1.6 dB; range = 0.7 to 2.4 dB) and did not improve to the extent of the unaided condition. This outcome is consistent with Marcoux et al.'s (2006) finding that SMNR processing had no effect on the discrimination of non-native speech contrasts in noise when 1.0 dB of SNR improvement was provided by the device they tested. Pittman (2011a, 2011b) also reported that there was no significant SMNR effect on speech perception by school-age children with an effective SNR improvement of 2.0 dB only. Gustafson et al. (2014) also measured speech recognition response of children for nonsense words processed by a hearing aid (Oticon Agil Pro) under SMNR-on and SMNR-off condition and found that speech recognition performance was not significantly different between SMNR-on and SMNR-off, despite effective SNR improvement of 3.8 dB and 1.5 dB for the 0 dB and +5 dB input SNR conditions, respectively. Similar to the hearing aid tested in the Gustafson et al. (2014) study, the hearing aid used in the current study was from the same manufacturer but a different model (Oticon Safari). Both hearing aids (Oticon Agil Pro and Oticon Safari) have MBNR. The results of this current study support the consistent findings in the literature that when SMNR processing provided an SNR improvement of less than 2.0 dB, it does not degrade novel speech sound identification (in noise) by adult listeners with normal hearing.

Even though SMNR had some acoustic effects on the amplitude and spectral mean of frication noise in the SMNR processed retroflex sounds (as describe in Chapter 3), these acoustic effects did not impede the learning of identification of these sounds. Median values of spectral mean difference for the Mandarin retroflex consonants (written as CH, ZH, SH in *PinYin*) spoken by the female talker ranged from -33 to -241 Hz, except in the vowel /u/ context where the median value was -834 Hz for CH, -418 Hz for ZH and -428 Hz for SH. For the Mandarin retroflex consonants spoken by the male talker, median values of spectral mean difference ranged from -44 to -277 Hz. Despite this amount of spectral mean reduction, the identification score still remained consistent for both groups of subjects. It is likely that listeners used other acoustic information in the vowel and transitions to distinguish the three Mandarin retroflex consonants. Studies have shown that two acoustic cues that can contribute to the distinction of affricate-fricative contrast include amplitude rise time and frication duration. For example, Tsao et al. (2006) found that amplitude rise-time is effective for non-native listeners to discriminate synthetic Mandarin alveolo-palatal fricatives and aspirated affricates, whereas frication duration is effective for discriminating aspirated and unaspirated affricates, as well as fricatives and unaspirated affricates. For English listeners, frication duration is perceptually more effective than rise time for distinguishing fricatives from affricates in English (Kluender & Walsh, 1992). However, both of these cues are not included in the analysis of the second experiment of this dissertation (Chapter 3) because they are not likely to be affected by SMNR processing. Because other acoustic cues might be used by the listeners to distinguish the Mandarin retroflex fricative and affricate contrasts, the effect of amplitude and spectral difference caused by SMNR processing may be overcome and therefore listeners in the SMNR-on group still can distinguish the three Mandarin retroflex consonants. It should be noted that the effect of frication-noise and release burst intensity reduction on audibility was not assessed in the current study since the stimuli were presented at a suprathreshold level.

The results also showed that there were improvements in the raw scores of the post-test sessions as compared to the baseline test. The results of the simple main effect testing for the three-way interaction of Training Voice x Test Voice x Session showed that the RAU scores in post-test 3 and post-test 4 were significantly higher than in baseline for all groups across conditions (see Figure 20). These results indicated that learning occurred and SMNR had no effect on the learning of novel speech sounds. These findings also confirmed the hypothesis that at least two training sessions would be required to see a significant improvement from the baseline results. This finding is consistent with previous research where short-term laboratory training with feedback has been shown to facilitate learning of new sound categories by naïve listeners (Flege, 1989; Flege, 1995; Jamieson & Morosan, 1986; Pisoni et al., 1982; Samuel & Kraljic, 2009).

The identification test results showed that the group trained with male voice had higher RAU scores than the group trained with female voice when tested with the male-voice stimuli. However, it is interesting to see that the group who received training in male voice did not perform as well in the female-voice test. The group that was trained with female voice did poorer in both female voice test and male voice test as compared to the other group. The fact that subjects who received training in male voice outperformed the group who received training in female voice when tested with VCV tokens spoken by the male talker and not when tested with the female VCV tokens indicated that the learning in male voice did not transfer to female voice. This could be explained by the talker variability effect: studies of non-native perception suggested that when non-native listeners were trained with single-talker stimuli, learning may not generalize to speech produced by other talkers. Likewise, non-native listeners trained with multiple-speaker stimuli will be able to generalize learning to unfamiliar speaker (Lively, Logan, & Pisoni, 1993; Logan et al., 1991; Nygaard & Pisoni, 1998). Recall that listeners in the current study were only trained with VCV tokens spoken by a single talker (either female or male) but were tested with VCV tokens spoken by both talkers. The transfer of learning to the other talker may not be as effective due to this talker variability effect. However, this could not explain why the group that learned with the female voice did not perform better in the female voice test. One possible underlying cause could be that there is a gender difference in the test stimuli. In the current study, the amplitude of the consonants spoken by the female talker was slightly lower than the male talker. Examination of the raw data for the frication noise amplitude measurement in Chapter 2 revealed that there was a 3 to 6 dB of difference between the frication noises produced by the female talker as compared to the male talker. This implied that although the RMS voltage of the VCV tokens was equalized to the same level between the female and male talker, there could be differences in the vowel-consonant amplitude. Further studies should take into account the vowel-consonant amplitude differences when using multiple talker stimuli. The HASQI nonlinear index for the retrieved word string containing the Mandarin retroflex consonants was slightly higher (better) for the male talker (0.64) than the female talker (0.60). Previous studies have reported talker gender difference in the acoustics of fricatives /s/ and /ʃ/ (Pittman & Stelmachowicz, 2000; Stelmachowicz et al., 2001). Pittman and Stelmachowicz (2000) reported that /s/ in a nonsense syllable /us/ spoken by a female talker had higher frequency peak energy (6 kHz) than a male talker (4 kHz). Stelmachowicz et al. (2001) reported that adult listeners with normal hearing had lower recognition score for the female voice when the stimuli was low-pass filtered at 6000 Hz (approximately 70% correct) and 5000 Hz (approximately 30-40% correct) as compared to the male voice (approximately

100% correct for both cut-off frequencies). Spectrograms of nonsense syllable /iʂi/, from the retrieved-speech file of the female and male talker, are shown in Appendix N. The spectrograms showed that the VCV syllables, retrieved from the speech-plus-noise recordings of HA#1 at -5dB SNR, were low-pass filtered. Other possible reasons may include (i) the particular articulation of each retroflex sound by the female talker, a stronger masking effect of the background noise on the female frication noise as compared to the male frication noise (e.g., due to the higher frequency energy in the female frication noise), (iii) or the interaction of these factors contributed to this talker-specific pattern. However, further investigation is clearly warranted to test these speculations.

4.5.2 Reaction time

Reaction time was collected as supplementary data and I hypothesized that the SMNR-on group would have shorter reaction time than the SMNR-off group. The shorter reaction time in the baseline test as compared to post-test 1 may suggest that the subjects were simply guessing and therefore did not require much processing time in order to respond. Therefore, the mean reaction time for the baseline test was excluded from the statistical analysis. Statistical test results indicated that SMNR had no significant effect on the reaction time of adult listeners in identifying novel speech contrasts and did not support the hypothesis. However, the results showed that on average, subjects started with a longer mean reaction time in post-test 1 and the mean reaction time decreased in subsequent post-test sessions. It is possible that subjects were still learning the new sound categories introduced to them during post-test 1 and therefore required longer processing time when responding. As subjects became more familiar with the Mandarin retroflex fricative and affricates, the time required for them to respond became shorter; this was reflected in the mean reaction time in post-test 2 through post-test 4. The fact that the mean reaction time in post-test 3 and 4 was significantly shorter than post-test 1 and post-test 2 but was not significantly different from the mean reaction time in the baseline session may suggest that the subjects' responses became more automated once they have learned the new speech categories. It was also noted that as the identification performance increased with post-test sessions, the reaction time also decreased with the post-test sessions. This pattern of reaction time and its relation to the identification performance was consistent with the findings reported by Logan et al. (1991) and Lively et al. (1993), in which Japanese listeners had to identify the English /l-r/ contrasts in both studies. The English /l-r/ contrasts are considered non-native to the Japanese listeners because there is no such contrast in the Japanese language. Logan et al. (1991) found that the subjects had longer mean reaction time during the second week of training sessions. However, the mean reaction time reversed (became shorter) during the third week of training. In another study, Lively et al. (1993) found that the reaction time decreased from the first week of training through the third week of training.

One caveat of the current study regarding reaction measurement was that subjects were unaware that they were timed. Nonetheless, the mean reaction times measured in the current study are comparable to the mean reaction time reported by Lee et al. (2012). Lee et al. (2012) measured reaction time and identification performance of English-speaking adult listeners in identifying Mandarin fricative consonants in background noise and found that reaction time increased as the SNR became poorer from 0 to -15 dB SNR. In their study, subjects were instructed to respond as fast as they could during the identification task and the mean reaction time at -5 dB SNR was approximately 1100 ms. In the current study, the mean reaction time for the subjects (in the SMNR-off group) to identify Mandarin retroflex fricative and affricates ranged from 908 to 1102 ms for all post-tests, which is comparable to the mean reaction time reported by Lee et al. (2012), suggesting that the lack of instruction to respond quickly was not likely a factor in the results of this current study.

4.5.3 Limitations

The present study has the following limitations. First, the perception of adult listeners with normal hearing may be different from children with normal hearing and listeners (both adults and children) with hearing impairment. Future research should test children and individuals with hearing impairments. If the testing and training paradigm in the current study is to be used to test children, modifications to the training paradigm are necessary. The training paradigm implemented in the current study included identification training with immediate feedback (e.g., feedback was provided for both correct and incorrect responses) and four sessions of training. Immediate feedback was intended to facilitate learning of new speech category during the training sessions in this study. In this training paradigm, each listener was exposed to 180 trials in each day of training. Each phoneme was presented 60 times (20 repetitions X 3 vowel contexts X 1 talker gender) during each training session. When testing children, modification such as shorter training time or other types of reinforcements (i.e., animated video games) can be incorporated. Recall that the stimuli were presented at comfortable listening level to the listeners, and thus the effect of audibility was not examined in the current study. Audibility should be quantified when testing the hearing-impaired population in future studies.

Second, there was no control group that did not receive any training. Therefore, the effect of repeated exposure to the VCV tokens as compared to the effect of training with feedback on the identification performance could not be separated. Third, the categorization task used in the current study not only involved the discrimination ability of the naïve listeners, but also required at least another higher order processes such as the learning of the task (e.g., motor learning) and the learning of the sound-to-picture mapping (e.g., working memory). Other factors such as fatigue and motivation were also not evaluated in the current study. Therefore, it is uncertain if these factors had any effect on the research findings. Other improvements that can be incorporated in the study are to include (i) native Mandarin listeners as the control group and (i) a retention test if one is interested in knowing whether the learning of novel speech sound could be retained at a later time.

The fourth limitation is that a low SNR condition (e.g., -5 dB SNR) may not occur very often in our everyday listening situation. However, it is reasonable to test listeners at the -5 dB SNR condition for two reasons. First, acoustic measurements in my second experiment (Chapter 3) showed that the effect of SMNR was stronger in the poorer SNR conditions (i.e., -5 and -10 dB SNR) as compared to the better SNR conditions (i.e., +10, +5, and 0 dB SNR). Second, pilot studies conducted showed that subjects performed at chance level in the identification test at the poorest SNR condition (i.e., -10 dB SNR). Thus, a balance point between the strongest SMNR effects without reaching the floor effect was the -5 dB SNR condition. In addition, one study also showed that adults with normal hearing can correctly identify the place of articulation of plosive and fricative consonants even at -5 dB SNR (Alwan, Jiang, & Chen, 2011). For these reasons, it was deemed appropriate to use the -5 dB SNR condition in the identification tests.

In conclusion, the results of this present study imply that modulation-based SMNR had no detrimental effect on the learning of novel speech sound identification by naïve listeners. The results could also indirectly imply that modulation-based SMNR has no detrimental effect on the identification of Mandarin retroflex consonants in background noise by Mandarin listeners, although further investigation is required to confirm this implication. This finding is in agreement with the previous SMNR studies, despite the increased difficulty of the identification task using Mandarin retroflex speech contrasts to test naïve listeners. This study also provided an extension of the cross-language approach used by Marcoux et al (2006) and the Turgeon et al. (2009) in examining the effect of SMNR on speech perception in noise. The different aspects of the current study as compared to the two previous studies included (i) the use of Mandarin speech contrasts (fricative-affricate) that are more susceptible to the acoustical effect of SMNR processing; (ii) the use of a three-response choice identification task vs. a same-different discrimination task; and (iii) the incorporation of training with feedback to facilitate the learning of novel speech categories. It should be noted that the current study tested only one type of SMNR and the results are only applicable modulation-based SMNR. In addition, the current study tested only a small set of Mandarin speech sounds. The results could have been different if a larger set of stimulus or an open set response was used. Future research should test other types of SMNR in different makes and models of hearing aids and to incorporate suggestions mentioned in the previous paragraphs to answer other research questions arise from the current research.

Chapter 5 Conclusion

In this dissertation, three studies were conducted to examine the acoustic and perceptual effects of single-microphone noise reduction (SMNR) on fricatives and affricates. The focus of the first study was to validate the inversion technique (Hagerman & Olofsson, 2004) for extracting speech post SMNR-processing. In order to examine the acoustic effects of SMNR on speech in noise, it is necessary to present speech-plus-noise signals to a hearing aid for processing. Subsequently, both signals have to be extracted post SMNR-processing for further acoustic analysis. The research questions of the first study were "Is the inversion technique a feasible tool for separating aided and unaided speech-plus-noise signals recorded in a sound field setting?" and "Does the inversion technique affect the fidelity of retrieved-speech signals?".

The aim of the second study was to examine the acoustic effects of SMNR on fricatives and affricates of the two most common languages in the world, Mandarin and English. Fricatives and affricates were chosen as test stimuli as compared to the other consonant categories because they have a noise-like feature that mimics the randomness of broadband noise and may be affected by SMNR processing when background noise is present. Mandarin and English were chosen because these two languages have the highest numbers of users in the world and fricatives and affricates are important in both languages.

The aim of the third study was to examine the perceptual effects of SMNR on novel speech sound identification in noise by naïve listeners. Mandarin retroflex fricative and affricates, processed with and without SMNR, were used to examine the effects of SMNR on novel speech sound identification in noise by native English-speaking adults with normal hearing. These speech sounds were selected because they are not within the phonemic inventory of the English talkers. The retrieved-speech signals from the recordings of HA#1 were used in Study 3 for a few reasons. First, the retrieved-speech signals had better speech fidelity (as indicated by higher amount of attenuation in Study 1) as compared to HA#2. Second, acoustic measurements in the second study showed that the acoustic change as a result of SMNR processing in this hearing aid were more consistent across consonants, vowel contexts, and talkers between the two languages. Thus, the results obtained in the third study can be generalized to other consonants tested in the second study. Although each of these studies can stand on its own, they were related in a way that the results of the first study were used to determine which sets of retrieved speech should be excluded from the second study and the results of the second study were used to determine which subset of speech stimuli would be used as the test stimuli in the third study.

5.1 Summary of Findings

The first study showed that the inversion technique (Hagerman & Olofsson, 2004) is a feasible tool for extracting speech from the speech-plus-noise recordings obtained in a sound field setting; this implementation also has good test-retest reliability. These results supported the two hypotheses of the study where (i) the amount of attenuation was greater when there were fewer sources of error and (ii) the inversion technique has good test-retest reliability. However, the hypothesis regarding the effect of the inversion technique on speech fidelity was only partially fulfilled because there was some variability in the findings between the two hearing aids tested in this study. Generally, fidelity of the retrieved-speech signals (aided and unaided) was not affected between 0 to +15 dB SNR but fidelity was affected when the recordings were made under poor SNR conditions (i.e., -10 and -5 dB SNR).

The second study showed that modulation-based SMNR caused acoustic changes to fricatives and affricates. The amount of acoustical change differed between hearing aids. The acoustic changes were more prominent in the amplitude domain; mixed results were obtained in the spectral domain, depending on the types of acoustic measurements conducted. First, the study found that modulation-based SMNR was effective in reducing steady-state broadband noise and the amount of gain reduction was level-dependent and manufacturer-dependent. Second, SMNR processing resulted in a change of the output SNR for speech-plus-noise signals, with one hearing aid exhibiting less SNR improvement than the other. These results were consistent with the previous findings reported in the literature. Third, speech fidelity measurement using the Hearing Aid Speech Quality Index (HASQI; Kates & Arehart, 2010) showed that SMNR processing in both hearing aids had minimal effect on the long-term spectra and the modulation envelope of the word strings containing Mandarin and English fricatives and affricates. Fourth, spectrographic analysis showed that SMNR reduced the amplitude of frication noise and release burst of fricatives and affricates; this finding is consistent with the trend observed in the amount of noise reduction measurement. However, spectrographic analysis showed that SMNR processing caused changes in the spectral mean of the frication noise and the amount of spectral mean change differed between hearing aids. This finding contradicted with the linear index of the HASQI measurement. It was unknown whether these acoustical changes would affect the perception of fricatives and affricates, therefore, the third study was designed to answer this question by using a subset of Mandarin nonsense words as the test stimuli.

The third study suggested that modulation-based SMNR did not impede the learning of Mandarin retroflex fricative and affricates identification by naïve listeners with normal hearing. The results showed that there was no significant SMNR main effect or other interaction effects involving SMNR on the identification performance of Mandarin retroflex fricative and affricates. This finding did not support the first hypothesis of the study. However, the finding may suggest that SMNR had no detrimental effect on learning of novel speech sound categorization. Overall, the results of this study were in agreement with the previous SMNR studies, despite the perceptual task being made more challenging by using Mandarin retroflex fricative and affricates that were novel to the listeners. Recall that the acoustical changes as a result of SMNR processing in HA#1 were consistent across consonants, vowel context, and talkers between Mandarin and English, the findings for Mandarin retroflex fricatives and affricates may be generalized to other fricatives and affricates tested in this dissertation.

5.2 Limitations and Future Studies

One of the limitations of the first study is that the residual-speech measurement may overestimate the effect of inversion technique on the fidelity of the retrieved-speech signals because the inversion technique was performed twice in the measurement (i.e., first to retrieve speech signals from the speech-plus-noise signals and second to obtain residual speech). This limitation also applies to the test-retest reliability measurement where the inversion technique was performed three times throughout the procedure. Therefore, the residual-speech signals for the speech fidelity measurement and the residual signal for test-retest reliability could have additional error introduced by the inversion technique. In other words, the actual residual speech could be lower than the values measured in the first study and the error in the test-retest reliability could be smaller.

For all three studies, only one type of SMNR (modulation-based noise reduction) implemented in two commercial hearing aids was tested. The results cannot be generalized to other types of SMNR processing (e.g., synchrony detection and spectral subtraction). Although the different findings between the two hearing aids could be attributed to the proprietary SMNR algorithm in each hearing aid, the studies did not control for number of processing channels within each hearing aid. In addition, all other advanced features such as compression and frequency lowering were turned off and it is unknown whether the interactions between SMNR and other advanced features will cause other acoustical changes to the speech signals. Future research should test hearing aids within and across manufacturers and to control for number of processing channels and technology levels (e.g., entry range, mid-range, and premium range).

It is also acknowledged that perception of adult listeners with normal hearing may be different from that of children or listeners with hearing impairment. Future research should test children because of the uncertainty regarding the use of SMNR with this population. If the testing and training paradigm in the third study is to be used to test the pediatric population, modifications such as shorter training time or other types of reinforcement need to be incorporated. In addition, future research should test individuals with hearing impairment and measures of audibility should be incorporated when testing this population.

Due to the redundancy of speech cues, future research should also examine the effects of SMNR processing on other acoustical cues that are associated with the perception of other speech categories. Future research should also test a larger data set with more talkers because the third study found that training using speech stimuli spoken by one talker does not generalize to the other talker. A control group that does not receive any training can be included to examine the effect of repeated exposure to the test stimuli and the effect of training paradigm on the identification performance of novel speech sounds. A group of native Mandarin talkers may also be included as a comparison in the cross language approach. Retention testing can be conducted if one is interested in knowing whether the learning of novel speech sound could be retained at a later time. Other factors such as ability to discriminate novel speech sounds and other higher order processes like memory can be evaluated in future studies.

5.3 Implications

The three studies conducted as part of this dissertation contribute further evidence about the acoustic and perceptual effects of SMNR on fricatives and affricates. First, the inversion technique was shown to be a feasible tool for extracting speech and noise post hearing-aid processing when the recordings were collected in a sound field setting under input SNR conditions between 0 to +15 dB SNR. Since the fidelity measurement of the aided retrieved speech showed some variability between hearing aids, it is recommended that verification of speech fidelity be conducted for aided recordings, particularly the ones recorded under a negative SNR condition, before proceeding with acoustic measurements. The inversion technique can be applied in a research setting to separate acoustic signals post hearing aid processing. It could be a potential measurement tool that can be incorporated in routine electroacoustic measurement of a hearing aid. However, further refinement on this measurement should be carried out before it can be used clinically.

Second, the acoustic effects of SMNR on Mandarin and English fricatives and affricates were systematically documented in the second study and the acoustic measurements adapted in the study may offer an alternative way to quantify the efficacy of

SMNR between hearing aids. The study showed that modulation-based SMNR caused acoustic changes on fricatives and affricates, particularly in the amplitude domain. This information from the second study can be used to counsel hearing aid users about the efficacy of a modulation-based SMNR on steady-state background noise and to build realistic expectations by the hearing aid users when listening to speech in a noisy environment. It is clear from the results of this study that gain reduction caused an amplitude change to the noise signal as well as the speech signals. This study also showed that the amount of acoustic change differed between SMNR in different make and models of hearing aids. This finding indicated that the efficacy of SMNR of one make and model of hearing aid may not be used to predict the efficacy of SMNR in other makes and models of hearing aids. Electroacoustic measurements such as suggested by Scollie et al. (2016) should be incorporated in the routine hearing aid evaluation.

Third, despite the acoustical changes documented in the second study, the perceptual test conducted in the third study showed that SMNR processing did not impede the learning of Mandarin retroflex sound identification in background noise by naïve listeners. Since the acoustic effects of SMNR in HA#1 were quite consistent on the Mandarin and English fricatives and affricates tested in the second study, we can infer that modulation-based SMNR processing in HA#1 is not likely to have a perceptual effect on other fricatives and affricates not tested in the third study. Because the strength of the SMNR in HA#1 was set to a maximum degree in the study, we can also infer that the SMNR is not likely to have a negative effect when the SMNR is set to a lower strength. In addition, because there was no perceptual effect when the listeners was tested at -5 dB SNR condition, we could also infer that SMNR in HA#1 do not have any detrimental effect when the input SNR for speech-plus-noise signal is higher than -5 dB SNR (i.e., at better input SNR condition).

The findings of the studies also raise a number of research questions: (i) Does SMNR affect other speech cues not measured in Study 2 (such as frication duration, formant transition, and rise time)? How do these speech cues contribute in the perception of fricatives and affricates in noise? (ii) Are there underlying stimulus or talker-gender factors that contribute to the learning effects for the different talker genders? (iii) Is the novel identification training paradigm feasible in children? (iv) What would be the outcome if Study 3 is repeated by testing other populations such as children with normal hearing or individuals with hearing loss? The results of the current studies can be used as the baseline for comparison for anyone who is interested to pursue these questions. The method used in these studies can serve as a model to examine the effect of other advanced hearing aid features on the acoustic and perception of speech sounds. One of such hearing aid features is the frequency lowering technology that aims to provide access of high frequency speech sounds to hearing aid users.

The significant contributions of the research conducted as part of this dissertation are (i) systematically document the acoustic effects of SMNR on Mandarin and English fricatives and affricates, (ii) providing evidence that SMNR has no detrimental effect on learning to categorize novel speech sounds (e.g., Mandarin retroflex sounds), (iii) providing an extension of the cross-language approach used in previous studies in examining the effect of SMNR on speech perception, and (iv) defining the non-meaningful acoustic changes. Overall, the results of the current research provide further evidence to support the use of SMNR in hearing aid users.

References

Alcantara, J. L., Moore, B. C., Kuhnel, V., & Launer, S. (2003). Evaluation of the noise reduction system in a commercial digital hearing aid. *International Journal of Audiology, 42*(1), 34-42.

Alwan, A., Jiang, J., & Chen, W. (2011). Perception of place of articulation for plosives and fricatives in noise. *Speech Communication, 53*(2), 195-209.

American Academy of Audiology. (2013, June). Clinical practice guidelines on pediatric amplification. 1-60. Retrieved May 4, 2014 from http://www.audiology.org /resources/documentlibrary/Pages/PediatricRehabilitation.aspx

ANSI S3.5-1997 (1997). *American National Standard: Methods for the Calculation of the Speech Intelligibility Index.* American national Standards Institute, New York.

Bagatto, M. P., Scollie, S. D., Hyde, M., & Seewald, R. (2010). Protocol for the provision of amplification within the Ontario infant hearing program. *International Journal of Audiology, 49 Suppl 1*, S70-9. doi:10.3109/14992020903080751

Bentler, R., & Chiou, L. K. (2006). Digital noise reduction: An overview. *Trends in Amplification, 10*(2), 67-82. Doi: 10.1177/1084713806289514

Bentler, R., Wu, Y. H., Kettel, J., & Hurtig, R. (2008). Digital noise reduction: Outcomes from laboratory and field studies. *International Journal of Audiology, 47*(8), 447-460. Doi: 10.1080/14992020802033091

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M.J. Munro & O.S. Bohn (Eds.), *Second Language Speech Learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Boymans, M., & Dreschler, W. A. (2000). Field trials using a digital hearing aid with active noise reduction and dual-microphone directionality. *Audiology, 39*, 260-268.

Bray, V., & Nilsson, M. (2001, December 5). Additive SNR benefits of signal processing features in a directional DSP aid. Retrieved June 25, 2015, from http://www.hearingreview.com/2001/12/additive-snr-benefits-of-signal-processing-features-in-a-directional-dsp-aid/

Brons, I., Houben, R., & Dreschler, W. A. (2013). Perceptual effects of noise reduction with respect to personal preference, speech intelligibility, and listening effort. *Ear and Hearing, 34*(1), 29-41. doi:10.1097/AUD.0b013e31825f299f [doi]

Chong, F.Y. & Jenstad, L.M. (2010, August). The effects of digital noise reduction algorithms on speech and noise. Poster presented at the International Hearing Aid Research Conference, Lake Tahoe, CA.

Chong, F.Y. & Jenstad, L.M. (2011, October). The effects of single microphone noise reduction on the acoustics of /s/ and /z/. Poster presented at the Canadian Academy of Audiology Conference, Victoria, BC, Canada.

Chen, F., Hu, Y., & Yuan, M. (2015). Evaluation of noise reduction methods for sentence recognition by mandarin-speaking cochlear implant listeners. *Ear and Hearing, 36*(1), 61-71. doi:10.1097/AUD.0000000000000074 [doi]

Chung, K. (2004). Challenges and recent developments in hearing aids. Part I. speech understanding in noise, microphone technologies, and noise reduction algorithms. *Trends in Amplification, 8*(3), 83-124.

Denes, P. B. (1963). On the statistics of spoken English. *The Journal of the Acoustical Society of America, 35*(6), 892-904. doi:10.1121/1.1918622

Edwards, B. W. (2000). Beyond amplification: Signal processing techniques for improving speech intelligibility in noise with hearing aids. *Seminars in Hearing, 21*(2), 137-156

Elberling, C. (2002, January). *About the VoiceFinder.* News from Oticon. Retrieved May 7, 2014, from http://phonicear.fr/eprise/main/Oticon/CA_en/Sec_Professionals /TechnologyandAudiology/AddditionalInformation/OticonPubllications/Publications/

Ellaham, N., Giguere, C., & Gueaieb, W. (2013). Evaluation of the phase-inversion signal separation method when using nonlinear hearing aids. *IEEE transactions on Audio, Speech, Language Processing, 21*(4), 879-888.

Falk, T. H., Parsa, V., Santos, J. F., Arehart, K., Hazrati, O., Huber, R., . . . Scollie, S. (2015). Objective quality and intelligibility prediction for users of assistive listening devices: Advantages and limitations of existing tools. *Signal Processing Magazine, IEEE, 32*(2), 114-124.

Flege, J. E. (1989). Chinese subjects' perception of the word-final English/t/–/d/contrast: Performance before and after training. *The Journal of the Acoustical Society of America, 86*(5), 1684-1697.

Flege, J. E. (1995). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics, 16*, 425-442.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America, 84*(1), 115-123.

Fredelake, S., Holube, I., Schlueter, A., & Hansen, M. (2012). Measurement and prediction of the acceptable noise level for single-microphone noise reduction algorithms. *International Journal of Audiology, 51*(4), 299-308. Doi: 10.3109/14992027.2011.645075

Gamst, G., Meyers, L. S., & Guarino, A. J. (2008). Analysis of variance designs: A conceptual and computational approach with SPSS and SAS. Cambridge University Press.

Ghent, R., Nilsson, M., & Bray, V. (2007, June). Validation of the noise reduction index (NRI). Poster presented at the 153rd Meeting of the Acoustical Society of America, Salt Lake City, Utah.

Gustafson, S., McCreery, R., Hoover, B., Kopun, J. G., & Stelmachowicz, P. (2014). Listening effort and perceived clarity for normal-hearing children with the use of digital noise reduction. *Ear and Hearing, 35*(2), 183-194. doi:10.1097/01.aud.0000440715.85844.b8 [doi]

Hagerman, B., & Olofsson, Å. (2004). A method to measure the effect of noise reduction algorithm using simultaneous speech and noise. *Acta Acustica United with Acustica, 90*(2), 356-361.

Hamacher, V., Chalupper, J., Eggers, J., Fischer, E., Kornagel, U., Puder, H., & & Rass, U. (2005). Signal processing in high-end hearing aids: State of the art, challenges, and future trends. *EURASIP Journal on Applied Signal Processing, 2005*(18), 2915-2929. Doi: 10.1155/ASP.2005.2915

Hedrick, M. (1997). Effect of acoustic cues on labeling fricatives and affricates. *Journal of Speech, Language, and Hearing Research: JSLHR, 40*(4), 925-938.

Hoetink, A. E., Korossy, L., & Dreschler, W. A. (2009). Classification of steady state gain reduction produced by amplitude modulation based noise reduction in digital hearing aids. *International Journal of Audiology, 48*(7), 444-455. Doi: 10.1080/14992020902725539

Howell, P., & Rosen, S. (1983). Production and perception of rise time in the voiceless affricate/fricative distinction. *The Journal of the Acoustical Society of America, 73*(3), 976-984.

Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard chinese). *Journal of Child Language, 27*(1), 3-42.

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English/ð/-/θ/contrast by francophones. *Perception & Psychophysics, 40*(4), 205-215.

Johannesson, R.B. (2006). *Output SNR Measurement method* (Report 052-08-04). Eriksholm: Oticon research centre.

Jenstad, L. M., & Zakis, J.A. (2011). Using an inversion technique to quantify non-linear hearing aid processing. Poster presented at the American Auditory Society, Scottsdale, Arizona.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108*(3), 1252-1263.

Kates, J. M. (2008). *Digital hearing aids*. San Diego: Plural Publishing Inc.

Kates, J.M. (2010, September). Intelligibility and Quality Metrics Version 1. GN ReSound Internal Report Version 1.5.

Kates, J. M., & Arehart, K. H. (2005). Coherence and the speech intelligibility index. *The Journal of the Acoustical Society of America, 117*(4), 2224-2237.

Kates, J. M., & Arehart, K. H. (2010). The hearing-aid speech quality index (HASQI). *Journal of the Audio Engineering Society, 58*(5), 363-381.

Kates, J. M., & Arehart, K. H. (2014). The hearing-aid speech quality index (HASQI) version 2. *Journal of the Audio Engineering Society, 62*(3), 99-117.

Keidser, G., Hartley, D., & Carter, L. (2008). Long-term usage of modern signal processing by listeners with severe or profound hearing loss: A retrospective survey. *American Journal of Audiology, 17*(2), 136-146.

King, A. M. (2010). The national protocol for paediatric amplification in Australia. *International Journal of Audiology, 49 Suppl 1*, S64-9. doi:10.3109/14992020903329422

Kluender, K. R., & Walsh, M. A. (1992). Amplitude rise time and the perception of the voiceless affricate/fricative distinction. *Perception & Psychophysics, 51*(4), 328-333.

Kochkin, S. (2010). MarkeTrek VIII: Consumer satisfaction with hearing aids is slowly increasing. *The Hearing Journal, 63*(1), 19-27.

Kortekaas, R. W., & Stelmachowicz, P. G. (2000). Bandwidth effects on children's perception of the inflectional morpheme /s/: Acoustical measurements, auditory detection, and clarity rating. *Journal of Speech, Language, and Hearing Research: JSLHR, 43*(3), 645-660.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*(2), F13-F21.

Kuk, F., Peeters, H., Lau, C., & Korhonen, P. (2011). Effect of maximum power output and noise reduction on speech recognition in noise. *Journal of the American Academy of Audiology, 22*(5), 265-273. Doi: 10.3766/jaaa.22.5.3

Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (sixth edition). Boston, USA: Wadsworth, Cengage Learning.

Lai, Y. (2009). Asymmetry in Mandarin affricate perception by learners of Mandarin Chinese. *Language and Cognitive Processes, 24*(7-8), 1265-1285.

Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication, 52*(11), 864-886.

Lee, S. I. (2011). Spectral analysis of Mandarin Chinese sibilant fricatives. In *The 17th International Congress of Phonetic Sciences (ICPhS XVII)* (pp. 1178-1181).

Lee, C. Y., Zhang, Y., Li, X., Tao, L., & Bond, Z. S. (2012). Effects of speaker variability and noise on Mandarin fricative identification by native and non-native listeners. *The Journal of the Acoustical Society of America, 132*(2), 1130-1140. Doi: 10.1121/1.4730883; 10.1121/1.4730883

Lewis, M. P., Gary F. S., & Charles D. F. (eds.). (2015). Ethnologue: Languages of the World, Eighteenth edition. Dallas, Texas: SIL International. Retrieved June 12, 2015, from http://www.ethnologue.com.

Li, F., Edwards, J., & Beckman, M. (2007). Spectral measures for sibilant fricatives of English, Japanese, and mandarin Chinese. *Proceedings of the XVIth International Congress of Phonetic Sciences, , 4* 917-920.

Li, J., Yang, L., Zhang, J., Yan, Y., Hu, Y., Akagi, M., & Loizou, P. C. (2011). Comparative intelligibility investigation of single-channel noise-reduction algorithms for Chinese, Japanese, and English). *The Journal of the Acoustical Society of America*, *129*(5), 3291-3301.

Liu, H., Zhang, H., Bentler, R., A., Han, D., & Zhang, L. (2012). Evaluation of a transient noise reduction strategy for hearing aids. *Journal of the American Academy of Audiology, 23*(8), 606-615. doi:10.3766/jaaa.23.8.4

Liu, Q., Zhou, N., Berger, R., Huang, D., & Xu, L. (2013). Mandarin consonant contrast recognition among children with cochlear implants or hearing aids and normal-hearing children. *Otology & Neurotology : Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology, 34*(3), 471-476. doi:10.1097/MAO.0b013e318286836b [doi]

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English/r/and/l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242-1255.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America, 89*(2), 874-886.

Loizou, P. C. (2013). *Speech enhancement: Theory and practice*. CRC press.

Loven, F. (2009). *Introduction to normal auditory perception*. Delmar Cengage Learning.

Marcoux, A. M., Yathiraj, A., Cote, I., & Logan, J. (2006). The effect of a hearing aid noise reduction algorithm on the acquisition of novel speech contrasts. *International Journal of Audiology, 45*(12), 707-714. Doi: 10.1080/14992020600944416

McCreery, R. W., Venediktov, R. A., Coleman, J. J., & Leech, H. M. (2012). An evidence-based systematic review of directional microphones and digital noise reduction hearing aids in school-age children with hearing loss. *American Journal of Audiology, 21*(2), 295-312. Doi: 10.1044/1059-0889(2012/12-0014)

Mitani, S., Kitama, T., & Sato, Y. (2006). Voiceless affricate/fricative distinction by frication duration and amplitude rise slope. *The Journal of the Acoustical Society of America, 120(3)*, 1600-1607.

Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., . . . Stelmachowicz, P. (2007a). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I—phonetic development. *Ear and Hearing, 28*(5), 605-627. Doi:10.1097/AUD.0b013e31812564ab

Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., . . . Stelmachowicz, P. (2007b). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part II—transition to words. *Ear and Hearing, 28*(5), 628-642.

Mueller, H. G., Weber, J., & Hornsby, B. (2006). The effects of digital noise reduction on the acceptance of background noise. *Trends in Amplification, 10*(2), 83-93.

Munro, M. (2008). CVC syllables for investigating the phonetic sensitivity of Mandarin and Chinese speakers. *Behavior Research Methods, 40*(1), 147-153.

Nabelek, A. K. (2006). Acceptable noise level: A clinical measure for predicting hearing aid outcome. *Journal of the American Academy of Audiology, 17*(9), [624-5].

Nabelek, A. K., Tampas, J. W., & Burchfield, S. B. (2004). Comparison of speech perception in background noise with acceptance of background noise in aided and unaided conditions. *Journal of Speech, Language & Hearing Research, 47*(5), 1001-1011.

Naylor, G., & Johannesson, R. B. (2009). Long-term signal-to-noise ratio at the input and output of amplitude-compression systems. *Journal of the American Academy of Audiology, 20*(3), 161-171.

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., & Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *International Journal of Audiology, 52*(7), 433-441.

Ng, E. H., Rudner, M., Lunner, T., & Ronnberg, J. (2015). Noise reduction improves memory for target language speech in competing native but not foreign language speech. *Ear and Hearing, 36*(1), 82-91. doi:10.1097/AUD.0000000000000080 [doi]

Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America, 112*(2), 711-719.

Nordrum, S., Erler, S., Garstecki, D., & Dhar, S. (2006). Comparison of performance on the hearing in noise test using directional microphones and digital noise reduction algorithms. *American Journal of Audiology, 15*(1), 81-91.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60*(3), 355-376.

Olofsson, Å., & Hansen, M. (2006). Objectively measured and subjectively perceived distortion in nonlinear systems. *The Journal of the Acoustical Society of America, 120*(6), 3759-3769.

Olsen, W. O. (1998). Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearson, Bennett, & Fidell (1977) report. *American Journal of Audiology, 7*(2), 21.

Palmer, C. V., Bentler, R., & Mueller, H. G. (2006). Amplification with digital noise reduction and the perception of annoying and aversive sounds. *Trends in Amplification, 10*(2), 95-104.

Parikh, G., & Loizou, P. C. (2005). The influence of noise on vowel and consonant cues. *The Journal of the Acoustical Society of America*, *118*(6), 3874-3888.

Pearson, K.S., Bennett, R.L., & Fidell, S. (1977). Speech levels in various noise environments (Report No. EPA-600/1-77-025). Washington, DC: U.S. Environmental Protection Agency.

Peeters, H., Kuk, F., Lau, C. C., & Keenan, D. (2009). Subjective and objective evaluation of noise management algorithms. *Journal of the American Academy of Audiology, 20*(2), 89-98.

Peng, S., Weiss, A. L., Cheung, H., & Lin, Y. (2004). Consonant production and language skills in mandarin-speaking children with cochlear implants. *Archives of Otolaryngology—Head & Neck Surgery, 130*(5), 592.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance, 8*(2), 297.

Pittman, A. (2011a). Age-related benefits of digital noise reduction for short-term word learning in children with hearing loss. *Journal of Speech, Language, and Hearing Research: JSLHR, 54*(5), 1448-63. Doi: 10.1044/1092-4388(2011/10-0341)

Pittman, A. (2011b). Children's performance in complex listening conditions: Effects of hearing loss and digital noise reduction. *Journal of Speech, Language, and Hearing Research: JSLHR, 54*(4):1224-39. Doi: 10.1044/1092-4388(2010/10-0225)

Pittman, A. L., & Stelmachowicz, P. G. (2000). Perception of voiceless fricatives by normal-hearing and hearing-impaired children and adults. *Journal of Speech, Language, and Hearing Research: JSLHR, 43*(6), 1389-1401.

Repp, B. H., Liberman, A. M., Eccardt, T., & Pesetsky, D. (1978). Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 621.

Rhebergen, K. S., & Versfeld, N. J. (2005). A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. *The Journal of the Acoustical Society of America, 117*(4), 2181-2192.

Ricketts, T. A., & Hornsby, B. (2005). Sound quality measures for speech in noise through a commercial hearing aid implementing digital noise reduction. *Journal of the American Academy of Audiology, 16*(5), 270-277.

Rudmin, F. (1983). The why and how of hearing /s/. *Volta Review, 85*, 263-269.

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics, 71*(6), 1207-1218.

Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language & Hearing Research, 52*(5), 1230-1240. Doi: 10.1044/1092-4388(2009/08-0111)

Schaub, A. (2008). *Digital hearing aids*. New York: Thieme.

Schum, D. J. (2003). Noise-reduction circuitry in hearing aids: (2) goals and current strategies. *Hearing Journal, 56*(6), 32-40.

Scollie, S., Levy, C., Pourmand, N., Abbasalipour, P., Bagatto, M., Richert, F., . . . Parsa, V. (2016). Fitting noise management signal processing applying the American Academy of Audiology pediatric amplification guideline: Verification protocols. *Journal of the American Academy of Audiology, 27*(3), 237-251.

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., . . . Pumford, J. (2005). The desired sensation level multistage input/output algorithm. *Trends in Amplification, 9*(4), 159-197.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America, 125*(2), 1103-1113. doi:10.1121/1.3058638

Souza, P. E., Jenstad, L. M., & Boike, K. T. (2006). Measuring the acoustic effects of compression amplification on speech in noise. *The Journal of the Acoustical Society of America, 119*(1), 41-44.

Stelmachowicz, P., Lewis, D., Hoover, B., Nishi, K., McCreery, R., & Woods, W. (2010). Effects of digital noise reduction on speech perception for children with hearing loss. *Ear and Hearing, 31*(3), 345-355. Doi: 10.1097/AUD.0b013e3181cda9ce

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., & Lewis, D. E. (2001). Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *The Journal of the Acoustical Society of America, 110*(4), 2183-2190.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., & Lewis, D. E. (2002). Aided perception of /s/ and /z/ by hearing-impaired children. *Ear and Hearing, 23*(4), 316-324.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., Lewis, D. E., & Moeller, M. P. (2004). The importance of high-frequency audibility in the speech and language development of children with hearing loss. *Archives of Otolaryngology—Head & Neck Surgery, 130*(5), 556-562. Doi: 10.1001/archotol.130.5.556

Studebaker, G.A. (1985). A “rationalized” arcsin transform. *Journal of Speech and Hearing Research,* 28, 455-462.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America, 120*(4), 2285-2294.

Turgeon, C., Dostaler, M., Yathiraj, A., & Marcoux, A. (2009). Hearing aid noise reduction algorithms and the acquisition of novel speech contrasts by young children. *Canadian Journal of Speech-Language Pathology & Audiology, 33*(3), 140-145.

Walden, B. E., Surr, R. K., Cord, M. T., Edwards, B., & Olson, L. (2000). Comparison of benefits provided by different hearing aid technologies. *Journal of the American Academy of Audiology, 11*(10), 540-560.

Walker, G., Dillon, H., & Byme, D. (1984). Sound field audiometry: Recommended stimuli and procedures. *Ear and Hearing, 5*(1), 13-21.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*(1), 49-63.

Wu, Y., & Bentler, R. A. (2007). Using a signal cancellation technique to assess adaptive directivity of hearing aids. *The Journal of the Acoustical Society of America, 122*, 496.

Yuen, K., Kam, A., & Lau, P. (2006). Comparative performance of an adaptive directional microphone system and a multichannel noise reduction system. *Journal of the American Academy of Audiology, 17*(4), 241-252.

Zakis, J. A., Hau, J., & Blamey, P. J. (2009). Environmental noise reduction configuration: Effects on preferences, satisfaction, and speech understanding. *International Journal of Audiology, 48*(12), 853-867. Doi: 10.3109/14992020903131117

Zeng, F. G., & Turner, C. W. (1990). Recognition of voiceless fricatives by normal and hearing-impaired subjects. *Journal of Speech and Hearing Research, 33*(3), 440-449.

Zhao, X., & Li, P. (2009). An online database of phonological representations for mandarin Chinese. *Behavior Research Methods, 41*(2), 575-583.

Appendices

Appendix A: Single-microphone Noise Reduction Studies in Adults and Children

Year; Author	**Subject**	**HA model**	**SMNR type**	**Field trial**	**Acoustic Analysis**	**Objective test**	**Subjective Tests**			
						Speech intelligibility	**ANL**	**Paired comparison**	**Scale ratings**	**Self-report questionnaires**
2000; Boysman & Dreschler	N=16; HI (age 40-68)	BTE: Siemens Prisma	MBNR	✓	✓ Spectra of speech-in-noise	✓		✓ Preference		✓ Benefit
2000; Walden et al.	N=40; HI (age 52-76)	BTE: Resound BZ5	MBNR	✓		✓			✓ Speech understanding; Sound comfort; Sound quality/ naturalness	
2001; Bray & Nillson	N=20; HI (age 34-84)	BTE: Sonic Innovations Natura 2E	MBNR			✓				
2003; Alcantara et al.	N=8; HI (age 49-83)	ITE: Phonak Claro 21 DAZ	MBNR	✓		✓			✓ Comprehension; Listening comfort; Listening effort; Sound quality; Sound clarity	

Appendix A: Single-microphone Noise Reduction Studies in Adults and Children (continue)

Year; Authors	Subjects	HA model	SMNR type	Field trial	Acoustic Analysis	Objective test	Subjective Tests			
						Speech intelligibility	**ANL**	**Paired comparison**	**Scale ratings**	**Self-report questionnaires**
2005; Dahlquist et al.	N=60; HI (age 33-81)	Computer-based noise reduction	Spectral subtraction			✓ (Note: Subjects listened to PC-NRA pre-processed stimuli through a linear aid.)		✓ Preference; Comfort; Clarity; Loudness		
2005; Ricketts & Hornsby	N=14; HI (age 42-83)	BTE: Siemens Triano S	MBNR+WF			✓		✓ Preference	✓ Strength of preference	
2006; Mueller et al.	N=22; HI (age 23-76)	BTE: Siemens Acuris S	MBNR+WF			✓	✓			
2006; Marcoux et al.	N=20; NH (age 21-41)	BTE: Widex Senso Diva	MBNR			✓ (Discrimination of Hindi dental and retroflex stops.)				
2006; Nordrum et al.	N=16; HI (age 58-90)	BTEs: GN ReSound Canta 7, Oticon Syncro, Phonak Perseo, Siemens Acuris.	MBNR; Synchrony detection			✓				

Appendix A: Single-microphone Noise Reduction Studies in Adults and Children (continue)

Year; Authors	Subjects	HA model	SMNR type	Field trial	Acoustic analysis	Objective test	Subjective Tests			
						Speech intelligibility	ANL	Paired comparison	Scale ratings	Self-report questionnaires
2006; Palmer et al.	N=49; HI (age 27-85) N=30; NH (mean age 34.8)	BTE: Siemens Triano	MBNR+WF	✓					✓ Annoyance	✓ Benefit; Aversiveness
2006; Yuen et al.	N=9; HI (age 39-79)	BTE: Phonak Perseo	MBNR	✓		✓				
2008; Bentler et al.	N=25; HI (age 42-79)	BTE: Starkey Axent	MBNR	✓		✓			✓ Listening comfort; Ease of listening; Sound quality	✓ Benefit
2008; Keidser et al.	N=30; HI (age 20-80)	BTE: GN ReSound Canta 780-D	MBNR	✓						✓ Satisfaction
2009; Peeters et al.	N=18;HI (age 44-89)	Open-ear BTE: Widex Inteo 9e ITC Widex Inteo IN-X	Speech enhancer			✓	✓			

Appendix A: Single-microphone Noise Reduction Studies in Adults and Children (continue)

Year; Authors	Subjects	HA model	SMNR type	Field trial	Acoustic analysis	Objective test	Subjective Tests			
						Speech intelligibility	ANL	Paired comparison	Scale ratings	Self-report questionnaires
2009; Sarampalis et al.	N=25;NH (age 18-27)	Computer-based noise reduction	Euphraim-Malah			✓ (Note: word recall and dual-task paradigm)				
2009; Turgeon et al.	N=19;NH (age 4-5)	BTE: Widex Senso Diva	MBNR			✓ (Note: discrimination of Hindi dental or retroflex stops.)				
2009; Zakis et al.	N=10; HI (age 45-82)	BTE: not specified	MBNR (variable maximum gain reduction configurations)	✓		✓			✓ Listening comfort; Ease of speech understanding Overall sound quality	✓ Preference
2010; Stelmachowicz et al.	N=16; HI (age 5-10)	BTE: Starkey Destiny 1200	Spectral subtraction		✓	✓				
2011; Kuk et al.	N=11; HI (age 30-88)	BTE: Widex mind440-19	Speech enhancer			✓				

Appendix A: Single-microphone Noise Reduction Studies in Adults and Children (continue)

Year; Authors	**Subjects**	**HA model**	**SMNR type**	**Field trial**	**Acoustic analysis**	**Objective test**	**Subjective Tests**			
						Speech intelligibility	**ANL**	**Paired comparison**	**Scale ratings**	**Self-report questionnaires**
2011; Ng et al.	N=14;NH (age 22-28) N=15;NH (age 6-18) N=14; HI (age 9-16)	BTE: not specified	MBNR+WF			✓				
2011b; Pittman	N=50; NH (age 8-12) N=30; HI (age 8-12)	BTE: Siemens Explorer 500	MBNR+WF		✓	✓ (Note: word categorization with competing complex visual task)				
2011a; Pittman	N=41; NH (age 8-9; 11-12) N=26; HI (age 8-12)	BTE: not specified	MBNR+WF		✓	✓ (Note: novel word learning)				
2012; Brons et al.	N=10, NH (age 19-23)	Four BTEs: Phonak Exelia M, ReSound Azure, Starkey Destiny 1200, Widex Mind 440	Multiple types		✓	✓		✓ Speech naturalness; Noise annoyance; Preferences	✓ Listening effort	

Appendix A: Single-microphone Noise Reduction Studies in Adults and Children (continue)

Year; Authors	**Subjects**	**HA model**	**SMNR type**	**Field trial**	**Acoustic analysis**	**Objective test**	**Subjective Tests**			
						Speech intelligibility	**ANL**	**Paired comparison**	**Scale ratings**	**Self-report questionnaires**
2012; Fredelake et al.	N=11; HI (age 13-67)	Computer-based noise reduction	Spectral subtraction				✓			
2013; Ng et al.	N=26; HI (age 32-65)	Computer-based noise reduction	Binary masking			✓ (sentence-final word identification and recall)				
2014; Gustafson et al.	N=24; NH (age 7-12)	Oticon Agil Pro Phonak naiad V SP	MBNR; Syncrony detection		✓	✓			✓ Sound clarity	
2015; Ng et al.		Computer-based noise reduction	Binary masking			✓ (sentence-final word identification and recall)				

Note: BTE = behind-the-ear; HI = hearing-impaired; ITC = in-the-canal; ITE = in-the-ear; MBNR = modulation-based noise reduction; N= number of participants; NH = normal-hearing; WF=Wiener Filtering.

Appendix B: Method and Materials for Speech-intelligibility testing in Single-microphone Noise Reduction Studies

Year; Authors	Test conditions	Test material	Speech		Noise	Loudspeaker(s) Position	Signal-to-Noise Ratio (SNR)	Results
			Type	Talker Gender	Type			
2000; Boymans & Dreschler	No NR DIR SMNR SMNR+DIR	Custom materials	Sentences	Male Female	Cocktail party noise Car noise	Speech: 0º Noise: 90º, 180º, 270º	Adaptive SNR (noise at 65 dBA)	(DIR = SMNR+DIR) > (SMNR = Omni)
2000; Walden et al.	UA No NR DIR SMNR+DIR	Connected speech test	Sentences	Female	Moderate: 2-talker babble Loud: 6-talker babble	Speech: 0º. Noise: 90º, 180º, 270º	0 dB SNR (speech and noise at 60 dBA) +2 dB SNR (speech at 75 dBA, noise at 73 dBA)	(DIR = SMNR+DIR) > (UA = Omni)
2001; Bray & Nillson	UA No NR DIR SMNR SMNR+DIR	Hearing in noise test (HINT)	Sentences	Male	Steady-state speech-shaped noise	Speech: 0º Noise-front: 0º Noise-diffuse: five loudspeakers (note: to avoid front-back or left-right advantage)	Adaptive SNR (noise at 65 dBA)	Noise front: (SMNR = SMNR+DIR) > (DIR = Omni) Noise diffuse: SMNR > Omni; SMNR+DIR > DIR > SMNR > Omni

Appendix B: Method and Materials for Speech-intelligibility testing in Single-Microphone Noise Reduction Studies (continue)

Year; Authors	Test conditions	Test material	Speech		Noise	Loudspeaker(s) Position	Signal-to-Noise Ratio (SNR)	Results
			Type/ variables	Talker Gender	Type			
2003; Alcantara et al.	Unaided SMNR-on SMNR-off	Adaptive sentence list	Sentences	n/a	Steady-state speech-shaped noise Modulated speech-shaped noise Steady-state noise with spectral dips Modulated noise with spectral dips	Speech: 0º Noise: 0º	Adaptive SNR (noise at 65 dB SPL)	SMNR-on = SMNR-off $SRT_{modulated} < SRT_{steady}$ $SRT_{dips} < SRT_{no\text{-}dips}$
2005; Mueller et al.	SMNR-on SMNR-off	HINT	Sentences	Male	Steady-state speech-shaped noise	Speech: 0º Noise: 0º	Adaptive SNR (noise at 65 dB SPL)	SMNR-on = SMNR-off
2005; Ricketts & Hornsby	No NR DIR SMNR SMNR+DIR	Connected speech test	Speech passages	Female	6-talker babble	Speech: 0º Noise: 0º, 160º, 180º, 300º	+6 dB SNR (speech at 71 dBA) +1 dB SNR (speech at 75 dBA).	(DIR = DIR+SMNR) > (SMNR = No NR)

Appendix B: Method and Materials for Speech-intelligibility testing in Single-Microphone Noise Reduction Studies (continue)

Year; Authors	**Test conditions**	**Test material**	**Speech**		**Noise**	**Loudspeaker(s) Position**	**Signal-to-Noise Ratio (SNR)**	**Results**
			Type	**Talker Gender**	**Type**			
2006; Marcoux et al.	SMNR-on SMNR-off	Custom materials	Vowel-consonant or vowel-consonant-vowel syllables (Note: Hindi dental or retroflex stops)	Female	Unmodulated International Collegium of Rehabilitative Audiology (ICRA) noise	Headphones were used (stimuli were pre-recorded from a hearing aid)	0 dB SNR (speech-plus-noise at 65 dB SPL).	SMNR-on = SMNR-off
2006; Nordrum et al.	No NR DIR SMNR SMNR+DIR	HINT	Sentences	Male	Steady-state speech-shaped noise	Speech: 0° Noise: 90°, 180°, 270°	Adaptive SNR (noise at 65 dBA)	(DIR = Dir+SMNR) > (Omni = SMNR)
2006; Yuen et al.	DIR SMNR SMNR+DIR	Cantonese HINT	Sentences	Male	Steady-state speech-shaped noise	Speech: 0° Noise-front: 0° Noise-side: 90°	Adaptive SNR Fixed SNR	Noise front: DIR = SMNR = SMNR+DIR Noise-side: (DIR = SMNR+DIR) > SMNR
2008; Bentler et al.	SMNR-off SMNR-4s SMBR-8s SMNR-16s	I-SPIN	Sentences (with and without visual cues)	Female	Multi-talker babble	Speech: 0° Noise:180°	Adaptive SNR (noise at 62 dB SPL and 78 dB SPL)	At 78 dB SPL: SMNR-off = SNMR-on At 62 dB SPL: (SMNR-off = SMNR-4s) > (SMNR-8s = SMNR-16s)

Appendix B: Method and Materials for Speech-intelligibility testing in Single-Microphone Noise Reduction Studies (continue)

Year; Authors	**Test conditions**	**Test material**	**Speech**		**Noise Type**	**Loudspeaker(s) Position**	**Signal-to-Noise Ratio (SNR)**	**Results**
			Type/ variables	**Talker Gender**				
2009; Peeters et al.	No NR DIR SMNR SMNR+DIR	HINT	Sentences	Male	Steady-state speech-shaped noise	Speech: 0^o Noise: 90^o, 180^o, 270^o	Adaptive SNR (noise at 75 dB SPL)	(DIR = SMNR+DIR) > Omni (DIR = SMNR+DIR) > SMNR SMNR > Omni.
2009; Sarampalis et al.	EM No EM	Experiment I: SPIN-R Experiment II: Institute of Electrical and Electronics Engineers (IEEE) sentences	Experiment I: Words Experiment II: Sentences	Male	Experiment I: Babble noise Experiment II: 4-talker babble (fixed at 65dB SPL)	Headphones were used.	Experiment I: +2 dB SNR -2 dB SNR Experiment II: -6 dB SNR -2 dB SNR +2 dB SNR	Exp I: % word identified (EM = no EM at +2dB SNR); word recall (EM > no EM at -2dB SNR). Exp II: Speech intelligibility (EM = no EM), visual task reaction time (EM ≥ no EM)
2009; Turgeon et al.	SMNR-on SMNR-off	Custom materials	VCV syllables containing Hindi dental or retroflex stops	Female	Unmodulated ICRA noise	Headphones were used (stimuli were pre-recorded from a hearing aid)	0 dB SNR + 5 dB SNR	SMNR-on = SMNR-off

Appendix B: Method and Materials for Speech-intelligibility testing in Single-Microphone Noise Reduction Studies (continue)

Year; Authors	**Test conditions**	**Test material**	**Speech**		**Noise**	**Loudspeaker(s) Position**	**Signal-to-Noise Ratio (SNR)**	**Results**
			Type/ variables	**Talker Gender**	**Type**			
2009; Zakis et al.	SMNR-off SMNR-on (constant gain reduction) SMNR-on (variable gain reduction)	HINT	Sentences	Male	Speech-shaped noise 8-talker babble	Speech: 0^{o} Noise: 180^{o}	Adaptive SNR (noise at 65 dBA)	All configurations: SMNR-off = SMNR-on
2011; Kuk et al.	SMNR-off SMNR-on (classic) SMNR-on (speech enhancer)	HINT	Sentences	Male	Speech-shaped noise	Speech: 0^{o} Noise:0^{o}	Adaptive SNR (noise at 68 dB SPL and 75 dB SPL)	SMNR-on (classic or speech enhancer) > SMNR-off
2010; Stelmachowicz et al.	SMNR-on SMNR-off	Phonetically Balanced Kindergarten lists Bamford-Kowal-Bench lists	Vowel-consonant-vowel nonsense syllables Monosyllabic words Sentences	Female	Speech-shaped noise	Speech: 0^{o} Noise:0^{o}	0 dB SNR +5 dB SNR +10 dB SNR (speech at 65 dB SPL).	Significant main effects of age and stimulus type. SMNR-on = SMNR-off
2011b; Pittman	SMNR-on SMNR-off	Custom materials	Words	Female	Steady-state noise	Speech: 0^{o} Noise:0^{o}	0 dB SNR (overall level: 50 dB SPL)	SMNR-on = SMNR-off.
2011a; Pittman	Quiet SMNR-on SMNR-off	Custom materials	Nonsense words (two-syllables)	Female	Steady-state noise	Speech: 0^{o} Noise:0^{o}	0 dB SNR (overall level: 50 dB SPL)	8 to 9-year-old: SMNR-on = SMNR-off 11 to 12-year-old: SMNR-on > SMNR-off

Appendix B: Method and Materials for Speech-intelligibility testing in Single-Microphone Noise Reduction Studies (continue)

Year; Authors	Test conditions	Test material	Speech		Noise	Loudspeaker(s) Position	Signal-to-Noise Ratio (SNR)	Results
			Type/ variables	Talker Gender	Type			
2012; Brons et al.	SMNR-on, SMNR-off	Custom materials	Sentences	Female	Multitalker babble	Headphones were used. (Note: stimuli were pre-recorded from four hearing aids)	-7 dB SNR -4 dB SNR (noise at 70 dBA; speech at 63 dBA and 66 dBA, respectively)	SMNR-on = SMNR-off
2013; Ng et al.	No NR SMNR-on SMNR-on (ideal)	Swedish HINT	Sentences (final-word identification & recall)	Male	Steady-state noise 4-talker babble (Swedish)	Insert earphones were used (stimuli pre-processed using a computer and a BTE)	Individualize SNR (SNR at which 95% intelligibility was achieved to ensure audibility)	SMNR-on improved word-recall performance for subjects with high working memory
2015; Ng et al.	No NR SMNR-on	Swedish HINT	Sentences (final-word identification & recall)	Male	4-talker babble (Swedish) 4-talker babble (Cantonese)	Insert earphones were used (stimuli pre-processed using a computer and a BTE)	Individualize SNR (SNR at which 95% intelligibility was achieved to ensure audibility)	SMNR-on improved word-recall performance only when competing noise was native.

Note: ">" = significantly better performance; "<" = significantly poorer performance; "=" = no significant difference; BTE = behind-the-ear; DIR = directional microphones; EM = Euphraim-Malah; NR = noise reduction; SMNR = single-microphone noise reduction; SRT = speech reception thresholds.

Appendix C: List of Speech-only Files

Phoneme	Name of speech-only sound files	
	Female talker	Male Talker
/s/ Mandarin alveolar fricative	s_Mfemale	s_Mmale
/ts/ Mandarin alveolar unaspirated affricate	z_Mfemale	z_Mmale
/ts^h/ Mandarin alveolar aspirated affricate	c_Mfemale	c_Mmale
/ʂ/ Mandarin retroflex fricative	sh_Mfemale	sh_Mmale
/tʂ/ Mandarin retroflex unaspirated affricate	zh_Mfemale	zh_Mmale
/tʂʰ/ Mandarin retroflex aspirated affricate	ch_Mfemale	ch_Mmale
/s/ English voiceless alveolar fricative	s_Efemale	s_Emale
z/ English voiced alveolar fricative	z_Efemale	z_Emale
/ʃ/ English voiceless palatal fricative	sh_Efemale	sh_Emale
/tʃ / English voiceless post-alveolar affricate	ch_Efemale	ch_Emale
/ʤ/ English voiced post-alveolar affricate	j_Efemale	j_Emale

Appendix D: Sets of Speech-plus-noise Files

1. Female Native Mandarin talker

Phoneme	Original speech-plus-Original Noise Files	Original speech-plus-Inverted Noise Files
/s/ Mandarin alveolar fricative	s_Mfemale_ori_Noise_ori-10dB	s_Mfemale_ori_Noise_inv-10dB
	s_Mfemale_ori_Noise_ori-5dB	s_Mfemale_ori_Noise_inv-5dB
	s_Mfemale_ori_Noise_ori0dB	s_Mfemale_ori_Noise_inv0dB
	s_Mfemale_ori_Noise_ori_+5dB	s_Mfemale_ori_Noise_inv+5dB
	s_Mfemale_ori_Noise_ori+10dB	s_Mfemale_ori_Noise_inv+10dB
/ts/ Mandarin alveolar unaspirated affricate	z_Mfemale_ori_Noise_ori-10dB	z_Mfemale_ori_Noise_inv-10dB
	z_Mfemale_ori_Noise_ori-5dB	z_Mfemale_ori_Noise_inv-5dB
	z_Mfemale_ori_Noise_ori0dB	z_Mfemale_ori_Noise_inv0dB
	z_Mfemale_ori_Noise_ori_+5dB	z_Mfemale_ori_Noise_inv+5dB
	z_Mfemale_ori_Noise_ori+10dB	z_Mfemale_ori_Noise_inv+10dB
/ts^h/ Mandarin alveolar aspirated affricate	c_Mfemale_ori_Noise_ori-10dB	c_Mfemale_ori_Noise_inv-10dB
	c_Mfemale_ori_Noise_ori-5dB	c_Mfemale_ori_Noise_inv-5dB
	c_Mfemale_ori_Noise_ori0dB	c_Mfemale_ori_Noise_inv0dB
	c_Mfemale_ori_Noise_ori_+5dB	c_Mfemale_ori_Noise_inv+5dB
	c_Mfemale_ori_Noise_ori+10dB	c_Mfemale_ori_Noise_inv+10dB
/ʂ/ Mandarin retroflex fricative	sh_Mfemale_ori_Noise_ori-10dB	sh_Mfemale_ori_Noise_inv-10dB
	sh_Mfemale_ori_Noise_ori-5dB	sh_Mfemale_ori_Noise_inv-5dB
	sh_Mfemale_ori_Noise_ori0dB	sh_Mfemale_ori_Noise_inv0dB
	sh_Mfemale_ori_Noise_ori_+5dB	sh_Mfemale_ori_Noise_inv+5dB
	sh_Mfemale_ori_Noise_ori+10dB	sh_Mfemale_ori_Noise_inv+10dB
/tʂ/ Mandarin retroflex unaspirated affricate	zh_Mfemale_ori_Noise_ori-10dB	zh_Mfemale_ori_Noise_inv-10dB
	zh_Mfemale_ori_Noise_ori-5dB	zh_Mfemale_ori_Noise_inv-5dB
	zh_Mfemale_ori_Noise_ori0dB	zh_Mfemale_ori_Noise_inv0dB
	zh_Mfemale_ori_Noise_ori_+5dB	zh_Mfemale_ori_Noise_inv+5dB
	zh_Mfemale_ori_Noise_ori+10dB	zh_Mfemale_ori_Noise_inv+10dB
/tʂ^h/ Mandarin retroflex aspirated affricate	ch_Mfemale_ori_Noise_ori-10dB	ch_Mfemale_ori_Noise_inv-10dB
	ch_Mfemale_ori_Noise_ori-5dB	ch_Mfemale_ori_Noise_inv-5dB
	ch_Mfemale_ori_Noise_ori0dB	ch_Mfemale_ori_Noise_inv0dB
	ch_Mfemale_ori_Noise_ori_+5dB	ch_Mfemale_ori_Noise_inv+5dB
	ch_Mfemale_ori_Noise_ori+10dB	ch_Mfemale_ori_Noise_inv+10dB

Appendix D: Sets of Speech-plus-noise Files (continue)

2. Male Native Mandarin talker

Phoneme	Original speech-plus-Original Noise Files	Original speech-plus-Inverted Noise Files
/s/ Mandarin alveolar fricative	s_Mmale_ori_Noise_ori-10dB	s_Mmale_ori_Noise_inv-10dB
	s_Mmale_ori_Noise_ori-5dB	s_Mmale_ori_Noise_inv-5dB
	s_Mmale_ori_Noise_ori0dB	s_Mmale_ori_Noise_inv0dB
	s_Mmale_ori_Noise_ori_+5dB	s_Mmale_ori_Noise_inv+5dB
	s_Mmale_ori_Noise_ori+10dB	s_Mmale_ori_Noise_inv+10dB
/ts/ Mandarin alveolar unaspirated affricate	z_Mmale_ori_Noise_ori-10dB	z_Mmale_ori_Noise_inv-10dB
	z_Mmale_ori_Noise_ori-5dB	z_Mmale_ori_Noise_inv-5dB
	z_Mmale_ori_Noise_ori0dB	z_Mmale_ori_Noise_inv0dB
	z_Mmale_ori_Noise_ori_+5dB	z_Mmale_ori_Noise_inv+5dB
	z_Mmale_ori_Noise_ori+10dB	z_Mmale_ori_Noise_inv+10dB
/ts^h/ Mandarin alveolar aspirated affricate	c_Mmale_ori_Noise_ori-10dB	c_Mmale_ori_Noise_inv-10dB
	c_Mmale_ori_Noise_ori-5dB	c_Mmale_ori_Noise_inv-5dB
	c_Mmale_ori_Noise_ori0dB	c_Mmale_ori_Noise_inv0dB
	c_Mmale_ori_Noise_ori_+5dB	c_Mmale_ori_Noise_inv+5dB
	c_Mmale_ori_Noise_ori+10dB	c_Mmale_ori_Noise_inv+10dB
/ʂ/ Mandarin retroflex fricative	sh_Mmale_ori_Noise_ori-10dB	sh_Mmale_ori_Noise_inv-10dB
	sh_Mmale_ori_Noise_ori-5dB	sh_Mmale_ori_Noise_inv-5dB
	sh_Mmale_ori_Noise_ori0dB	sh_Mmale_ori_Noise_inv0dB
	sh_Mmale_ori_Noise_ori_+5dB	sh_Mmale_ori_Noise_inv+5dB
	sh_Mmale_ori_Noise_ori+10dB	sh_Mmale_ori_Noise_inv+10dB
/tʂ/ Mandarin retroflex unaspirated affricate	zh_Mmale_ori_Noise_ori-10dB	zh_Mmale_ori_Noise_inv-10dB
	zh_Mmale_ori_Noise_ori-5dB	zh_Mmale_ori_Noise_inv-5dB
	zh_Mmale_ori_Noise_ori0dB	zh_Mmale_ori_Noise_inv0dB
	zh_Mmale_ori_Noise_ori_+5dB	zh_Mmale_ori_Noise_inv+5dB
	zh_Mmale_ori_Noise_ori+10dB	zh_Mmale_ori_Noise_inv+10dB
/$tʂ^h$/ Mandarin retroflex aspirated affricate	ch_Mmale_ori_Noise_ori-10dB	ch_Mmale_ori_Noise_inv-10dB
	ch_Mmale_ori_Noise_ori-5dB	ch_Mmale_ori_Noise_inv-5dB
	ch_Mmale_ori_Noise_ori0dB	ch_Mmale_ori_Noise_inv0dB
	ch_Mmale_ori_Noise_ori_+5dB	ch_Mmale_ori_Noise_inv+5dB
	ch_Mmale_ori_Noise_ori+10dB	ch_Mmale_ori_Noise_inv+10dB

Appendix D: Sets of Speech-plus-noise Files (continue)

3. Female native English talker

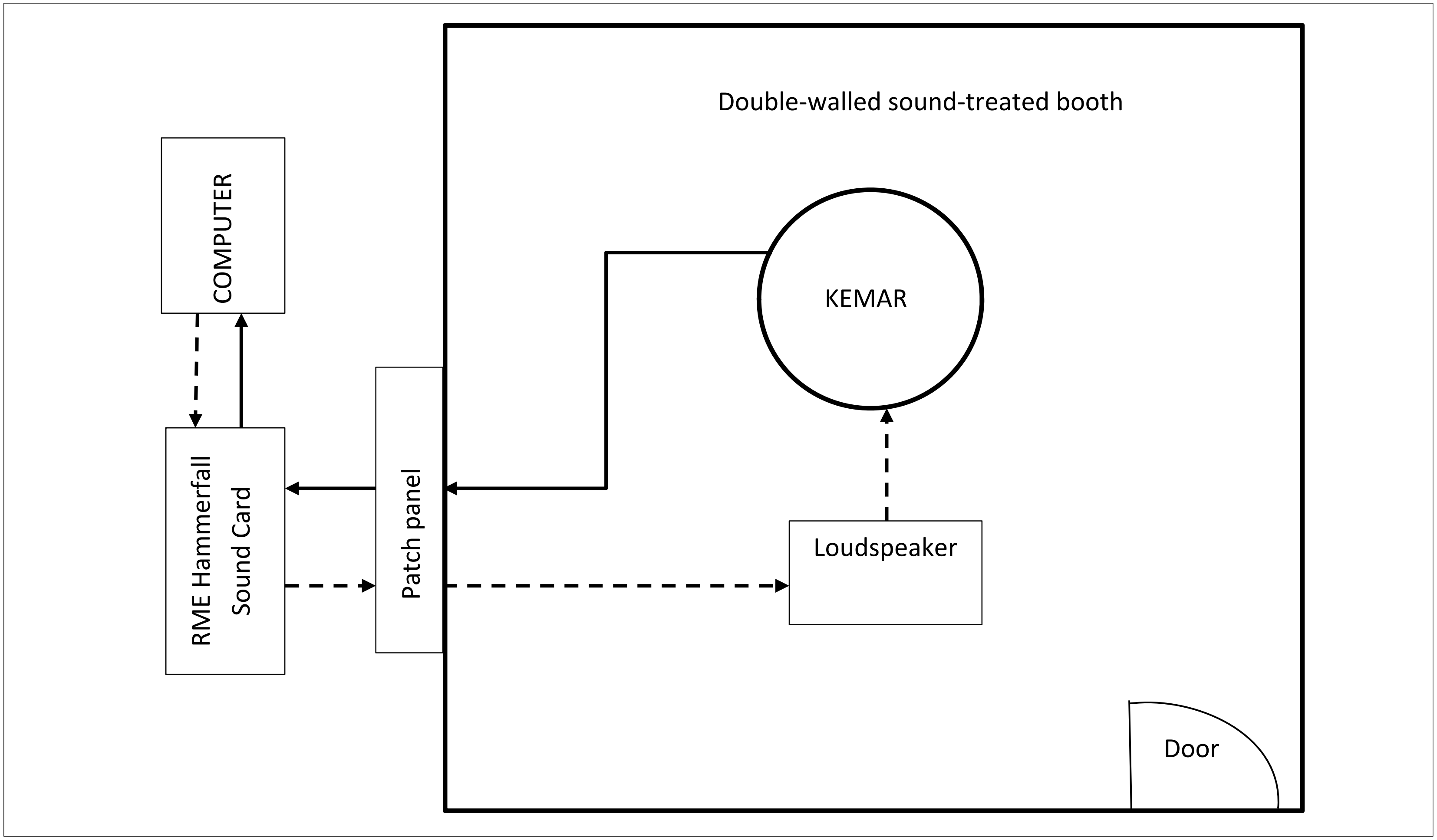
Phoneme	Original speech-plus-Original Noise Files	Original speech-plus-Inverted Noise Files
/s/	s_Efemale_ori_Noise_ori-10dB	s_Efemale_ori_Noise_inv-10dB
English	s_Efemale_ori_Noise_ori-5dB	s_Efemale_ori_Noise_inv-5dB
voiceless	s_Efemale_ori_Noise_ori0dB	s_Efemale_ori_Noise_inv0dB
alveolar	s_Efemale_ori_Noise_ori_+5dB	s_Efemale_ori_Noise_inv+5dB
fricative	s_Efemale_ori_Noise_ori+10dB	s_Efemale_ori_Noise_inv+10dB
/z/	z_Efemale_ori_Noise_ori-10dB	z_Efemale_ori_Noise_inv-10dB
English	z_Efemale_ori_Noise_ori-5dB	z_Efemale_ori_Noise_inv-5dB
voiced	z_Efemale_ori_Noise_ori0dB	z_Efemale_ori_Noise_inv0dB
alveolar	z_Efemale_ori_Noise_ori_+5dB	z_Efemale_ori_Noise_inv+5dB
fricative	z_Efemale_ori_Noise_ori+10dB	z_Efemale_ori_Noise_inv+10dB
/ʃ/	sh_Efemale_ori_Noise_ori-10dB	sh_Efemale_ori_Noise_inv-10dB
English	sh_Efemale_ori_Noise_ori-5dB	sh_Efemale_ori_Noise_inv-5dB
voiceless	sh_Efemale_ori_Noise_ori0dB	sh_Efemale_ori_Noise_inv0dB
palatal	sh_Efemale_ori_Noise_ori_+5dB	sh_Efemale_ori_Noise_inv+5dB
fricative	sh_Efemale_ori_Noise_ori+10dB	sh_Efemale_ori_Noise_inv+10dB
/tʃ /	ch_Efemale_ori_Noise_ori-10dB	ch_Efemale_ori_Noise_inv-10dB
English	ch_Efemale_ori_Noise_ori-5dB	ch_Efemale_ori_Noise_inv-5dB
voiceless	ch_Efemale_ori_Noise_ori0dB	ch_Efemale_ori_Noise_inv0dB
post-alveolar	ch_Efemale_ori_Noise_ori_+5dB	ch_Efemale_ori_Noise_inv+5dB
affricate	ch_Efemale_ori_Noise_ori+10dB	ch_Efemale_ori_Noise_inv+10dB
/ʤ/	zh_Efemale_ori_Noise_ori-10dB	zh_Efemale_ori_Noise_inv-10dB
English	zh_Efemale_ori_Noise_ori-5dB	zh_Efemale_ori_Noise_inv-5dB
voiced	zh_Efemale_ori_Noise_ori0dB	zh_Efemale_ori_Noise_inv0dB
post-alveolar	zh_Efemale_ori_Noise_ori_+5dB	zh_Efemale_ori_Noise_inv+5dB
affricate	zh_Efemale_ori_Noise_ori+10dB	zh_Efemale_ori_Noise_inv+10dB

Appendix D: Sets of Speech-plus-noise Files (continue)

4. Male native English talker

Phoneme	Original speech-plus-Original Noise Files	Original speech-plus-Inverted Noise Files
/s/ English voiceless alveolar fricative	s_Emale_ori_Noise_ori-10dB	s_Emale_ori_Noise_inv-10dB
	s_Emale_ori_Noise_ori-5dB	s_Emale_ori_Noise_inv-5dB
	s_Emale_ori_Noise_ori0dB	s_Emale_ori_Noise_inv0dB
	s_Emale_ori_Noise_ori_+5dB	s_Emale_ori_Noise_inv+5dB
	s_Emale_ori_Noise_ori+10dB	s_Emale_ori_Noise_inv+10dB
/z/ English voiced alveolar fricative	z_Emale_ori_Noise_ori-10dB	z_Emale_ori_Noise_inv-10dB
	z_Emale_ori_Noise_ori-5dB	z_Emale_ori_Noise_inv-5dB
	z_Emale_ori_Noise_ori0dB	z_Emale_ori_Noise_inv0dB
	z_Emale_ori_Noise_ori_+5dB	z_Emale_ori_Noise_inv+5dB
	z_Emale_ori_Noise_ori+10dB	z_Emale_ori_Noise_inv+10dB
/ʃ/ English voiceless palatal fricative	sh_Emale_ori_Noise_ori-10dB	sh_Emale_ori_Noise_inv-10dB
	sh_Emale_ori_Noise_ori-5dB	sh_Emale_ori_Noise_inv-5dB
	sh_Emale_ori_Noise_ori0dB	sh_Emale_ori_Noise_inv0dB
	sh_Emale_ori_Noise_ori_+5dB	sh_Emale_ori_Noise_inv+5dB
	sh_Emale_ori_Noise_ori+10dB	sh_Emale_ori_Noise_inv+10dB
/tʃ / English voiceless post-alveolar affricate	ch_Emale_ori_Noise_ori-10dB	ch_Emale_ori_Noise_inv-10dB
	ch_Emale_ori_Noise_ori-5dB	ch_Emale_ori_Noise_inv-5dB
	ch_Emale_ori_Noise_ori0dB	ch_Emale_ori_Noise_inv0dB
	ch_Emale_ori_Noise_ori_+5dB	ch_Emale_ori_Noise_inv+5dB
	ch_Emale_ori_Noise_ori+10dB	ch_Emale_ori_Noise_inv+10dB
/dʒ/ English voiced post-alveolar affricate	zh_Emale_ori_Noise_ori-10dB	zh_Emale_ori_Noise_inv-10dB
	zh_Emale_ori_Noise_ori-5dB	zh_Emale_ori_Noise_inv-5dB
	zh_Emale_ori_Noise_ori0dB	zh_Emale_ori_Noise_inv0dB
	zh_Emale_ori_Noise_ori_+5dB	zh_Emale_ori_Noise_inv+5dB
	zh_Emale_ori_Noise_ori+10dB	zh_Emale_ori_Noise_inv+10dB

Appendix E: Diagram of Equipment Set-up for Sound Field Recording



Note. The arrows with dash lines represent stimulus presentation; the arrows with solid lines represent recordings of the aided (HA#1 or HA#2) and unaided (KEMAR only) speech-plus-noise signals.

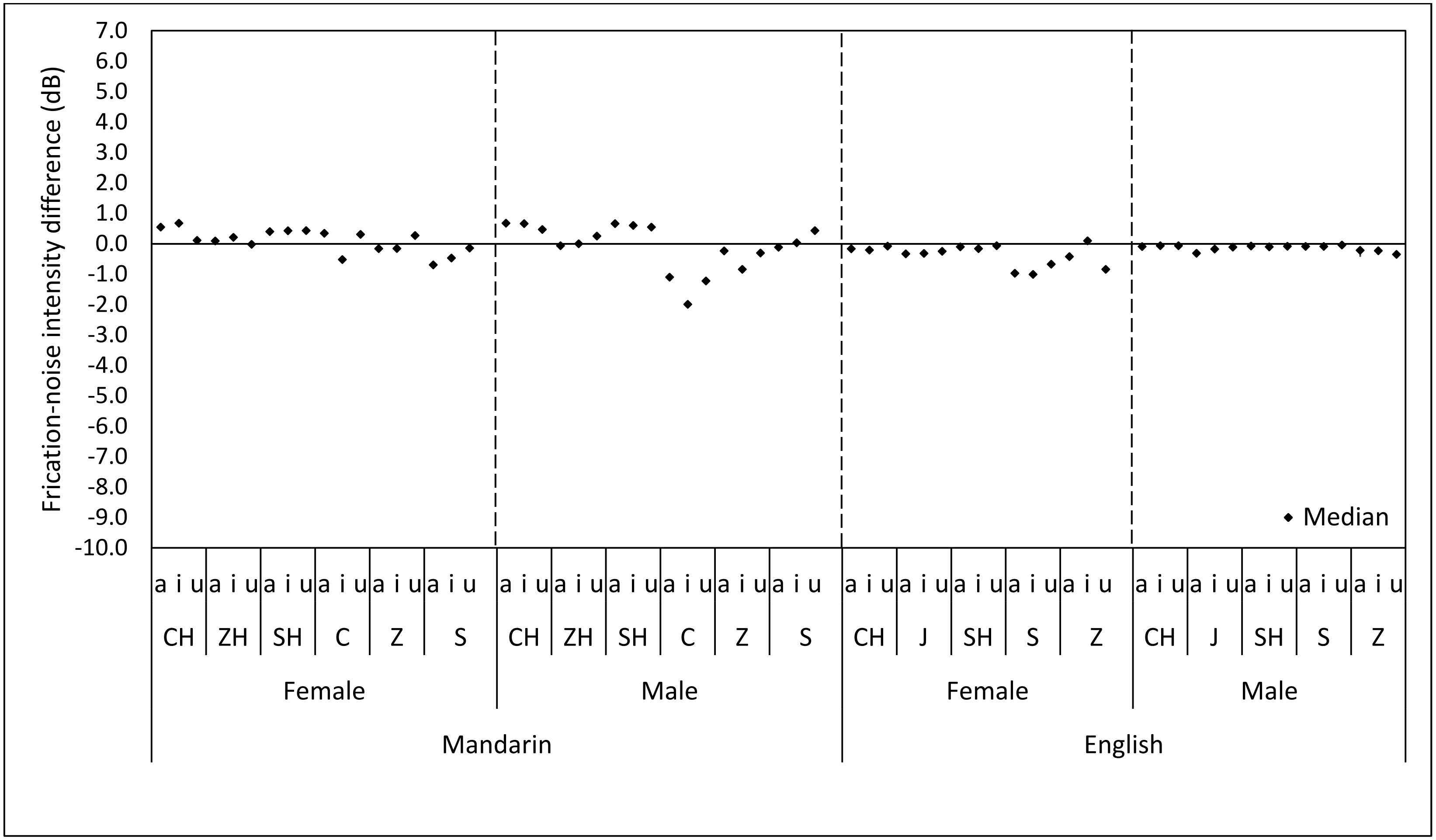
Appendix F: Settings in Praat Program

1. Spectrogram setting
    a. View range (Hz): 0 to 10000 z
    b. Window length (s): 0.005
    c. Dynamic range (dB): 70.0

2. Advanced spectrogram setting
    a. Time and frequency resolutions:
        i. Number of time steps: 1000
        ii. Number of frequency steps: 250
    b. Spectrogram analysis setting:
        i. Method: Fourier
        ii. Window Shape: Gaussian
    c. Spectrogram view setting:
        i. Autoscalling: ticked
        ii. Maximum (dB/Hz): 100
        iii. Pre-emphasis (dB/oct): 6.0
        iv. dynamic compression (0-1): 0.0

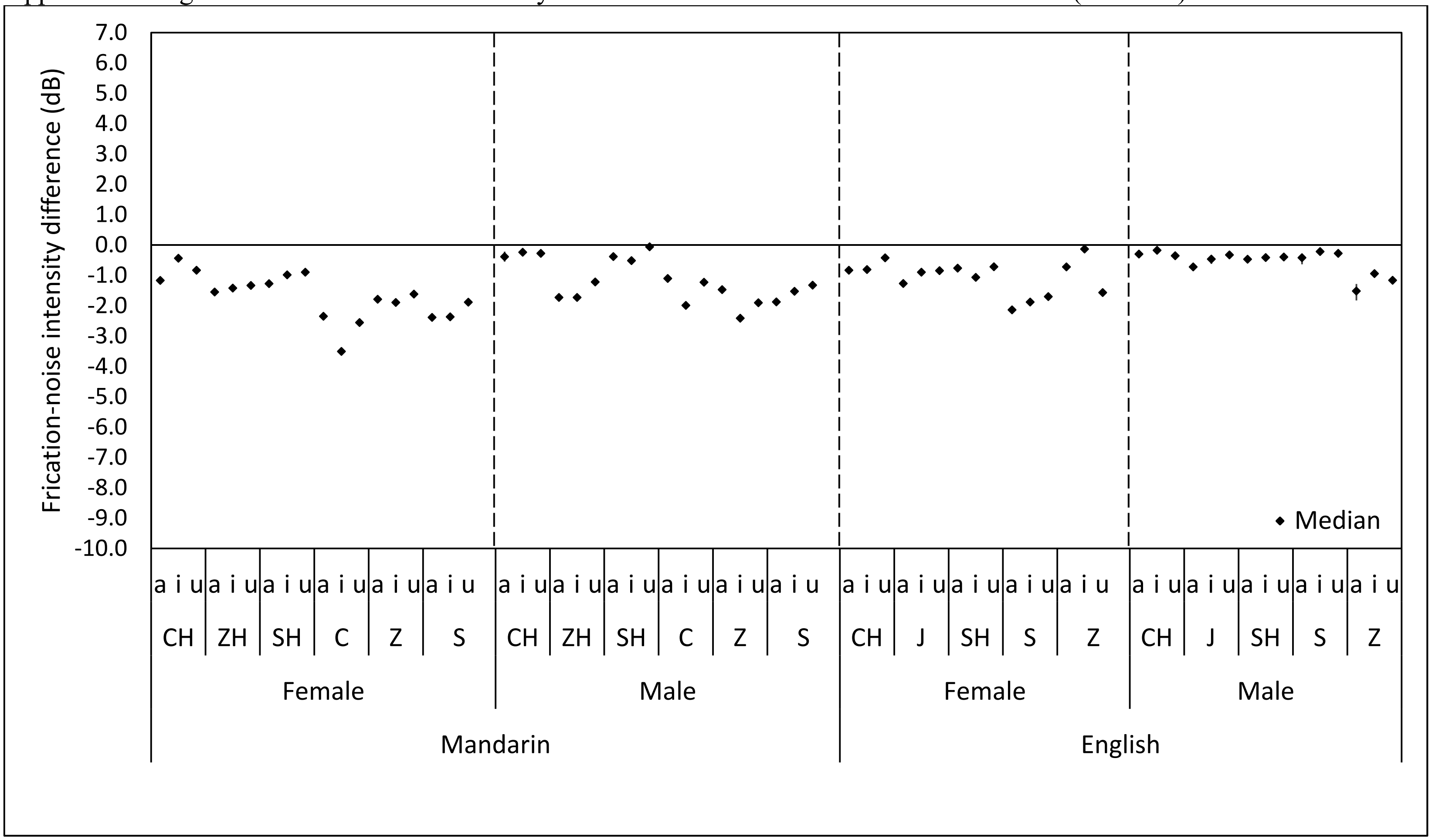
3. Intensity setting
    a. View range (dB): 0 to 100
    b. Averaging method: mean energy

Appendix G: Figures of Frication-noise Intensity Difference for HA#1 at Five SNR Conditions



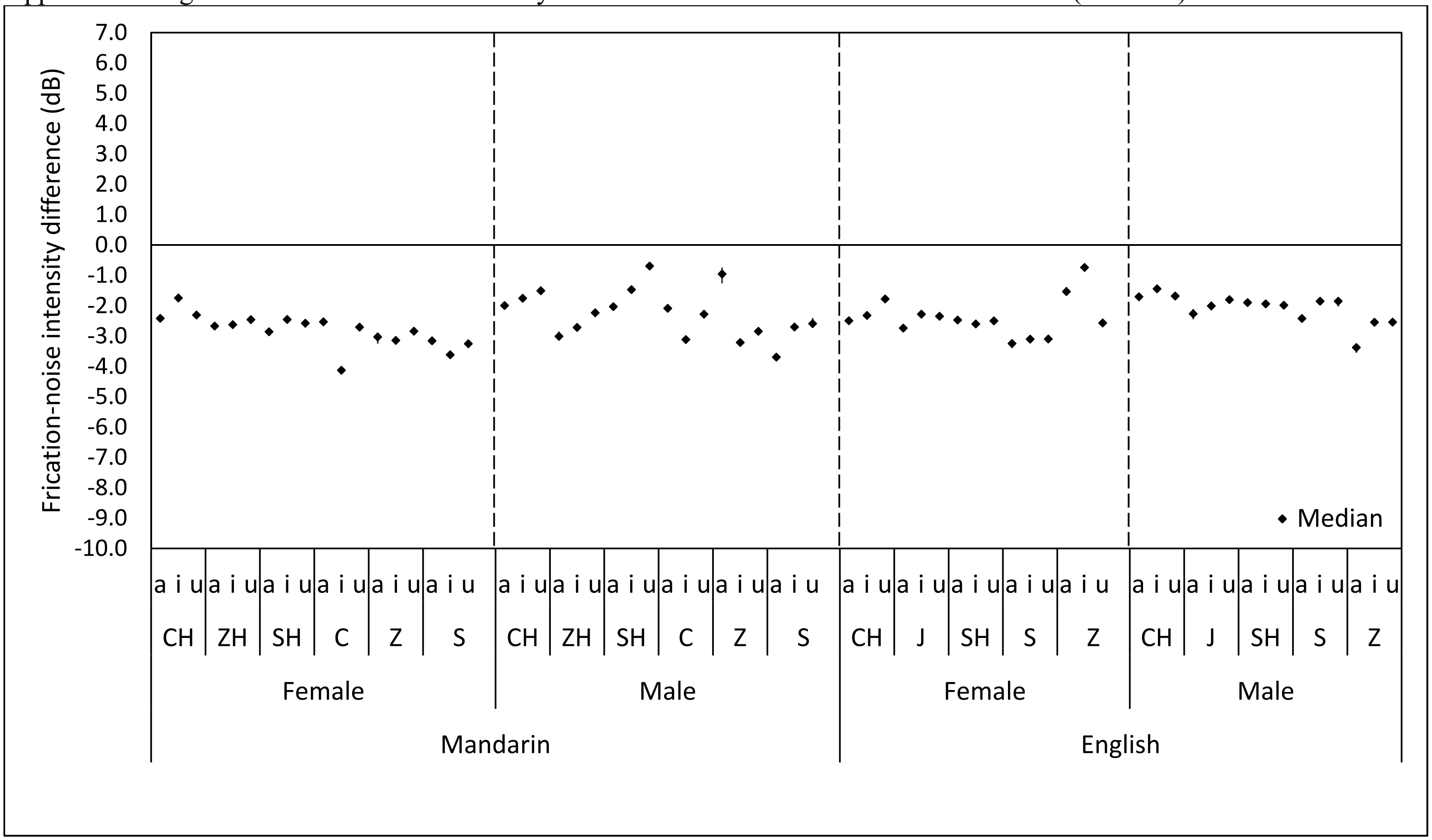
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#1 at +10 dB SNR condition. Positive values indicate intensity increment and negative values indicate frication-noise intensity reduction when SMNR was turned on. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix G: Figures of Frication-noise Intensity Difference for HA#1 at Five SNR Conditions (continue)



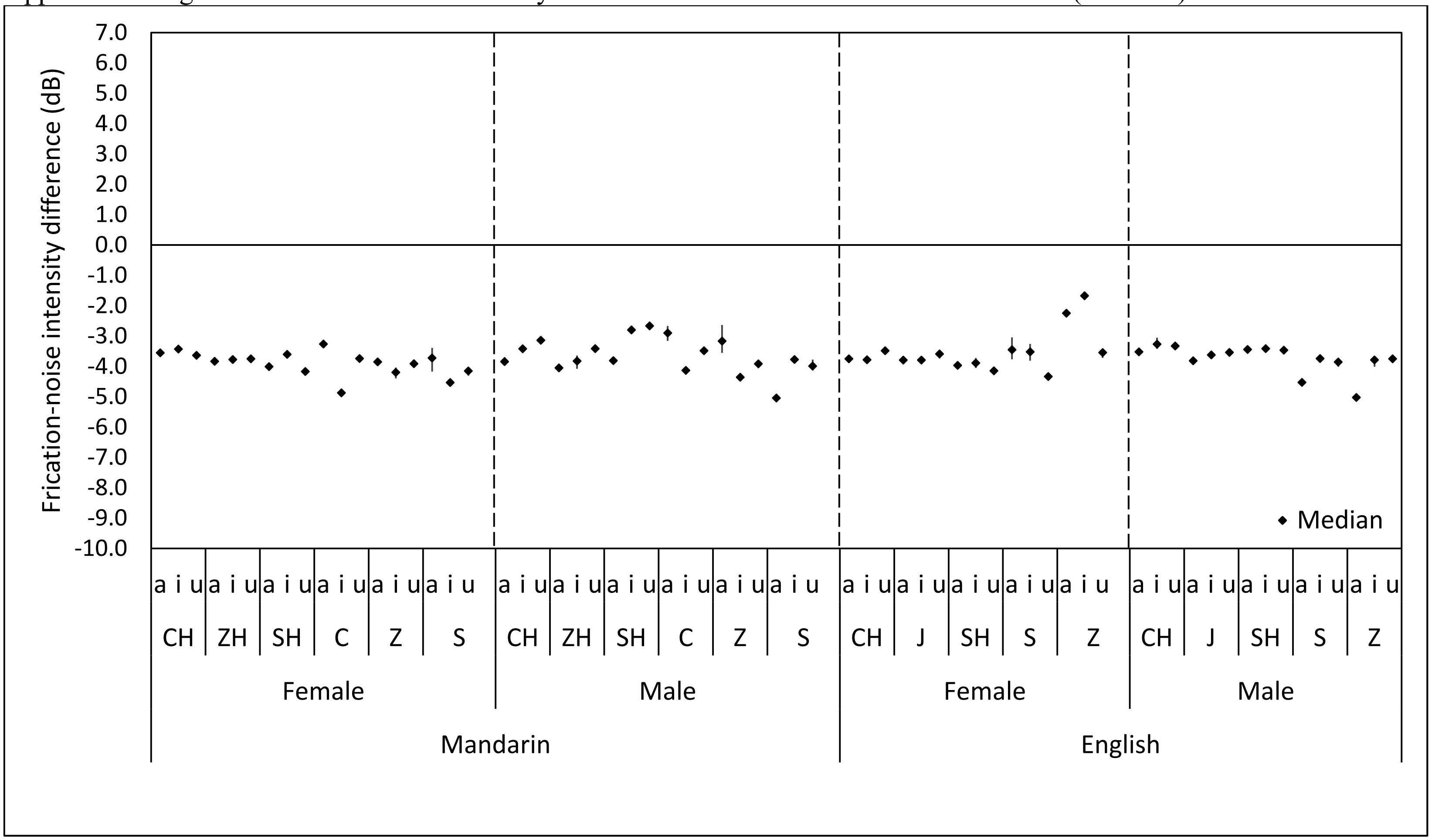
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#1 at +5 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix G: Figures of Frication-noise Intensity Difference for HA#1 at Five SNR Conditions (continue)



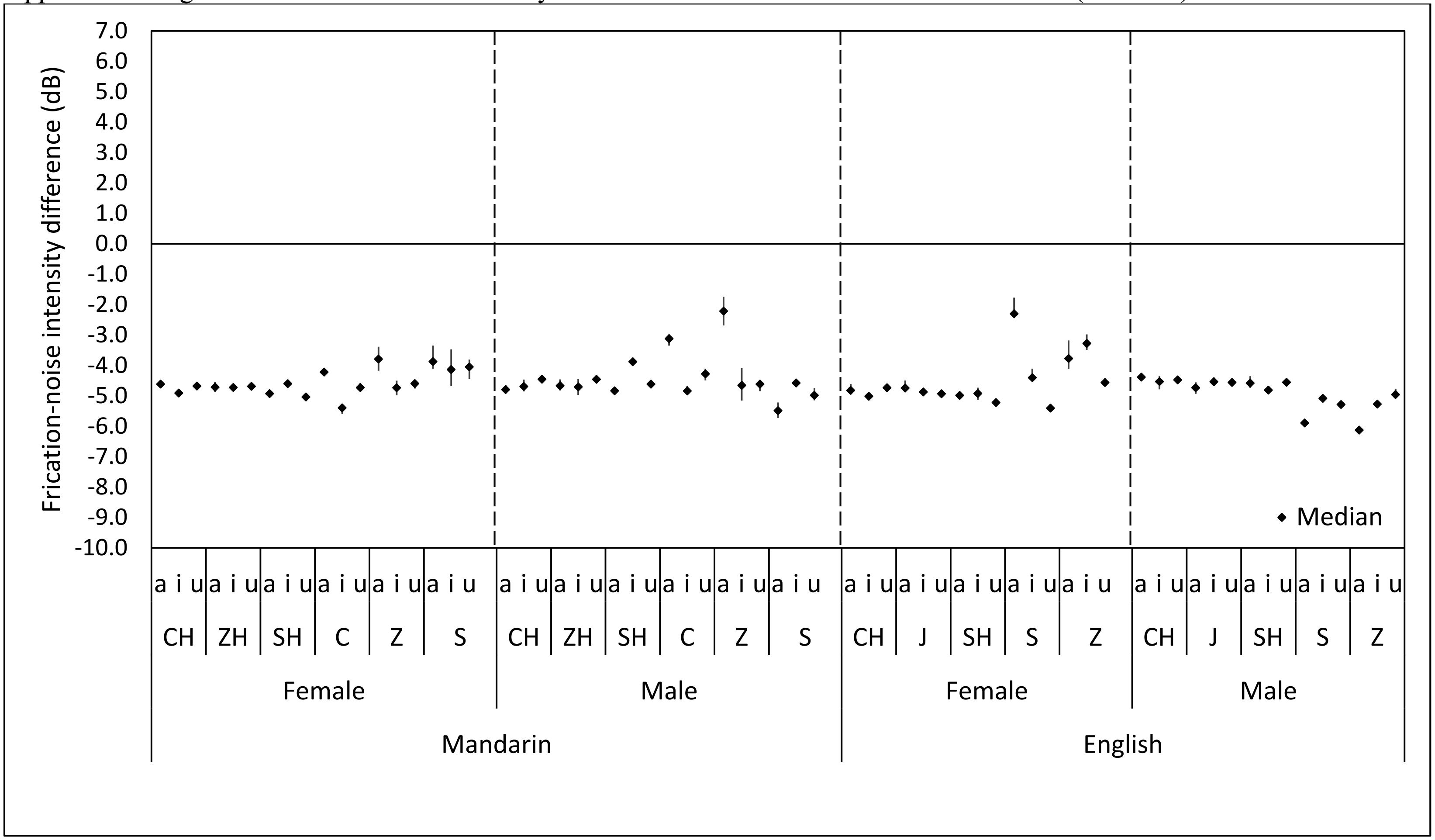
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#1 at 0 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix G: Figures of Frication-noise Intensity Difference for HA#1 at Five SNR Conditions (continue)



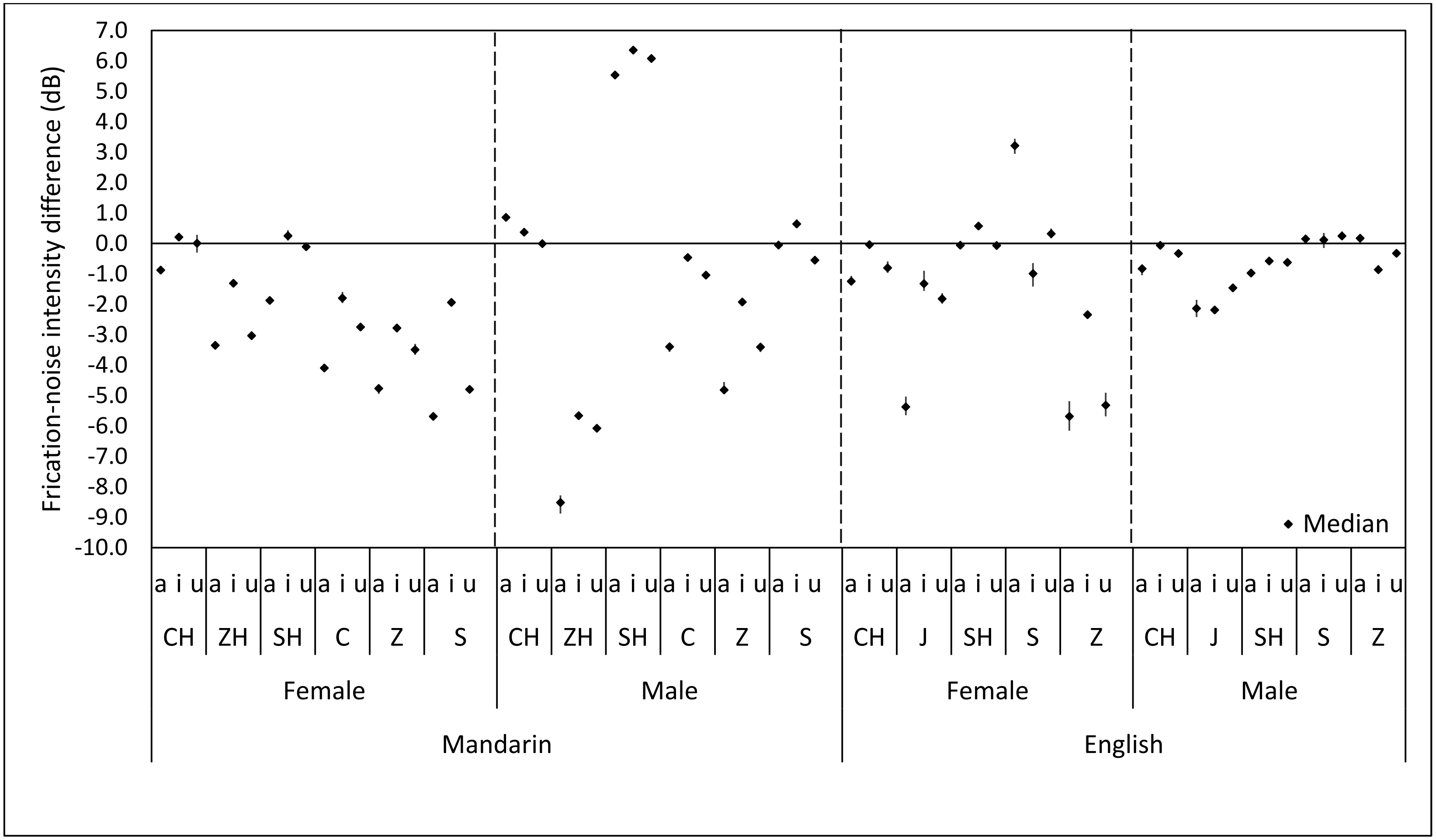
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#1 at -5 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix G: Figures of Frication-noise Intensity Difference for HA#1 at Five SNR Conditions (continue)



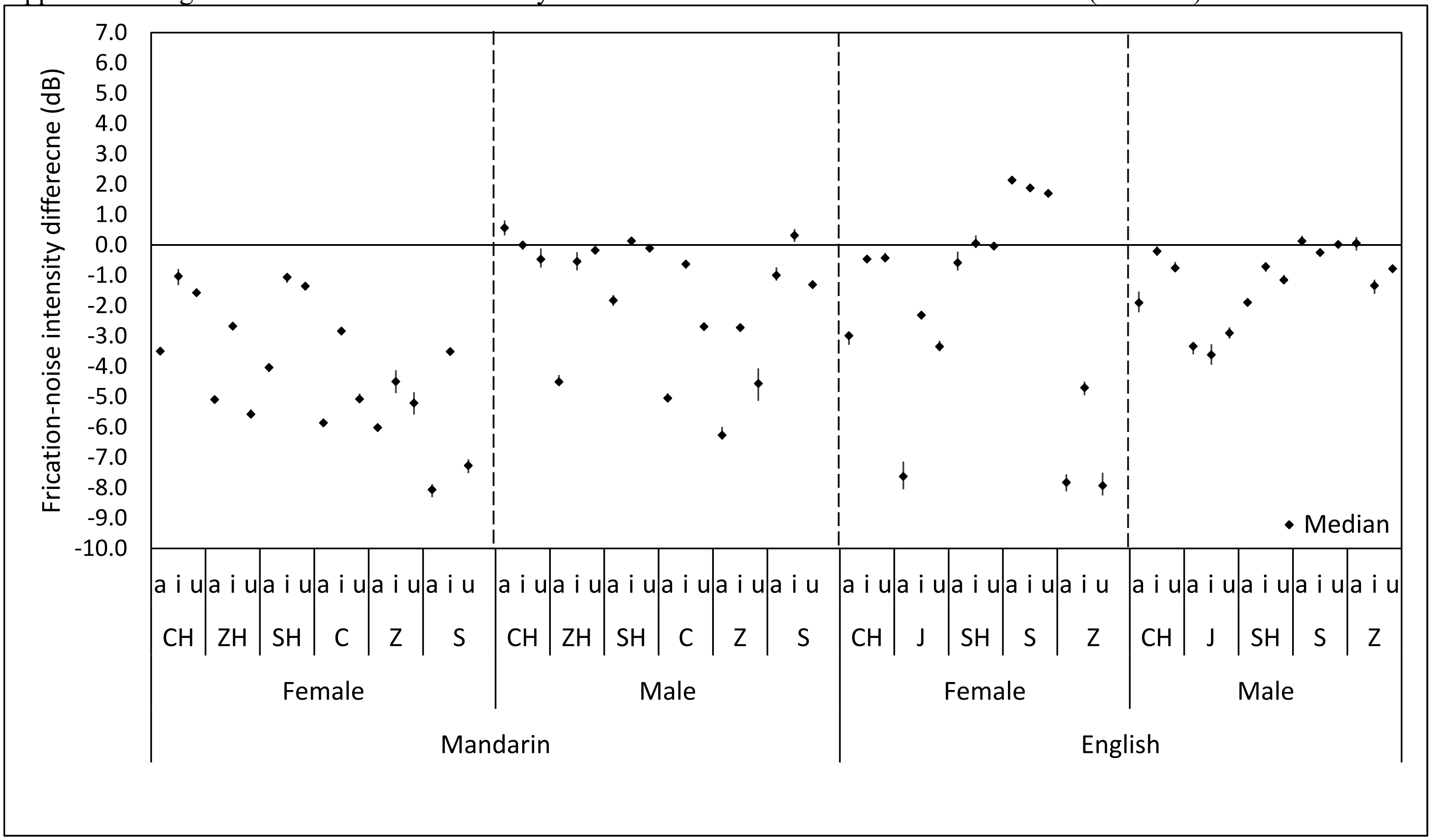
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#1 at -10 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix H: Figures of Frication-noise Intensity Difference for HA#2 at Three SNR Conditions



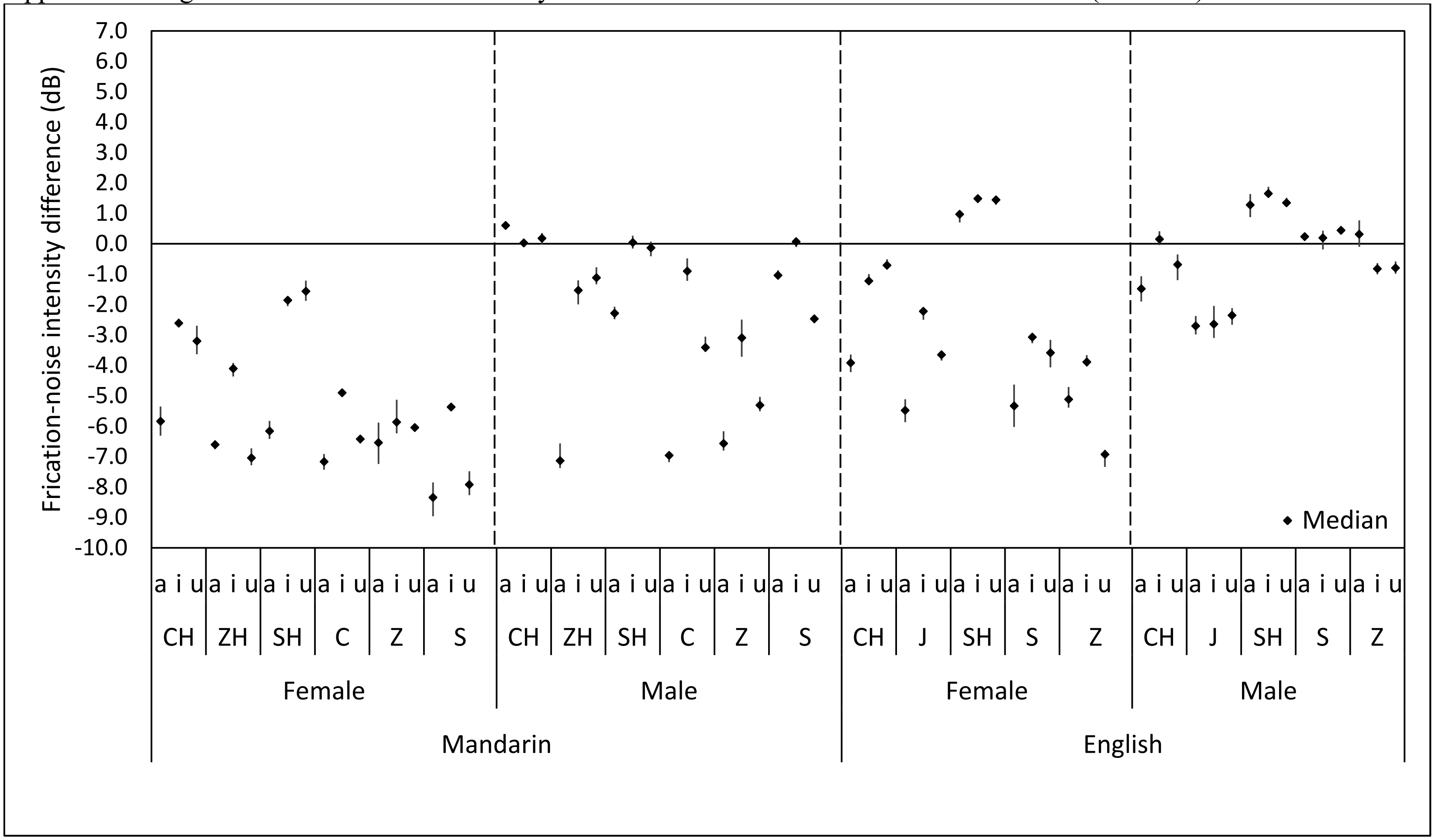
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#2 at +10 dB SNR condition. Positive values indicate intensity increment and negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix H: Figures of Frication-noise Intensity Difference for HA#2 at Three SNR Conditions (continue)



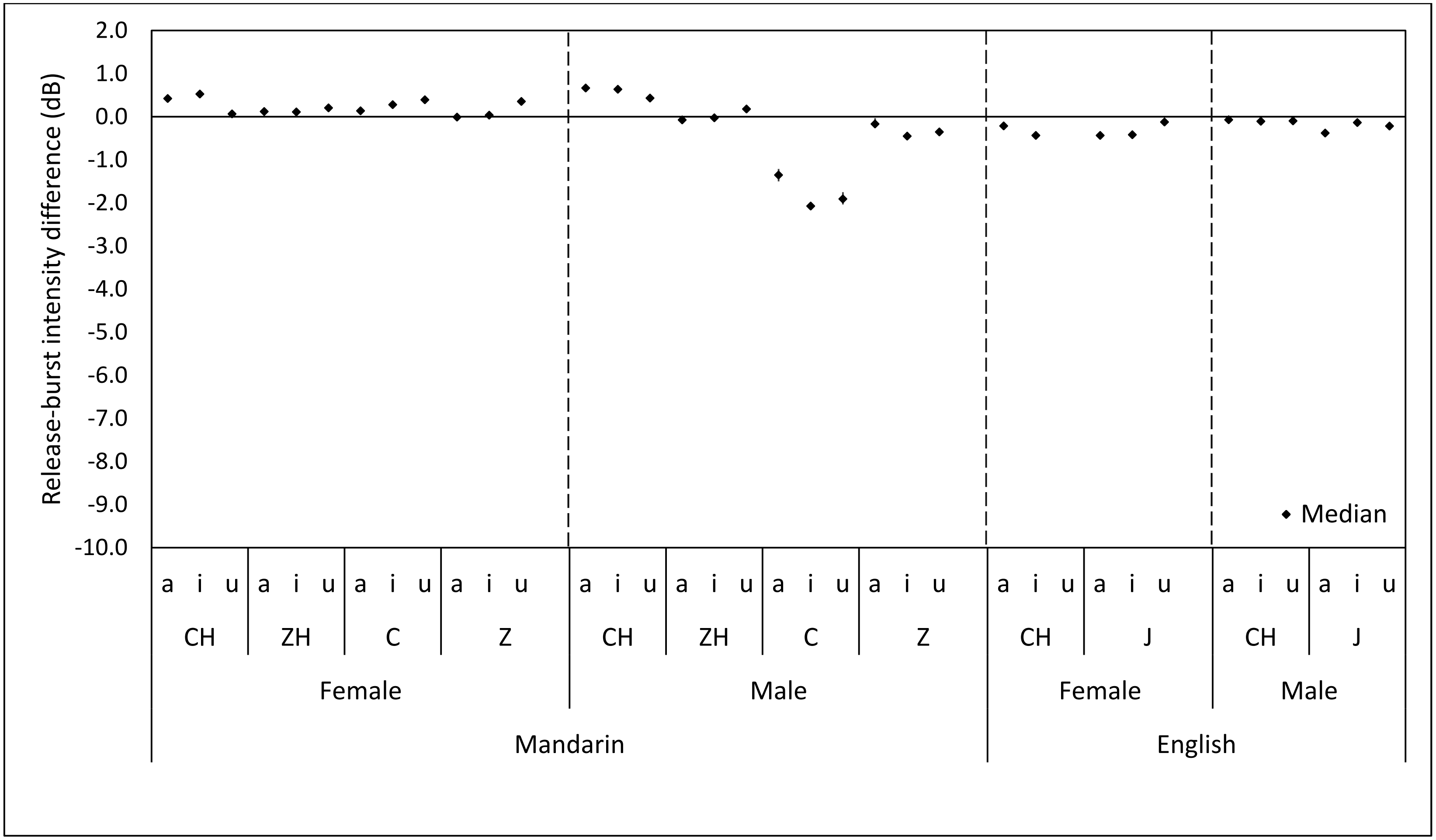
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#2 at +5 dB SNR condition. Positive values indicate intensity increment and negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix H: Figures of Frication-noise Intensity Difference for HA#2 at Three SNR Conditions (continue)



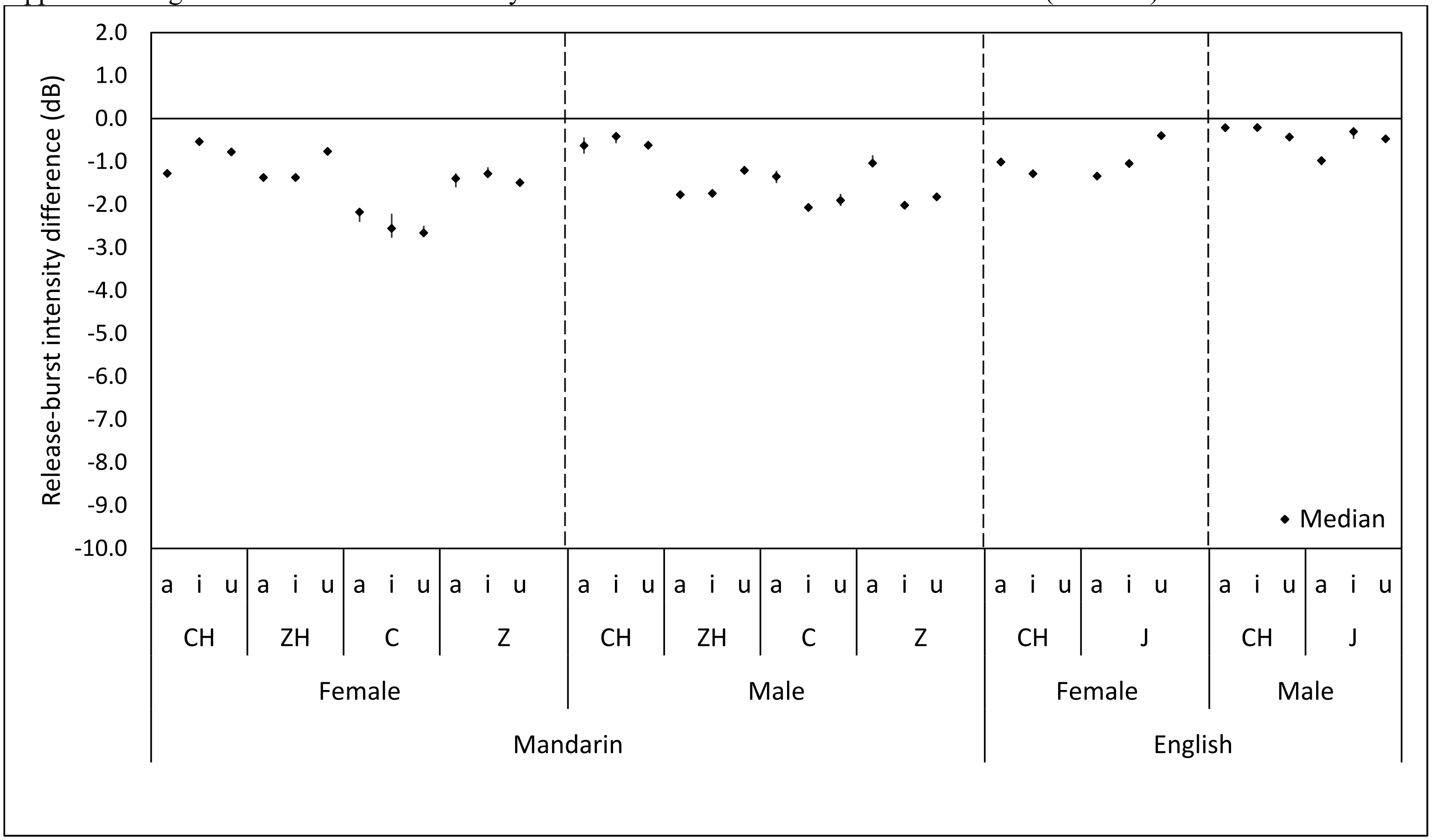
Note. Frication-noise intensity difference for fricatives and affricates processed with HA#2 at 0 dB SNR condition. Positive values indicate intensity increment and negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, SH, S, Z). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix I: Figures of Release-burst Intensity Difference for HA#1 at Five SNR Conditions



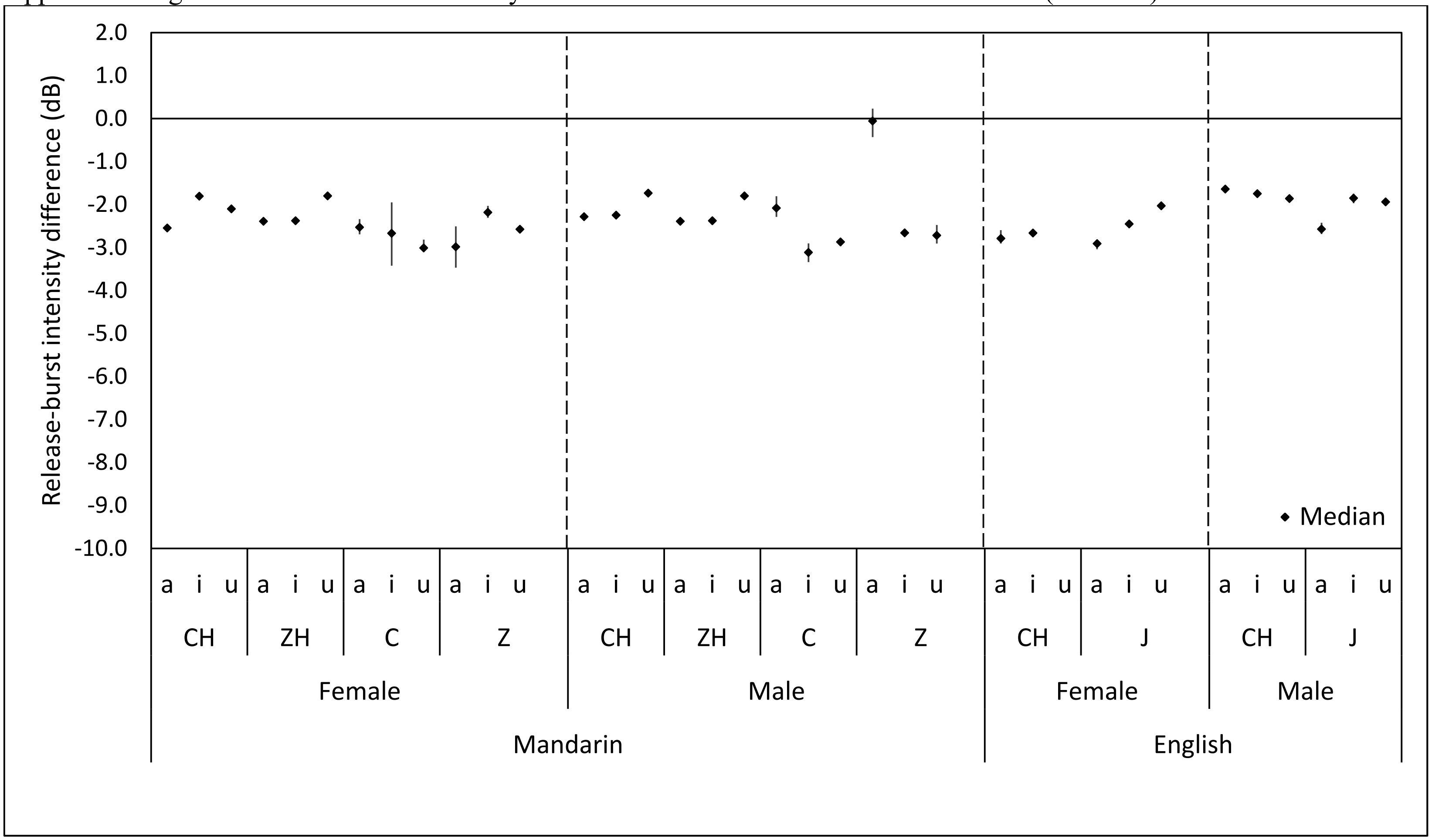
Note. Release-burst intensity difference for affricates processed with HA#1 at +10 dB SNR condition. Positive values indicate intensity increment and negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/)

Appendix I: Figures of Release-burst Intensity Difference for HA#1 at Five SNR Conditions (continue)



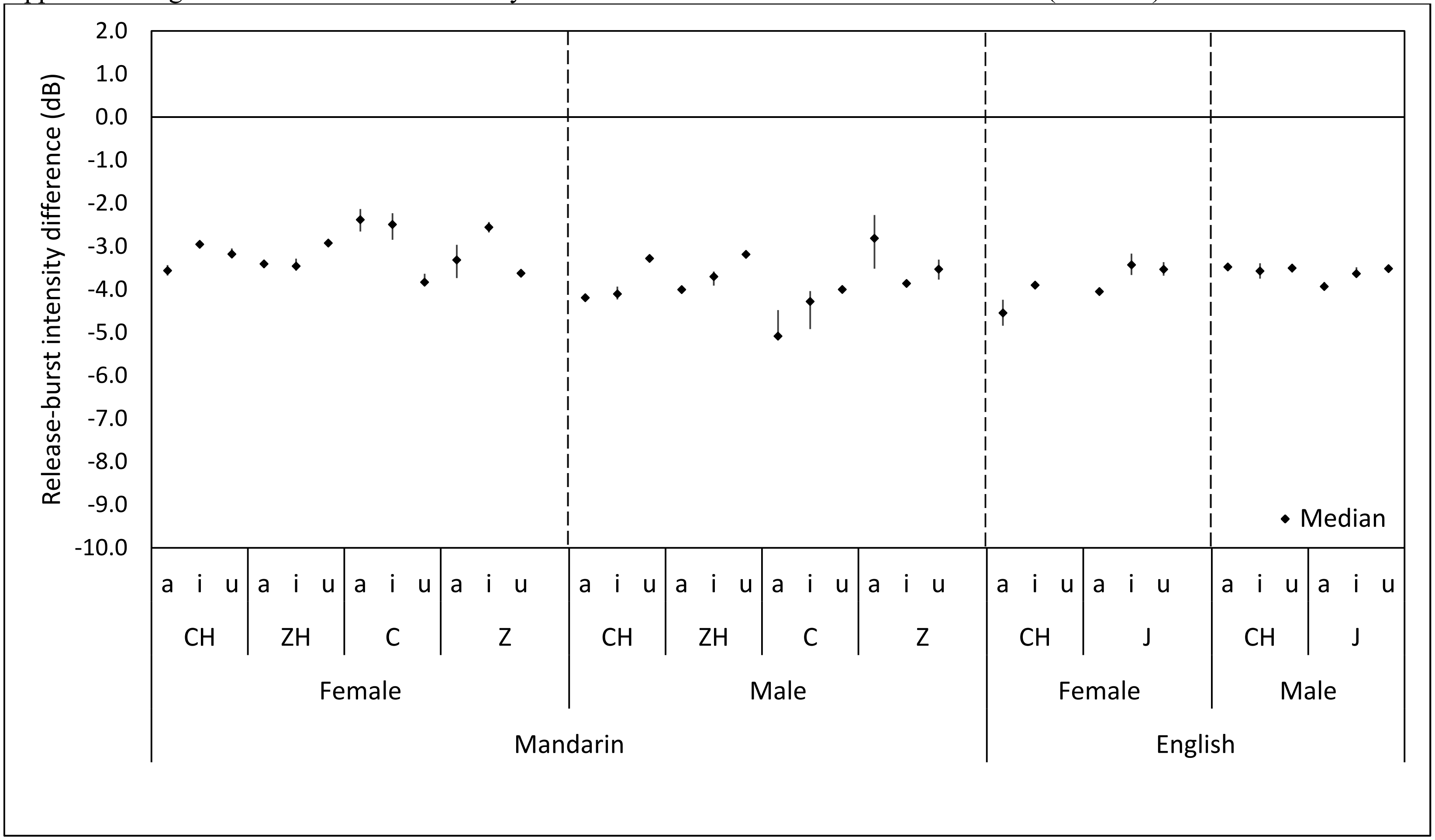
Note. Release-burst intensity difference for affricates processed with HA#1 at +5 dB SNR condition. Positive values indicate intensity increment and negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/).

Appendix I: Figures of Release-burst Intensity Difference for HA#1 at Five SNR Conditions (continue)



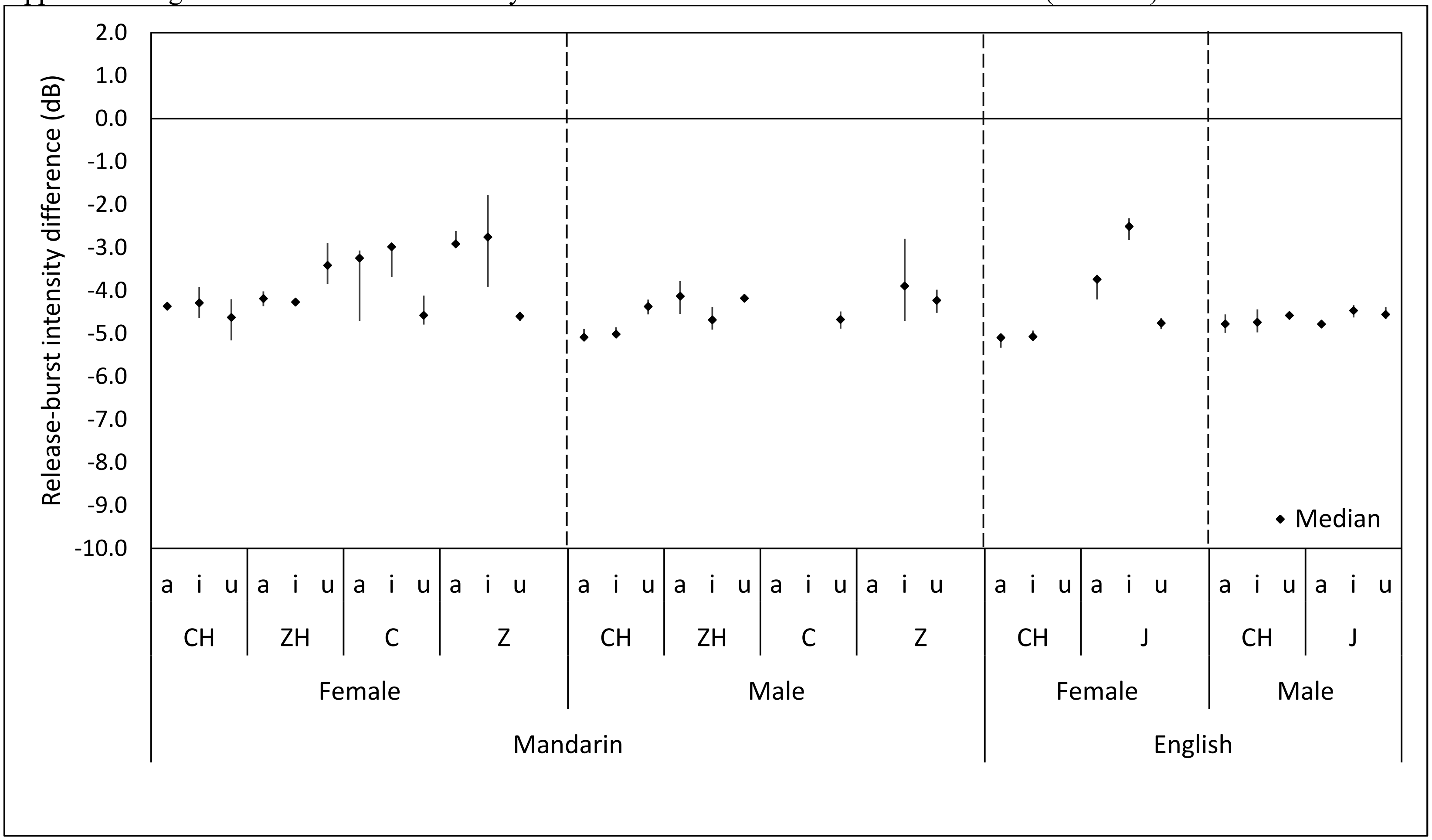
Note. Release-burst intensity difference for affricates processed with HA#1 at 0 dB SNR condition. Positive values indicate intensity increment and negative values indicate intensity reduction when SMNR was turned on. The bars represent the interquartile range and the diamond shape markers represent the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/).

Appendix I: Figures of Release-burst Intensity Difference for HA#1 at Five SNR Conditions (continue)



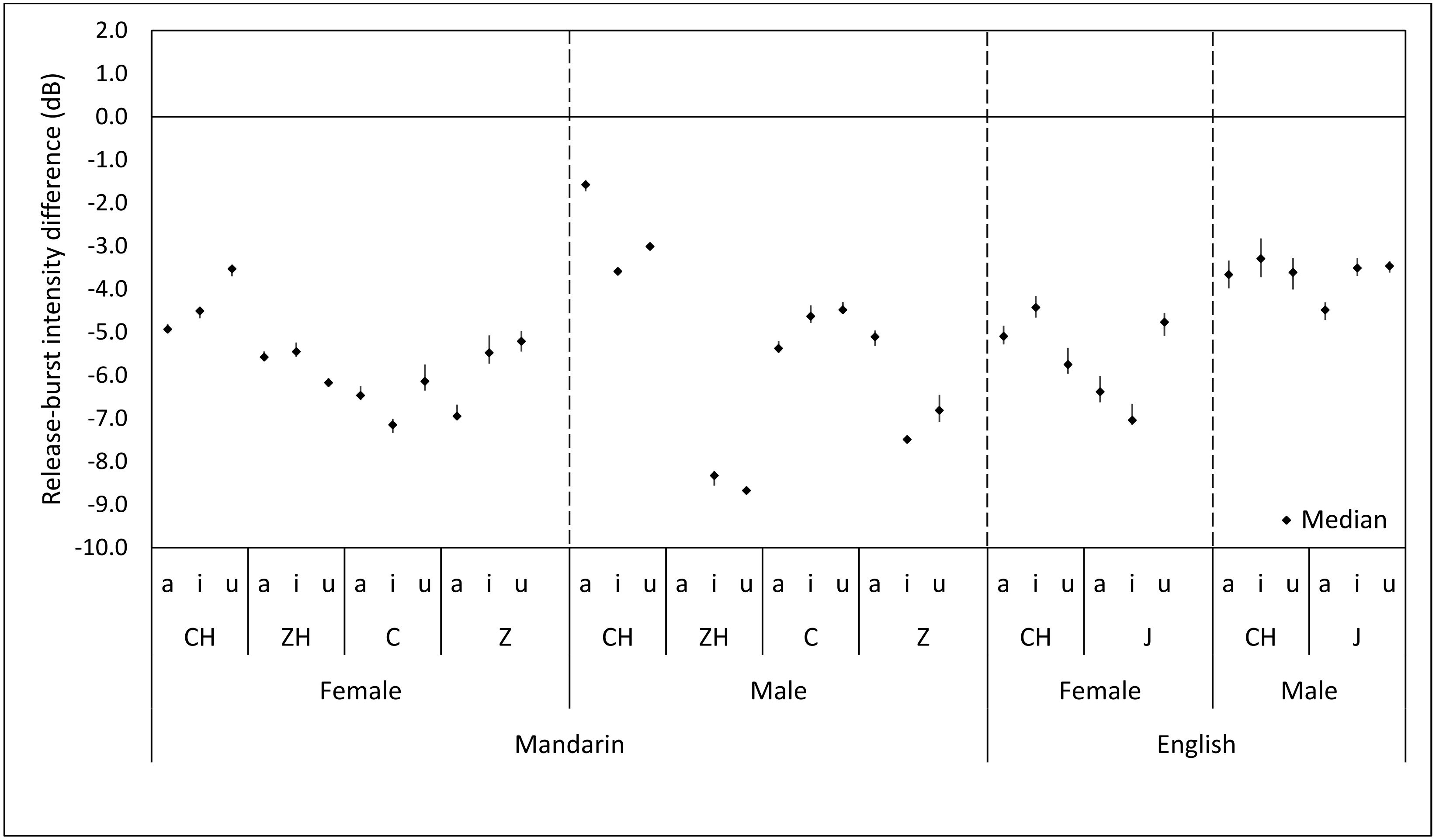
Note. Release-burst intensity difference for affricates processed with HA#1 at -5 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/)

Appendix I: Figures of Release-burst Intensity Difference for HA#1 at Five SNR Conditions (continue)



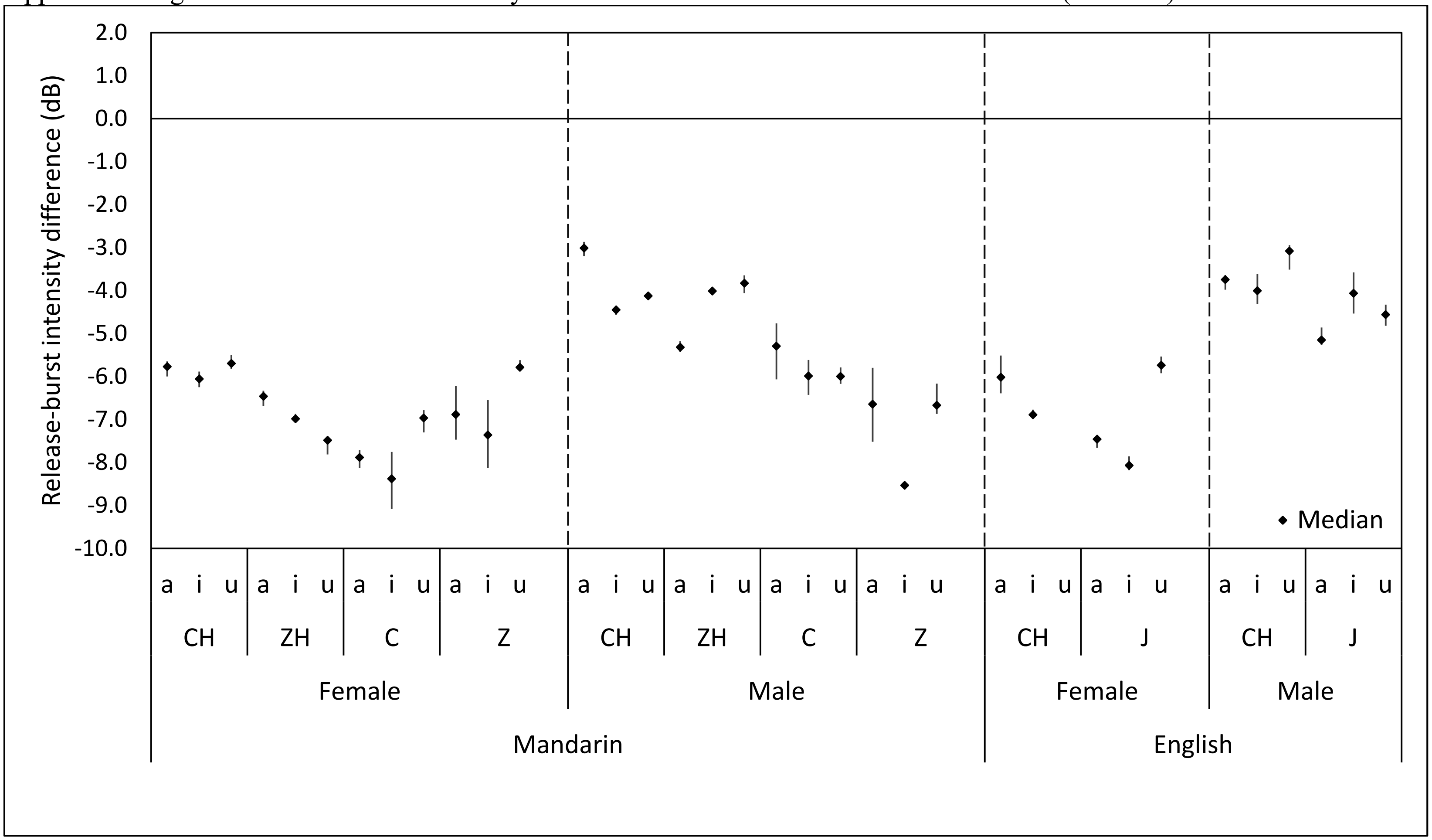
Note. Release-burst intensity difference for affricates processed with HA#1 at -10 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/).

Appendix J: Figures of Release-burst Intensity Difference for HA#2 at Three SNR Conditions



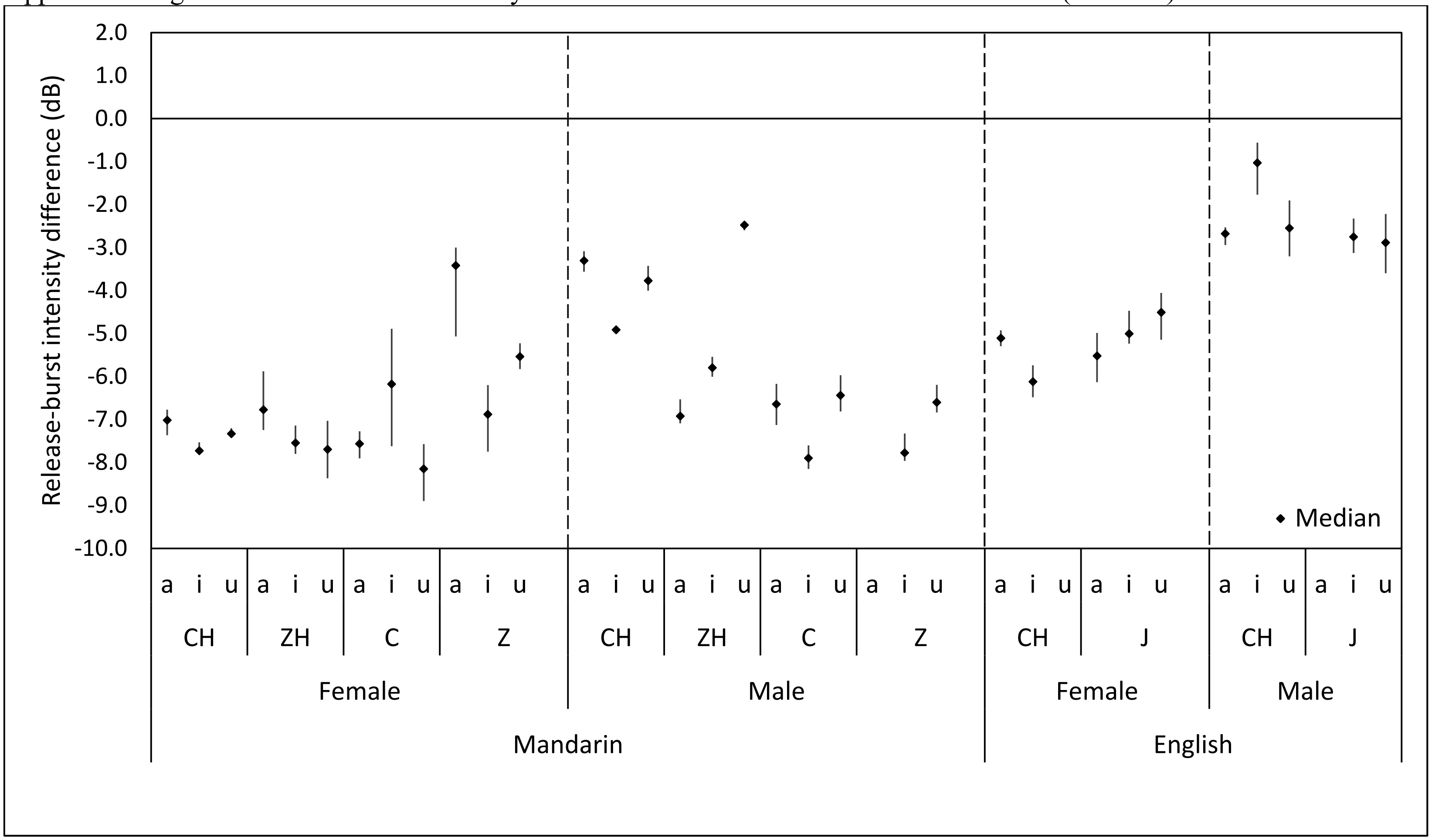
Note. Release-burst intensity difference for affricates processed with HA#2 at +10 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/).

Appendix J: Figures of Release-burst Intensity Difference for HA#2 at Three SNR Conditions (continue)



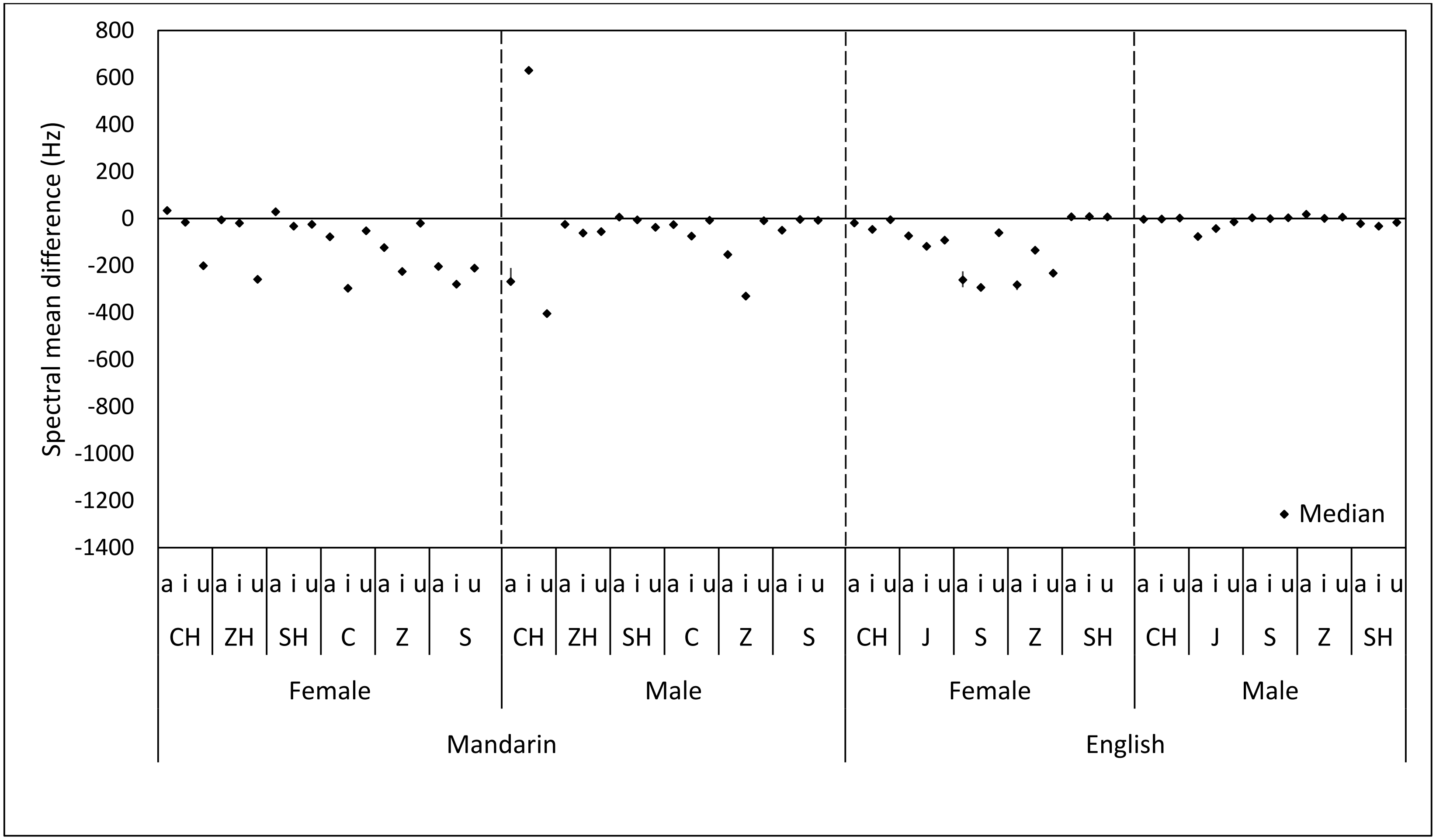
Note. Release-burst intensity difference for affricates processed with HA#2 at +5 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The bar represents the range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/)

Appendix J: Figures of Release-burst Intensity Difference for HA#2 at Three SNR Conditions (continue)



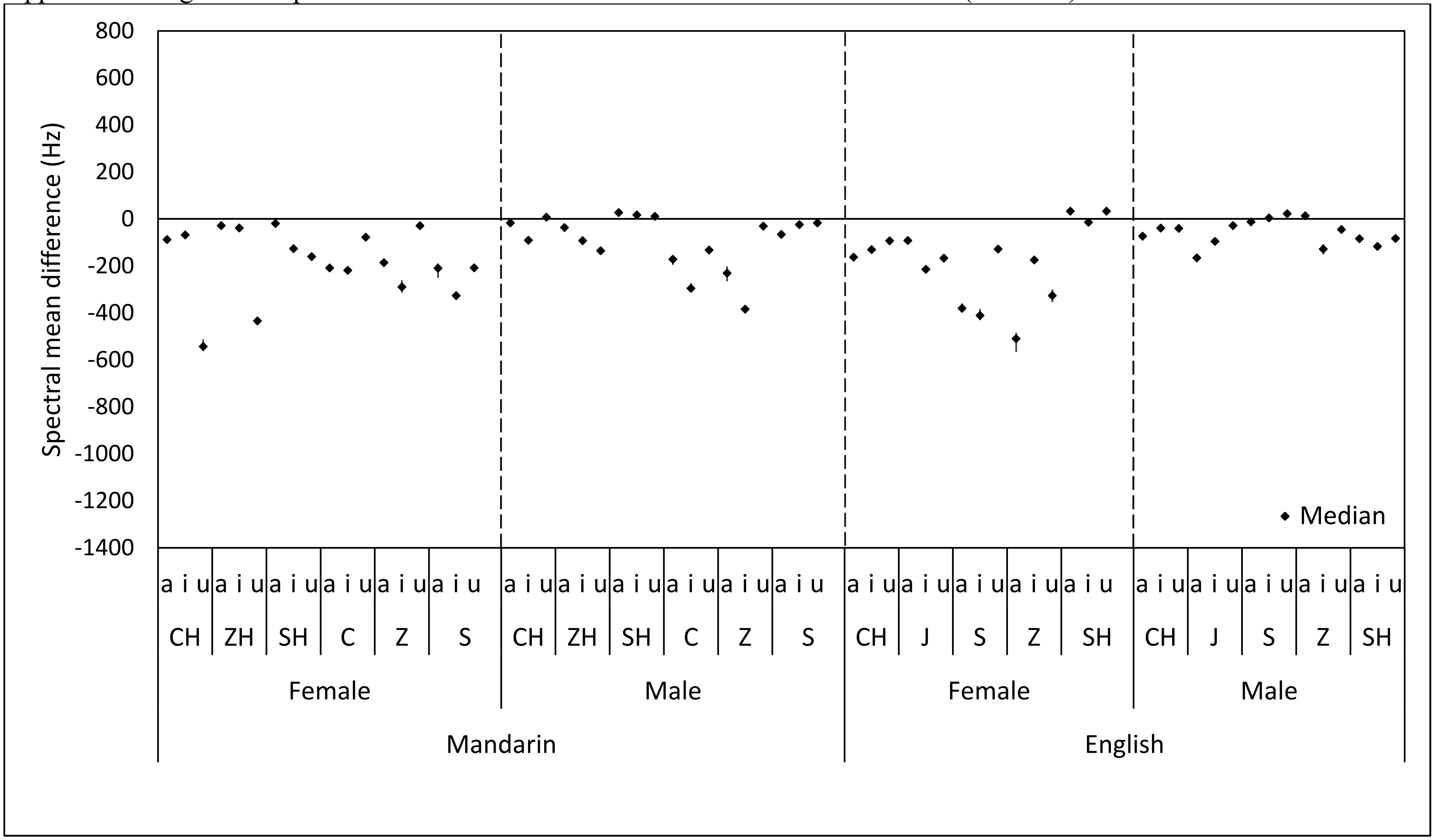
Note. Release-burst intensity difference for affricates processed with HA#2 at 0 dB SNR condition. Negative values indicate intensity reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were four Mandarin affricates (CH, ZH, C, Z) and two English affricates (CH, J). Each affricate occurred in three vowel contexts (/a, i, u/).

Appendix K: Figures of Spectral Mean Difference for HA#1 at Five SNR Conditions



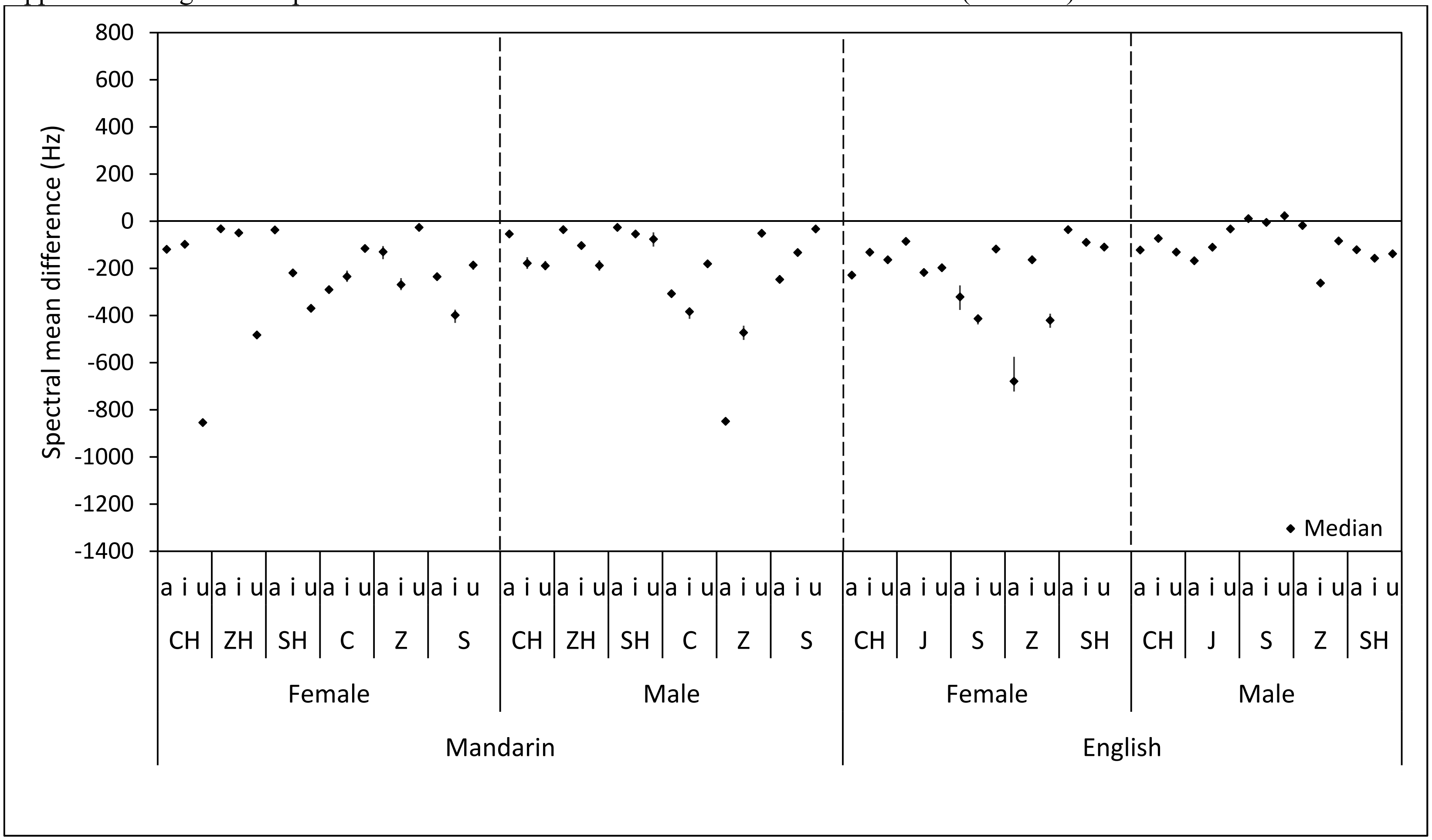
Note: Spectral mean difference for fricatives and affricates processed by HA#1 at +10 dB SNR condition. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix K: Figures of Spectral Mean Difference for HA#1 at Five SNR Conditions (continue)



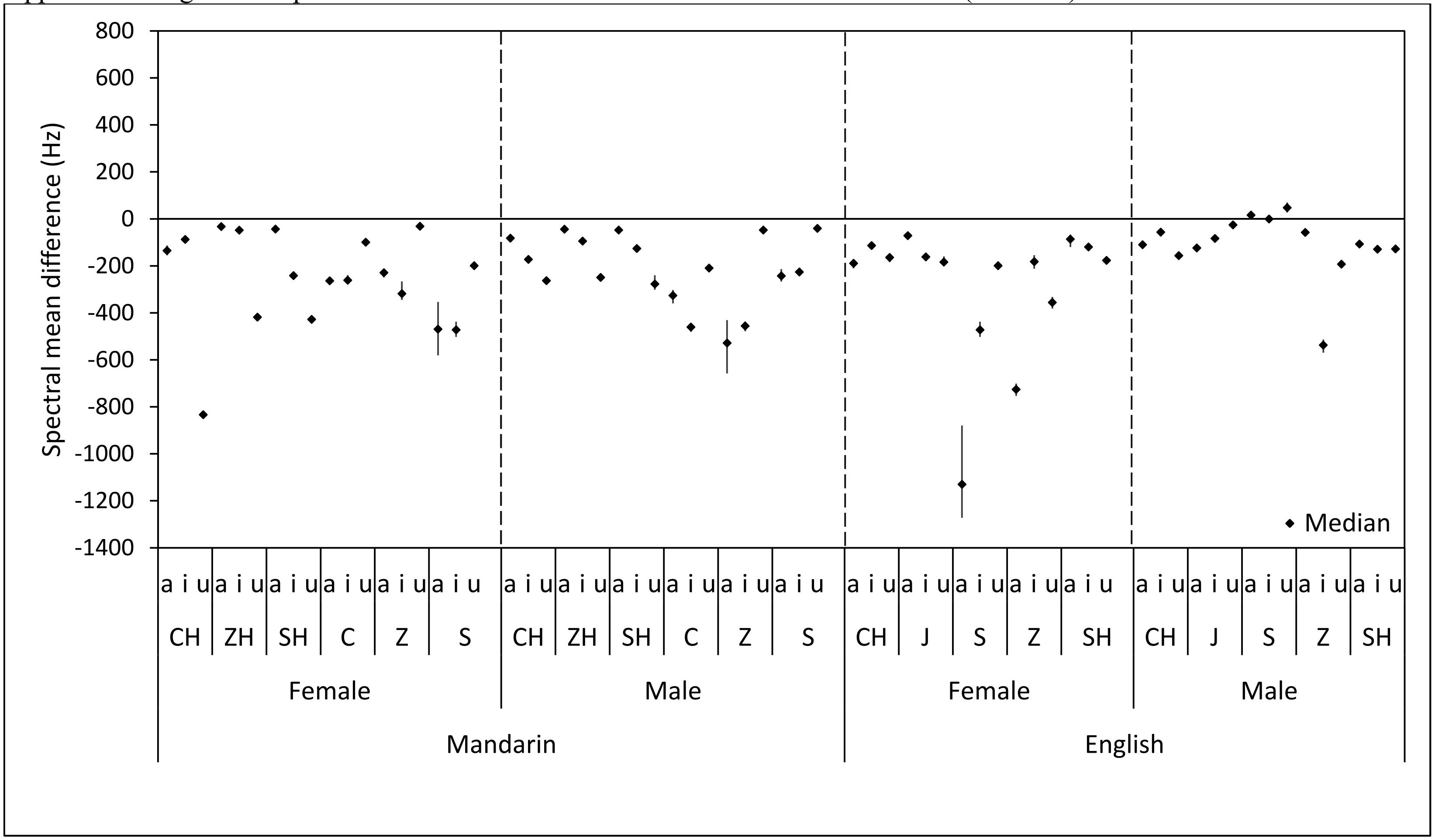
Note. Spectral mean difference for fricatives and affricates processed by HA#1 at +5 dB SNR condition. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix K: Figures of Spectral Mean Difference for HA#1 at Five SNR Conditions (continue)



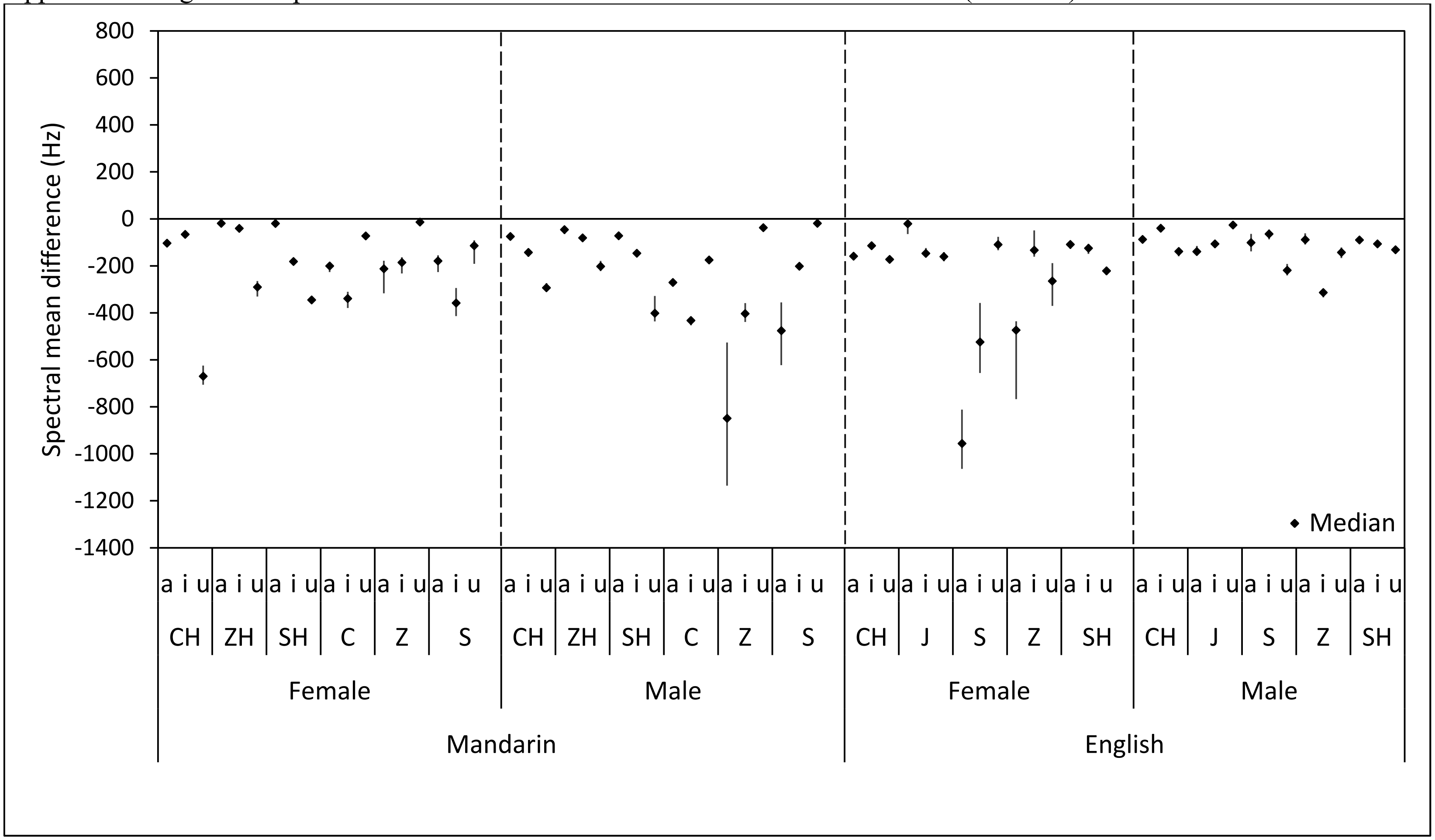
Note. Spectral mean difference for fricatives and affricates processed by HA#1 at 0 dB SNR condition. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix K: Figures of Spectral Mean Difference for HA#1 at Five SNR Conditions (continue)



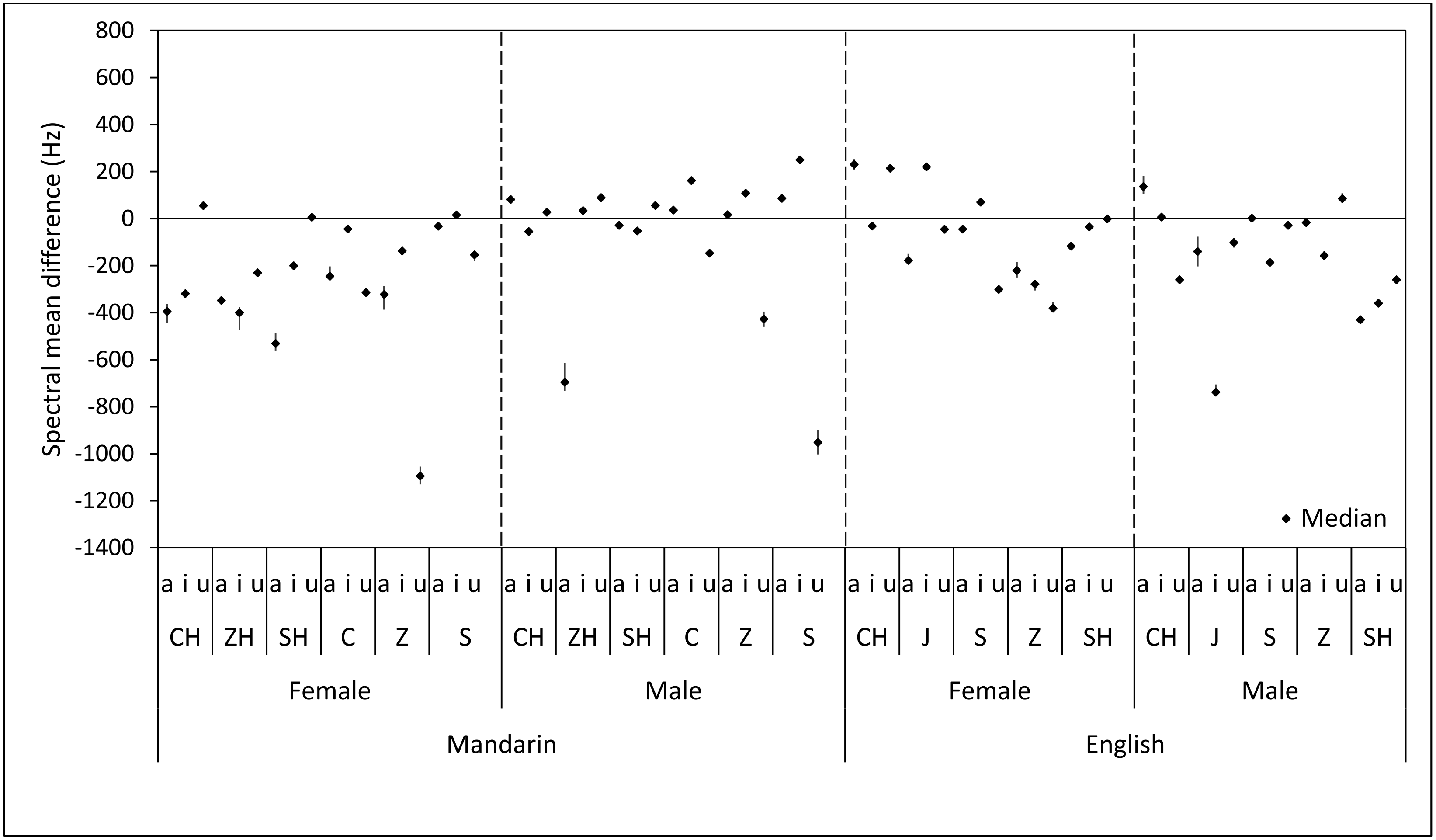
Note. Spectral mean difference for fricatives and affricates processed by HA#1 at -5 dB SNR condition. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix K: Figures of Spectral Mean Difference for HA#1 at Five SNR Conditions (continue)



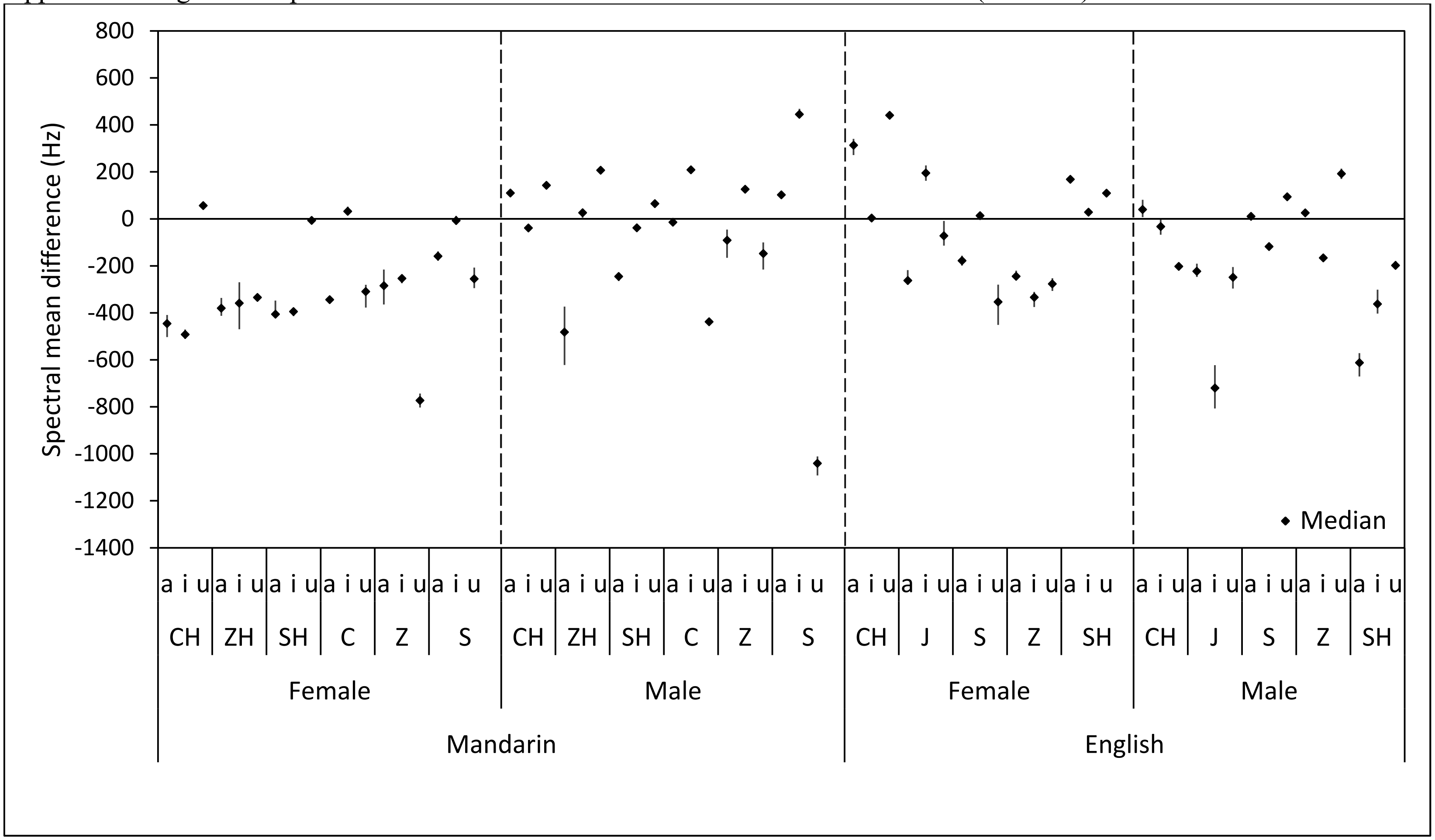
Note. Spectral mean difference for fricatives and affricates processed by HA#1 at -10 dB SNR condition. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix L: Figures of Spectral Mean Difference for HA#2 at Three SNR Conditions



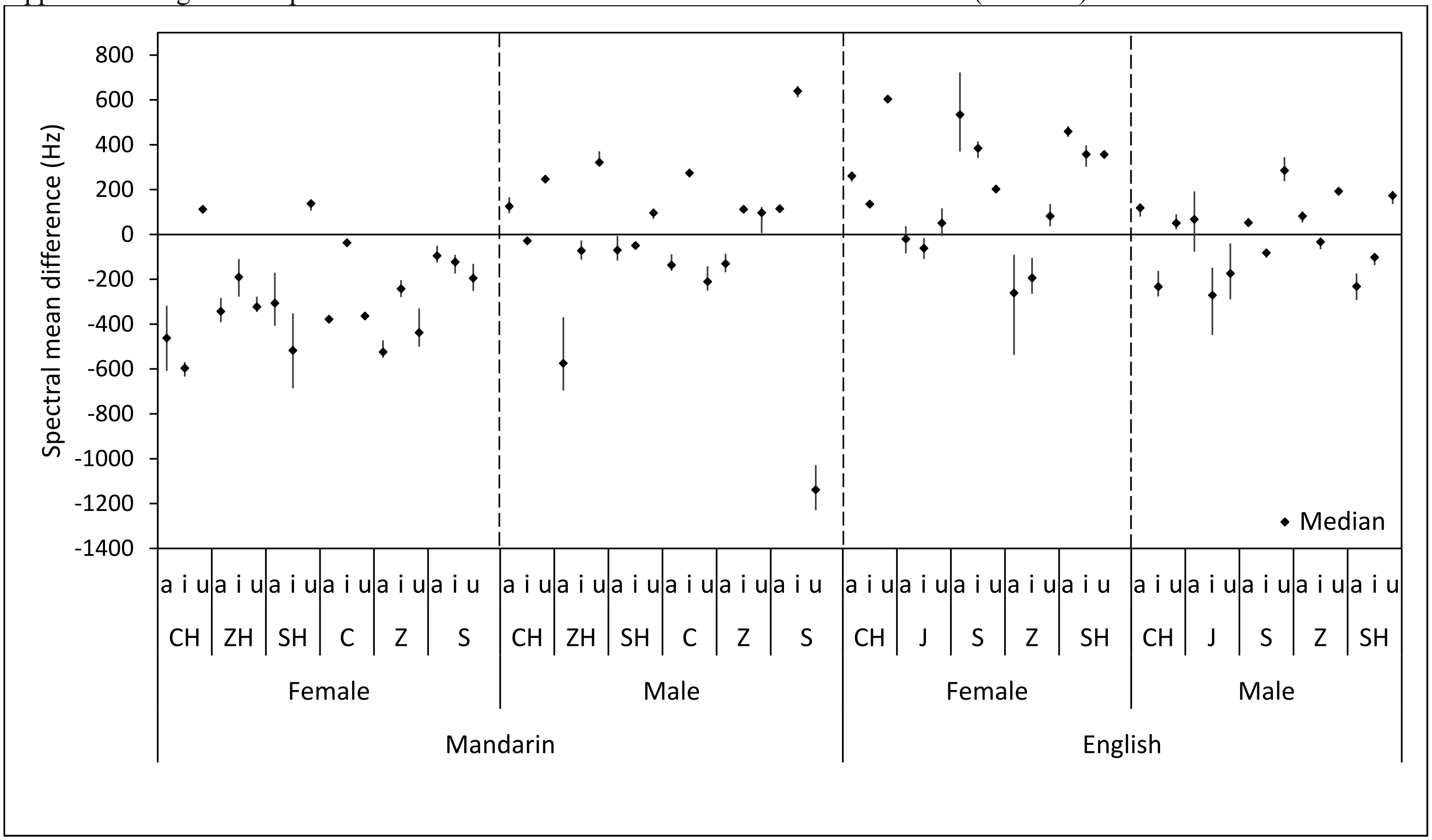
Note. Spectral mean difference for fricatives and affricates processed by HA#2 at +10 dB SNR input conditions. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix L: Figures of Spectral Mean Difference for HA#2 at Three SNR Conditions (continue)



Note. Spectral mean difference for fricatives and affricates processed by HA#2 at +5 dB SNR input conditions. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/).

Appendix L: Figures of Spectral Mean Difference for HA#2 at Three SNR Conditions (continue)



Note. Spectral mean difference for fricatives and affricates processed by HA#2 at 0 dB SNR input conditions. Positive values indicate spectral mean increment and negative values indicate spectral mean reduction when SMNR was turned on. The bar represents the interquartile range and the diamond shape marker represents the median of each category. The dotted lines separate the consonants spoken by each talker. There were six Mandarin consonants (CH, ZH, SH, C, Z, S) and five English consonants (CH, J, S, Z, SH). Each consonant occurred in three vowel contexts (/a, i, u/)

Appendix M: Tables of Average Spectral Mean Values for Consonants Processed Under SMNR-on and SMNR-off conditions

Consonant	Average spectral mean (Hz)	
	SMNR-off	SMNR-on
Mandarin Female Talker		
Retroflex aspirated affricate /tʂʰ/	3452	3198
Retroflex unaspirated affricate /tʂ/	3154	3006
Retroflex fricative /ʂ/	3413	3259
Alveolar aspirated affricate /tsʰ/	3919	3732
Alveolar unaspirated affricate /ts/	3717	3556
Alveolar fricative /s/	4184	3890
Mandarin Male Talker		
Retroflex aspirated affricate /tʂʰ/	3483	3379
Retroflex unaspirated affricate /tʂ/	3194	3101
Retroflex fricative /ʂ/	3873	3787
Alveolar aspirated affricate /tsʰ/	3413	3172
Alveolar unaspirated affricate /ts/	4000	3680
Alveolar fricative /s/	4008	3884
English Female Talker		
Voiceless post-alveolar affricate /ʧ/	3542	3403
Voiced post-alveolar affricate /ʤ/	3560	3430
Voiceless alveolar fricative /s/	4190	3885
Voiced alveolar fricative /z/	2807	2419
Voiceless post-alveolar fricative /ʃ/	3880	3808
English Male Talker		
Voiceless post-alveolar affricate /ʧ/	3091	3017
Voiced post-alveolar affricate /ʤ/	3091	3014
Voiceless alveolar fricative /s/	4878	4852
Voiced alveolar fricative /z/	4625	4500
Voiceless post-alveolar fricative /ʃ/	3001	2902

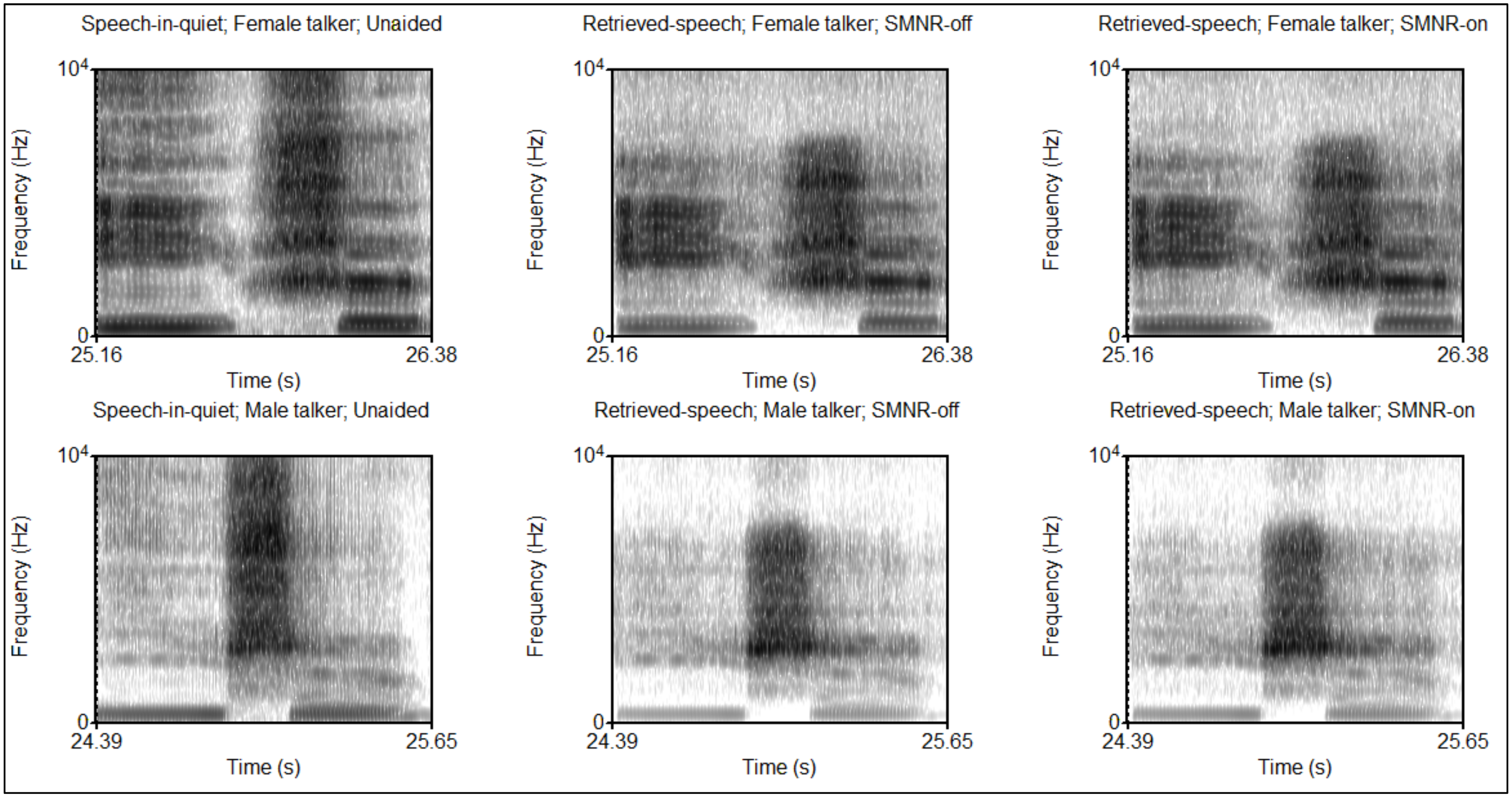
Note: The average spectral mean (Hz) for fricative and affricates retrieved from recordings of HA#1. The spectral mean value is an average of 75 measurements across three vowel contexts and five SNR conditions. These values may not reflect the unaided spectral mean values due to the gain frequency response set in each hearing aid.

Appendix M: Tables of Average Spectral Mean Values for SMNR-on and SMNR-off conditions (continue)

Consonant	Average spectral mean (Hz)	
	SMNR-off	SMNR-on
Mandarin Female Talker		
Retroflex aspirated affricate /tʂʰ/	5269	4979
Retroflex unaspirated affricate /tʂ/	4633	4319
Retroflex fricative /ʂ/	5205	4955
Alveolar aspirated affricate /tsʰ/	5662	5439
Alveolar unaspirated affricate /ts/	5541	5090
Alveolar fricative /s/	6249	6133
Mandarin Male Talker		
Retroflex aspirated affricate /tʂʰ/	5194	5255
Retroflex unaspirated affricate /tʂ/	5710	5609
Retroflex fricative /ʂ/	5483	5451
Alveolar aspirated affricate /tsʰ/	6005	5988
Alveolar unaspirated affricate /ts/	5828	5757
Alveolar fricative /s/	5849	5669
English Female Talker		
Voiceless post-alveolar affricate /ʧ/	5322	5555
Voiced post-alveolar affricate /ʤ/	5696	5684
Voiceless alveolar fricative /s/	6201	6093
Voiced alveolar fricative /z/	6069	5824
Voiceless post-alveolar fricative /ʃ/	5567	5716
English Male Talker		
Voiceless post-alveolar affricate /ʧ/	4731	4687
Voiced post-alveolar affricate /ʤ/	4750	4484
Voiceless alveolar fricative /s/	5717	5717
Voiced alveolar fricative /z/	5896	5908
Voiceless post-alveolar fricative /ʃ/	4784	4517

Note: The average spectral mean (Hz) for fricative and affricates retrieved from recordings of HA#2. Each spectral mean value is an average of 45 measurements across three vowel contexts and three SNR conditions. These values may not reflect the unaided spectral mean values due to the gain frequency response set in each hearing aid.

Appendix N: Spectrograms of Nonsense Syllables Retrieved From HA#1 Speech-plus-noise Recordings



Note: The top row shows the spectrograms of VCV syllable /işi/ spoken by a female talker. The bottom row shows the spectrograms of VCV syllables /işi/ spoken by a make talker. The left column shows the spectrograms of VCV syllables recorded in quiet (unaided). The middle column shows the spectrograms of the VCV syllables that were retrieved from the speech-plus-noise recordings from HA#1 when SMNR was turned off. The right column shows the spectrograms of the VCV syllables that were retrieved from the speech-plus-noise recordings from HA#1 when SMNR was turned on.