EFFECTS OF SPECTRAL DISTORTION ON SPEECH INTELLIGIBILITY

by

Lara Georgis

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2014

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Fine-structure cues play a role in understanding speech in fluctuating noise, but do not appear to be accessible to listeners with sensorineural hearing loss (SNHL) (Lorenzi et al., 2006). Moore (2008) proposed that fine-structure cues may help mitigate deleterious effects of multichannel compression. Recent evidence suggests that effects of compression are not influenced by fine-structure and detailed spectral information (Hopkins et al., 2012). The present study measured impact of multichannel compression and the role of fine-structure. Two indices were considered to quantify spectral distortion introduced by compression. Speech understanding in noise was measured in normal-hearing subjects. Fine-structure was removed from the signal and spectral smearing was used to simulate broadened auditory filters characteristic of SNHL. Fast multichannel compression reduced intelligibility compared to single-channel and no compression. There was no interaction between compression and fine-structure or smearing. The results support Hopkins et al. (2012), and do not support Moore's (2008) hypothesis.

LIST OF ABBREVIATIONS USED

SII	Speech Intelligibility Index
DSL	Desired Sensation Level
NAL	National Acoustics Laboratories
SNHL	Sensorineural Hearing Loss
WDRC	Wide Dynamic Range Compression
ASMC	Across-Signal Modulation Correlation
WSMC	Within-Signal Modulation Correlation
CSE	Cochlea-scaled Spectral Entropy
ERB	Equivalent Rectangular Bandwidth
SNR	Signal-to-Noise Ratio

ACKNOWLEDGEMENTS

During my thesis I was privileged to have help and support from many individuals for whom I am vey thankful. Foremost, I would like to express my sincere gratitude to my supervisor, Dr. Steve Aiken, for his support, motivation and enthusiasm. During my three years under his supervision, he generously provided me with his guidance, assistance, and expertise.

I am also sincerely thankful to my committee, Dr. Michael Kiefte, Dr. Rachel Caissie and Neil Hockley for their encouragement, insightful comments and questions.

Last, but not least, I would like to thank my family for their unconditional love, support and for teaching me to strive and to pursue my goals. I am especially grateful for my sister, Mariam Georgis, who has inspired and encouraged me to learn and grow.

Chapter 1: Introduction

1.1 Introduction To Hearing Aids

The main goal of hearing aid amplification is to make speech sounds audible to support communication. There are challenges associated with doing this because sensorineural hearing loss does not affect perception of all sounds equally. The hearing aid must provide enough gain for soft sounds without providing too much gain for louder sounds, which requires some distortion of the speech signal—generally some type of compression. Although compression plays an important role in hearing aid function, it can also degrade the signal in ways that significantly impact speech intelligibility (Plomp, 1988). Unfortunately, clinical tools used for fitting hearing aids (e.g., programming software and real-ear measurement systems) do not provide measurements of this signal degradation in a way that is relevant to speech quality or intelligibility. It would be helpful to have a better understanding of the distortion imposed by compression and particularly helpful if this could be quantified in a meaningful way for clinicians during the fitting process. This would make it possible to understand the implications of hearing aid programming decisions for both audibility and signal degradation, and could improve fittings, particularly for individuals with hearing losses that are difficult to fit (e.g., those with narrow dynamic ranges).

1.2 The Speech Intelligibility Index

Currently, the Speech Intelligibility Index (SII; ANSI S3.5, 1997, R2007) is the standard for relating speech intelligibility to audibility. The SII is based on the level of speech in a series of weighted frequency bands in relation to background noise or hearing threshold. Although the SII includes a level distortion factor, it may not be sensitive to all distortions of the speech signal that might occur in the process of increasing its audibility through compression. Moreover, although the SII can be used in a way that likely has some sensitivity to compression-related distortion, the version of the SII that is implemented in some hearing aid measurement systems (e.g., the Audioscan Verifit 3.10.60) do not do this. When using the simpler approach to meeting SII targets that is implemented in the Verifit, amplifying the average speech level above the threshold level maximizes the SII. The method of determining the SII assumes that no distortion is introduced in the process of raising the signal level apart from the level distortion factor and upward spread of masking. Current prescription methods for fitting hearing aids, such as DSL 5.0 and NAL NL2, (e.g., DSL[i/o], Seewald, Moodie, Scollie, & Bagatto, 2005; NAL-NL, Byrne et al., 2001) are similarly based on raising the level of speech above threshold without requiring measurement of signal distortions that might occur in the process, although there are recommendations for limits on the amount of compression used. Unfortunately it is not safe to assume that hearing aids adjusted to targets are free of compression-related signal degradation.

1.3 Speech Acoustics

Speech is acoustic energy produced by movement of the vocal folds or turbulent energy that is subsequently shaped by the positions of the primary articulators (i.e., the tongue, jaw and lips). The shaping results in amplitude peaks at certain frequencies that vary over time, called formants. The dips in energy interposed between these formant peaks are called troughs. The formants are more important for speech understanding than the troughs (Kiefte, Enright, & Marshall, 2010; Assman & Summerfield, 1998). Formant frequencies provide critical information for speech understanding, while the troughs do not provide independent information apart from serving as a contrast to the formants. The first two to three formants are most important (Nearey, 1989; Syndal, & Gopal, 1986), and convey information about vowel identity and consonant place of articulation. Collectively, this spectral shaping (i.e., the changing pattern of formant peaks and troughs) is called the speech envelope, and this can be differentiated from information about the speech source (i.e., harmonics and turbulent noise), which is called speech fine-structure. If speech levels are measured in short time periods that approximate the temporal integration of the human ear (e.g., 125 ms), and in relatively narrow frequency bands that approximate human critical bands (e.g., 1/3rd octave bands), the peaks of the envelope (i.e., the most intense percentile of the speech level distribution) are about 12 dB above the average speech level, and the speech troughs (generally the 20th or 30th percentile of the level distribution) are about 18 dB below the average speech level (French & Steinberg, 1947). Therefore, short-term running speech levels in narrow bands vary by approximately 30 dB. These levels fluctuate at rates between about 2-8 Hz (Drullman, Festen, & Plomp, 1994 a,b) with a peak near the syllabic rate (roughly 3-5 Hz). The overall spectro-temporal envelope of speech seems to be important for speech recognition, while the fine-structure has often been thought to be unimportant (e.g., Shannon, Zeng, Kamath, Wygnoski, & Ekelid, 1995). However, newer research suggests that fine-structure plays a significant role in hearing in noise, and reduced access to the fine structure in speech may contribute to hearing in noise difficulties experienced by people with sensorineural hearing loss (SNHL) (Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006; Moore, 2008b)

The spectral contrast between peak and trough levels required for accurate vowel identification varies among studies. Some studies have found that very low amounts of spectral contrast (about 1-3 dB) are required to achieve roughly 75% accuracy for discrimination of vowel-like harmonic complexes in normally hearing subjects (Alcantara & Moore, 1995, Leek, Dorman, & Summerfield, 1987), whereas other studies have found that more spectral contrast is required (about 8 dB) (Liu & Eddins, 2008, Dreisbach, Leek, & Lentz, 2005). The variance among findings may be contributed to differences in stimuli. Liu and Eddins (2008) manipulated other aspects of the stimuli in addition to spectral contrast, used more natural sounding vowels compared to previous studies and included more vowels to choose from, making the task more difficult and possibly requiring more spectral contrast. Individuals with SNHL require more spectral contrast for vowel identification than those with normal hearing (6-7 dB compared 1-2 dB for greater than 75% accuracy) (Leek et al., 1987). Leek et al. (1987) suggested the possibility that individuals with SNHL require more spectral contrast in the signal because broadened auditory filters lead to smaller differences between peaks and valleys internally. Interestingly, amplitude relationships between formant peaks are also important for speech understanding (Schvarts, Chatterjee, & Gordon-Salant, 2008; Liu & Eddins, 2008), especially in noise (Duquesnoy, 1983). In fact, vowel identification can be maintained when formant peaks are suppressed as long as the amplitude relationships across the spectrum are maintained (Ito, Tsuchida, & Yano, 2001; Kiefte et al., 2010). When listening in background noise, a listener's sensitivity to spectro-temporal modulations is more important than audibility alone in predicting speech intelligibility (Bernstein, Summers, Grassi, & Grant, 2013).

1.4 Compression

Sensorineural hearing loss leads to a reduced dynamic range. For example, for someone with a 70 dB HL hearing loss and a 100 dB HL upper limit of comfort, which roughly corresponds to 75.5 and 105.5 dB SPL (depending on the transducer and individual ear acoustics), the average speech level at 4000 Hz might be roughly 45 dB SPL, with peaks that are 12 dB above (57 dB SPL) and troughs that are about 18 dB below (27 dB SPL). If enough gain is provided to make the troughs audible (roughly 50 dB), the peaks of speech (107 dB SPL) will exceed the upper limit of comfort. Moreover, the hearing aid will be too soft for someone speaking more softly, and far too loud for someone speaking more loudly. Output limiting (e.g., peak clipping or high-ratio compression) prevents the hearing aid from producing sound that is too loud by clipping the peaks of the signal that reach saturation of the hearing aid. Peak clipping causes distortion and is associated with reduced intelligibility (Crain & Tassell, 1994). The solution is to provide more gain for soft sounds than loud sounds, which is called 'compression.'

Compression is accomplished by a gain reduction that is proportional to signal level, generally above a minimum signal level, called the 'threshold kneepoint.' The time over which the gain reduction occurs is called the 'attack time.' Attack time is usually short (<5 ms) so that the hearing aid can respond quickly to high-level sounds that might be uncomfortable for the listener. The 'release time' is the time it takes the hearing aid to recover from the gain reduction, and is generally longer than attack time. Compression can be classified as fast-acting or slow-acting based on attack and release times. If the gain changes quickly in response to changing input levels (i.e., fast attack and release times), the troughs will receive more gain than the peaks, reducing the depth of the energy modulations over time. If the gain changes too slowly to follow the speech modulations, peaks and troughs will receive similar gain and speech modulation depths will not be reduced. Instead, the overall level will be adjusted downwards. Fast compression is quick enough to provide adequate gain for lower levels, allowing for the more intense parts of speech to be selectively reduced. If coupled with a gain increase, a fast compressor can provide additional gain for soft sounds without providing too much gain for moderate and loud sounds. Fast-acting wide dynamic range compression (WDRC) is used in this way in modern hearing aids to make low level components of speech, such as consonants, more audible for hearing aid users while maintaining comfort for loud sounds.

Hearing loss tends to vary as a function of frequency and loudness growth varies as a function of hearing loss. This requires the use of frequency- and level-dependent gain (compression). Compression can be applied independently in different frequency regions, which are referred to as 'channels.' This 'multichannel' compression can enhance audibility; multiple discrete channels allow for an increase in gain for soft short-term speech components in one frequency region while maintaining a comfortable level for speech in other frequency regions. Hearing aids that are commercially available in 2014 offer between 3 and 20 channels, with more channels available at higher technology levels.

Although multichannel compression is useful for providing audibility for thresholds that vary across frequencies, it can alter the spectral shape of the speech. While single channel compression may lead to temporal distortions, multichannel compression can create spectral as well as temporal distortions in the signal. The troughs in the speech spectrum may be amplified to a greater extent than the peaks, reducing spectral contrast that is important for intelligibility. This is called spectral degradation. Although reducing the variance in intensity is helpful for accommodating a reduced dynamic range, distorting the spectral shape may be detrimental to intelligibility. Amplitude relationships across the spectrum provide useful information for speech understanding, as explained in the previous section on speech acoustics.

1.4.1 Evidence For Fast- And Slow-Acting Compression

It is not clear in the literature whether fast- or slow-acting compression is more beneficial, but there is some evidence that this is listener-dependent. Stone and Moore (2003) showed that fast compression has deleterious effects on speech understanding, while slow compression does not. In this study the fast compressor had an attack time of 1-2 ms. The authors investigated attack times further in 2004 in an attempt to define a point at which fast compression deteriorates speech intelligibility. It was found that, with a compression ratio of 7:1, intelligibility began to deteriorate for attack times below 8 ms, but attack times less than 2 ms were found to have the most negative effects.

Gatehouse, Naylor, and Elberling (2006a) reviewed the literature on the effect of different time-constants in compression. Of the 13 studies examined, 4 studies found no effect of time constants, 3 reported fast compression to be better than slow, 3 found slow to be better than fast, and 3 found mixed results. The results were dependent on the outcome measures used (speech intelligibility in quiet vs. in noise, or subjective ratings vs. objective measures of intelligibility and sound quality). Participants reported that slow-acting compression was more comfortable while fast-acting compression yielded better speech intelligibility scores as well as perceived intelligibility. The authors found that none of the fittings were superior in all outcome measures and even within a single outcome measure there was individual variability among participants. Fast-acting WDRC provides an opportunity for maximizing audibility, but this might be at the expense of greater distortion. Once audibility is accounted for, fast compression effects changes in amplitude over time and leads to decreased speech recognition compared to linear processing in individuals with severe to profound SNHL (Souza, Jenstad, & Folino, 2005).

1.5 Temporal And Spectral Distortion Caused By Compression

Both fast and slow acting compression cause temporal distortion. However, fast-acting multichannel compression, although it improves audibility and comfort for a wide range of speech sounds, may lead to distortions in the spectral envelope as well as reduction in temporal modulations. Fast-acting multichannel compression raises the low-level short-term portions of the speech spectrum and reduces the high-intensity portions, effectively flattening it (Plomp, 1988). Plomp (1988) showed that compression distorts the spectro-temporal contrasts in speech, reducing intelligibility, although a high compression ratio of 10:1 was used in this experiment. Stone and Moore (2003) suggest that another potential reason why fast compression degrades performance is, in instances of listening in background noise, it introduces correlated fluctuations in amplitude, or comodulation, between the target signal and competing speech or background noise. For instance, if there is a peak in one signal, the compressor reduces the gain for both signals, which leads to partially correlated modulations and the potential for perceptual fusion of the sources. Stone and Moore (2007) propose a method for calculating this correlation, called across-signal modulation correlation (ASMC).

Listeners with SNHL have reduced frequency selectivity (Glasberg & Moore, 1986; Moore & Glasberg, 1986b), which can increase difficulty in coping with distortions in the spectral pattern. Furthermore, since individuals with hearing loss have been shown to benefit from increased spectral contrast (Leek et al, 1987; Dreisbach et al., 2005), compression that reduces spectral contrast could have an adverse effect for these individuals. Reduced temporal and spectral contrast may be especially detrimental for people with SNHL in demanding listening environments, such as in background noise (ter Keurs, Festen, & Plomp, 1993; Baer & Moore, 1993; 1994, Turner, Chi, & Flock, 1999).

Age and level of cognitive functioning are factors that may influence the listener's ability to overcome spectral degradation introduced by compression (Schvarts et al., 2008). In this study, younger adults performed better than older adults, while middle-aged and older adults performed very similarly. Age was the biggest predictor of speech intelligibility performance across all 3 age groups. However, among the middle-aged and older adults, processing speed and verbal memory abilities were better predictors of performance. The findings suggest that as listeners get older, cognitive ability plays a larger role in speech intelligibility and can interact with hearing loss.

Gatehouse, Naylor, & Elberling (2006b) investigated the role of individual variability beyond auditory capability in perceived benefit from WDRC and also found cognitive ability to play a role. Fast compression tends to be more beneficial to adults whose auditory environments change rapidly and who have relatively high cognitive capacities, while slow acting compression seems better suited for adults with slowly changing auditory environments and lower cognitive capacity. The authors reason that adults with high cognitive capacity are able to compensate for the reduced spectral and temporal contrast caused by fast-acting WDRC and are therefore able to take advantage of the increased short-term audibility provided by fast compression. Furthermore, fast compression provides more benefit than slow compression to listeners with high cognitive ability when listening in background noise with a temporal envelope because it amplifies the speech signal found in the temporal "dips."

1.5.1 Number Of Compression Channels

Intelligibility decreases as compression speed and the number of compression channels increases (Stone & Moore, 2008). There appears to be a trade-off between compression speed and the number of channels used in multi-channel compression. Decreasing the number of channels improves spectral contrast and counteracts the effects of reduced modulations caused by fast-acting compression. In one study, increasing the number of compression channels negatively affected performance with fast compression, but not with slow compression (Stone & Moore, 2008). In another study with 15 compression channels, subjects with hearing loss subjectively preferred slow compression to fast compression based on speech intelligibility and sound quality (Hansen, 2002). When fewer channels are used (up to 4 channels), speech recognition performance is similar to that observed with single channel compression (Keidser & Grant, 2001; Plyler, 2013).

1.6 Speech Fine-Structure

Speech stimuli can be processed to isolate envelope or fine-structure cues in order to find the impact of each on its own. When speech is processed to preserve only the envelope cues and remove fine-structure, 50% speech intelligibility can be achieved in quiet with as little as 3 spectral bands (Shannon et al., 1995). Even for listeners with SNHL, envelope cues are sufficient for achieving adequate speech intelligibility in quiet (Souza & Boike, 2006). However, when background noise is introduced, steady or modulated, more bands are required to achieve good speech intelligibility when only envelope cues are available (Stone & Moore, 2003; Friesen, Shannon, Baskent, & Wang, 2001).

Cochlear hearing loss can impair the processing of fine-structure cues to a greater extent than envelope cues, such that the ability to process fine-structure is reduced with even mild to moderate hearing loss, although the degree of fine-structure deficit varies among individuals (Lorenzi, Husson, Ardoint, & Debruille, 2006; Ardoint, Sheft, Fleuriot, Garnier, & Lorenzi, 2010). A reduced ability to process fine-structure may account for the degraded speech performance observed in people with SNHL (Lorenzi et al., 2006; Moore, 2008b; Moore, Glasberg, & Hopkins, 2006).

1.7 Release From Masking

Listening in background noise is a significant concern for people with SNHL. People with normal hearing experience some difficulty in noise as well, but appear to benefit from temporal modulations in noise with a fluctuating envelope, such as a competing speaker. The relatively silent “dips” in the spectrum of a single competing talker provide the listener with important information from the target speech signal. The same advantage is not observed with steady-state noise (Eisenberg, Dirks, & Bell, 1995; Festen & Plomp, 1990). Improvement in speech intelligibility in the presence of fluctuating background noise is referred to as “release from masking” or “masking release” and can vary with the depth and rate of fluctuation (Nelson, Jin, Carney, & Nelson, 2003). This release from masking is reduced or absent in people with SNHL, whose performance tends to be similar in steady-state and fluctuating background noise (Eisenberg et al., 1995; Duquesnoy, 1983; Gatehouse, Naylor, & Elberling, 2003; Festen & Plomp, 1990; Peters, Moore, & Baer, 1998, Nelson et al., 2003). Masking release is observed in some listeners with SNHL, but it not as robust as in normal hearing listeners (Lorenzi et al., 2006). Even individuals with mild SNHL (Takahashi & Bacon, 1992) and nearly normal hearing (Middelweerd, Festen, & Plomp, 1990) show a deficit with respect to masking release compared to controls with normal hearing.

Several factors likely contribute to the reduced release from masking observed in people with hearing loss. One possibility is that people with SNHL cannot access information in the “dips” due to reduced audibility; speech information in the dips falls below their threshold (Desloge, Reed, Braida, Perez, & Delhorne, 2010; George, Festen, & Houtgast, 2006). Peters et al. (1998) found that with the use of linear amplification with a frequency-gain characteristic meeting NAL-R targets, some of the low-level parts of the speech spectrum were below the listeners' absolute thresholds. The use of fast-acting WDRC could theoretically increase audibility of low-level parts of speech and improve audibility in the dips. Nevertheless, performance of listeners with hearing loss was disproportionately lower than age-matched controls and could not be explained by the limitations in audibility alone, suggesting that suprathreshold factors must also be involved. This is supported by studies showing that suprathreshold factors such as reduced frequency selectivity and temporal resolution contribute to poor performance (Baer & Moore, 1993; 1994; Dubno et al., 2003; Leger et al., 2012). Increased forward masking is also observed in listeners with hearing loss and is correlated with decreased masking release (Thibodeau, 1991; Bacon et al., 1998; Dubno, Horwitz, & Ahlstrom, 2002).

1.7.1 Release From Masking And Fine-Structure

Reduced audibility, temporal envelope processing and frequency selectivity do not fully account for decreased masking release in listeners with SNHL (Leger, Moore, & Lorenzi, 2012a; 2012b). With good frequency resolution representative of normal auditory function, Hopkins, King, & Moore (2009) were able to manipulate masking release consistently by increasing access to speech fine-structure. Leger and colleagues (2012) reasoned that since fine-structure contributes generally to speech intelligibility in quiet as well as modulated and unmodulated noise, it must be involved in masking release. There is evidence to suggest that fine-structure is important for speech recognition and speaker identification in background noise. When listening in modulated noise, changes in fine-structure occurring in the dips may provide information about the target speech to help the listener identify the speech and its characteristics (Moore, 2008b; Friesen et al., 2001;

Hopkins, Moore, & Stone, 2008). Consequently, reduced ability to process speech fine-structure may limit how much information can be extracted from the dips in modulated noise (Lorenzi et al., 2006; Nelson et al., 2003, Moore, 2008b). When listeners with normal hearing are provided with temporal envelope cues but not fine-structure, their performance is similar to those with SNHL - they are unable to take advantage of the speech information available in the dips (Nelson et al., 2003). This supports the hypothesis that access to fine-structure plays a role in extracting speech information from the temporal dips in noise. Furthermore, there is a significant correlation between the ability to process fine-structure and degree of masking release (Lorenzi et al., 2006): listeners with hearing loss who were able to derive some benefit from speech that was processed to contain only fine-structure cues in quiet demonstrated relatively large degree of masking release.

If reduced ability to process fine-structure is related to speech-in-noise issues faced by people with hearing loss, it would be of value to try to understand what causes it. Hopkins and Moore (2011) investigated whether aging or broadened auditory filters could lead to the reduced fine-structure sensitivity observed in people with SNHL. Interestingly, it was found that even older adults without hearing loss showed a deficit in fine-structure sensitivity. It seems that reduced fine-structure sensitivity might not be caused by factors related to hearing loss, such as reduced frequency selectivity, but rather factors related to aging. Moreover, studies have shown that audiometric thresholds alone do not explain reduced fine-structure sensitivity. Speech intelligibility is negatively impacted by suprathreshold auditory deficits in instances of high frequency hearing loss and low-pass filtered stimuli, where audiometric thresholds are normal in the frequency regions tested (Lorenzi, Debruille, Garnier, Fleuriot, & Moore, 2009; Strelcyk & Dau, 2009; Leger et al., 2012a,b; Feng, Yin, Kiefte, & Wang, 2009). Leger et al. (2012b) used a test of sensitivity to fine-structure at low frequencies developed by Hopkins and Moore (2010) to verify the role of fine-structure in masking release and, contrary to previous findings, found that sensitivity to fine-structure is not related to speech intelligibility scores in noise once audiometric thresholds are accounted for. However, fine-structure sensitivity was not measured for frequencies above 750 Hz.

Hopkins et al. (2008) investigated the effect of varying fine-structure information on speech recognition in fluctuating background noise. Listeners with SNHL did not benefit from added fine-structure. Their performance improved significantly less than normal hearing subjects with increased fine-structure information. These results support the hypothesis that fine-structure information is important for speech recognition in fluctuating background noise and individuals with SNHL cannot access it as well as listeners with normal hearing. However, there were individual differences in the amount of benefit among subjects with hearing loss and the authors found no correlation between audiometric thresholds and benefit from added fine-structure information. Lunner, Hietkamp, Anderson, Hopkins, and Moore (2012) verified these findings and followed up Hopkins et al.'s (2008) study. Their results provide support for the role of fine-structure in segregating target speech from background noise.

Table 1 summarizes several studies that assessed the impact of either compression or loss of fine-structure on speech intelligibility. Some of the studies (Stone & Moore 2003; 2004; 2008) used noise vocoding with a relatively low number of spectral channels and found decreases in speech intelligibility with increases in number of compression channels or speed. Other studies (Hopkins & Moore 2009, Hopkins et al., 2008, Lunner et al., 2012) used tone vocoding and 30-32 channels (without compression), and have also found fine-structure to be important for speech understanding in modulated noise. Only two studies (Hopkins et al., 2012; Stone, Fullgrabe, & Moore, 2009) looked at both compression and fine-structure. These are discussed in more detail in Section 1.8.

Table 1. Comparison of studies looking at the impact of compression and reduced fine-structure on speech understanding in noise. The studies that manipulated compression (Stone & Moore 2003; 2004; 2008) found reduced intelligibility with increased channels and compression speed. The studies that manipulated fine-structure (Hopkins & Moore, 2009; Hopkins et al., 2008; Lunner et al., 2012) found that performance did not improve with increased fine-structure for individuals with hearing loss. Hopkins et al. (2012) and Stone et al. (2009) tested both compression and fine-structure and are discussed in section 1.8.

	Compression Channels	**Hearing Loss Simulation**	**Type of Noise**	**Main Finding**
Stone & Moore 2003	6, 8, 11, 16	Noise vocoding – 4, 6, 8, 11, or 16 channels	Single speaker reading naturally from a script	Fast-acting compression reduced performance; slow-compression did not
Stone & Moore 2004	6, 11	Noise vocoding – 6 or 11 channels	Single speaker reading naturally from a script	Fast-acting, compression reduced intelligibility, performance degraded with comodulation
Stone & Moore 2008	1, 3, 6, 12	Noise vocoding – 8, 12 or 18 channels	Single speaker, highly modulated speech in long and short term	Intelligibility decreased with increased compression speed and number of compression channels
Stone et al. 2009	1, 2, 4, 8	Tone vocoding – 8 or 16 channels	Single speaker with f0 0.5 octaves than target	High-rate envelope cues in many vocoder channels reduces negative effect of fast compression

Hopkins et al. 2012	**6**	Tone vocoding – 30 channels	Single speaker reading a passage with f0 ranging from 100-200 Hz (target speaker f0 75-150 Hz)	Better intelligibility with fast compression, regardless of fine-structure cues and detailed spectral information
Current Study	1, 18	Spectral smearing, fine-structure completely removed from signal	Single speaker, reversed target sentence, f0 shifted up	
Hopkins & Moore 2009	N/A	Tone vocoding above varying cut-off channel – 32 channels	Steady and amplitude modulated noise	Fine-structure cues improve intelligibility in modulated noise
Hopkins et al. 2008	N/A	Subjects with SNHL; Exp. 1: fine-structure removed via noise vocoding above varying cut-off channel – 32 channels Exp. 2: fine-structure removed via tone vocoding above varying cut-off channel – 8 or 16 channels	Competing speaker with f0 ranging from 130-280 Hz (target speaker f0 130-200 Hz)	Listeners with hearing loss do not benefit from added fine-structure cues in presence of fluctuating noise
Lunner et al. 2012	N/A	Subjects with SNHL; fine-structure removed via tone vocoding above varying cut-off channel – 32 channels	Competing speech of continuous prose with f0 overlapping that of target speaker	Listeners with hearing loss do not benefit from added fine-structure cues in fluctuating noise

1.8 Fine-Structure And Compression

As discussed above, speech fine-structure appears to be important for speech recognition in noise; fine-structure heard in the dips of fluctuating background noise helps to segregate the speaker's formant patterns from the background noise. Therefore, if there is a deficit in processing fine-structure, simply amplifying the signal above threshold for listeners with hearing loss may not improve the ability to listen in background noise. However, if a hearing aid user has good sensitivity to fine-structure, fast-acting compression might be more beneficial than slow if it can improve audibility of the signal in the dips (Moore, Peters, & Stone, 1999). Using background noise with spectral and temporal dips, the authors looked at the potential benefit of fast-acting compression on speech understanding in noise. Although compression was better than linear amplification, aided performance for listeners with SNHL was still worse than for listeners with near-normal hearing. The authors suggest that deficits in cochlear hearing loss that affect sound processing may be contributing to the poor benefit gained from amplification in this listening environment. Amplification does not fully remediate deficits such as reduced frequency selectivity, reduced temporal resolution for sounds with fluctuating envelopes and reduced access to speech fine-structure. It has been suggested that with reduced access to fine-structure, those with cochlear hearing loss likely rely more on spectro-temporal envelope cues for speech understanding; which can be degraded by both single- and multi-channel fast-acting compression more so than slow compression (Stone & Moore, 2008), although this was not tested directly. Moore (2008) similarly reviewed the role of individual differences in choosing the best compression speed and suggested that individuals with hearing loss who have reduced access to fine-structure will consequently rely more heavily on temporal envelope cues which may be distorted by fast-acting compression. Moore (2008) also suggested that an individual's sensitivity to fine-structure might be related to the potential benefit gained from fast-acting compression, particularly for those with moderate to profound SNHL listening in background noise, since fast-acting compression could provide audibility for the dips in fluctuating noise and access to fine-structure appears to be necessary for masking release (Nelson et al., 2003). Therefore, Moore (2008) suggested that a method for measuring an individual's ability to process fine-structure might be useful in deciding the most suitable compression speed.

Only two studies have tested the hypothesis that access to fine-structure relates to the optimum compression speeds. Stone et al., (2009) used a tone-vocoder to limit access to fine structure and presented compression at various speeds. However, in some conditions the low-pass filter used with the envelope was raised from 50 Hz, which only provides access to the low-frequency speech envelope (e.g., formant peaks and troughs), to 200Hz, which also provides access to the periodicity envelope. The periodicity envelope is derived from speech fine-structure when two or more harmonics occur within a single vocoder channel. They found that the effects of compression speed were reduced when high-rate envelope cues were available. In contrast, Hopkins et al. (2012) found no relationship between the impact of compression speed and fine-structure on speech intelligibility. Fast multichannel compression yielded better speech intelligibility than slow compression, regardless of fine-structure availability. The authors provide several possible explanations for the discrepancy between their findings and previous studies. For instance, the authors used tone vocoding rather than noise vocoding, which may not completely remove fine-structure information and provide some information about the signal's spectro-temporal structure. Furthermore, the stimuli in the study were processed in 30 overlapping 1- or 2- ERB wide vocoder channels, resulting in good spectral resolution. This may not accurately represent SNHL, which has been associated with broadened auditory filters and, as a result, reduced frequency resolution (Glasberg & Moore, 1986). Hopkins et al used young adults with presumably high cognitive function, which might explain why they benefited more from fast compression than slow compression, as compared with older adults who are more likely to have reduced cognitive ability and tend to prefer slow compression (Gatehouse et al. 2006a).

1.9 Summary Of Issues In Current Literature And Rationale

Although compression is useful for providing gain that addresses both elevated thresholds and loudness recruitment, fast compression also introduces distortions that may be problematic, particularly for some listeners when listening in background noise. Most studies which have shown this have used vocoders which reduced or eliminated access to fine structure and simulated a loss in spectral resolution (e.g., Stone & Moore 2003, 2004, 2008), inspiring the hypothesis that the deleterious effects may be due to the lack of access to fine-structure; in the absence of fine-structure listeners have to rely more heavily on speech envelope modulations, which are negatively impacted by compression (Moore, 2008). Similarly, it has been proposed that fast compression might be beneficial because it makes it possible for people with hearing loss to listen in the dips of modulated noise—something that has been shown to be influenced by the presence of fine structure.

However, Hopkins et al. (2012) found no relationship between access to fine structure and the effects of fast compression.

It is possible that the reason that Hopkins et al. (2012) did not find an effect of fine-structure was that the audibility benefit outweighed the deleterious effects of fast compression that are found when audibility is equated across conditions (as in Stone & Moore 2003, 2004, 2008). Perhaps access to fine-structure does not affect the benefit of fast compression (increased audibility) but does reduce the detrimental effects of fast compression when audibility is equated.

The present study tested this by controlling access to fine structure with fast compression but with no simulation of elevated thresholds. The vocoder studies showing problems with fast compression also simulated poorer-than-normal spectral resolution; so reduced spectral resolution was also controlled, by smearing the speech spectrum of the signal in some conditions (Moore & Glasberg, 1990). Given that no deleterious effects of slow compression have been demonstrated, the present study used only fast compression, but compared single-channel compression with multichannel compression in order to partially distinguish temporal distortion (single-channel compression) from combined spectrotemporal distortion (multichannel compression). A non-compressed condition was also used. A single speaker, reversed-speech masker was used to maximize opportunities for listening in the dips. Reversed speech had the benefit of containing the same spectral content as the target and being non-intelligible so it did not provide informational masking.

The study also tested several indices that might be useful for estimating the impact of fast-compression on intelligibility. An objective index could be a useful tool for minimizing distortions that may negatively impact intelligibility during the hearing aid fitting process. One of the tested indices was based on comodulation of the signal and the background noise, which could impair intelligibility. Stone and Moore (2008) proposed a measure of this, called across-signal modulation correlation (ASMC), as described in Section 1.5. Compression also causes a reduction of modulation depth over time (for single- and multi-channel fast compression) and a distortion of spectral shape (for multichannel fast compression) as differences in level across frequency are reduced. Stone and Moore (2008) suggested that increased correlation of modulations across frequency might improve auditory object formation and be helpful for speech understanding and proposed a formula for measuring this, called within-signal modulation correlation (WSMC). However, increased correlation across frequencies is also a reduction in spectral information, something that is likely important for speech intelligibility. Indeed, the authors found that increases in WSMC were associated with poorer intelligibility (Stone & Moore, 2008), suggesting that the loss of information outweighs any benefits of comodulation. Stilp, Kiefte, Alexander, and Kluender (2010) have proposed a measure of spectral information called Cochlea-scaled Spectral Entropy (CSE) that might be useful for indexing this information and the loss of information that might occur with multichannel compression.

In the present study, both ASMC and CSE were calculated for all 240 sentences at four signal-to-noise ratios and the average scores were compared to measured intelligibility thresholds in order to determine whether they might be suitable for further development as clinical measures of compression-related distortion.

Chapter 2: Methods

2.1 Listeners

Ten normal-hearing listeners (eight female) were tested. All listeners were university students with audiometric thresholds within normal limits (< 25 dB HL) at frequencies between 250 and 8000 Hz bilaterally. Thresholds averaged across subjects are illustrated in Figure 2.1. Ages ranged from 24 to 27 years. All were native speakers of English. None of the participants had previous experience with the Hearing In Noise Test (HINT) sentences. Participants attended one session and were not compensated for their time.

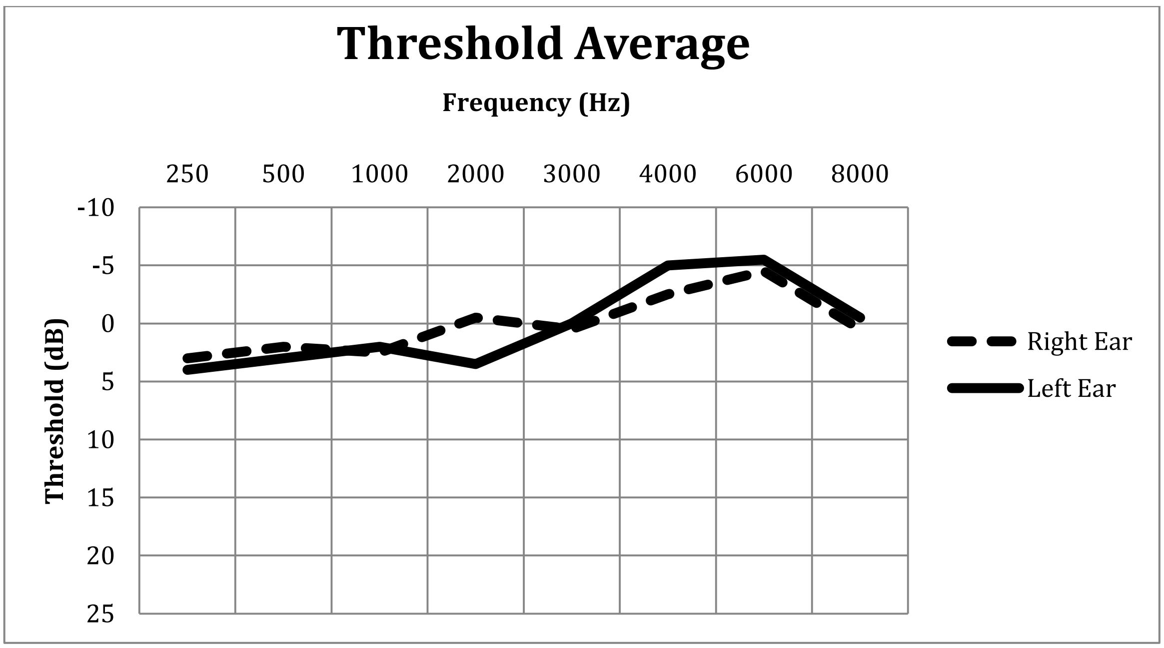

Figure 2.1 Average hearing thresholds for right and left ear across subjects. Thresholds fall within normal limits.

2.2 Stimuli

Speech intelligibility testing was completed using sentences from the Hearing in Noise Test (HINT, Nilsson et al., 1994) processed to simulate the effects of single and multichannel compression, as well as spectral smearing and loss of access to fine-structure. Signal to noise ratio (SNR) thresholds were measured for each condition. Twenty-four HINT lists were used, each containing 10 sentences, for a total of 240 sentences. Two lists (20 sentences) were used for each condition. Each HINT sentence was reversed to create a competing-talker noise source. The fundamental frequencies of the reversed HINT sentences were shifted upwards by 20% to facilitate separation of speech and noise (i.e., the competing speech).

The processing of the stimuli is illustrated in a flow chart in Figure 2.2. All target and competing sentences were processed using STRAIGHT (Kawahara, Mroise, Takahashi, Nisimura, 2008) software in MATLAB 2014a (The Mathworks, Natick, MA). This software decomposes speech into source and filter components. In the current study, the filter information was decomposed into spectra calculated with 21.5 Hz resolution and updated in one-millisecond increments. The decomposition was processor intensive and was thus completed for all sentences and reversed (competing) sentences in advance of the experiment. All subsequent processing was completed online, based on the signal-to-noise ratios at each step in the experiment. The online threshold measurement is described below in the Procedure section.

For the unprocessed condition, the sentence and noise were simply resynthesized using the STRAIGHT algorithm and summed for presentation. Target sentences were resynthesized using the original source track obtained during the STRAIGHT decomposition, and competing sentences were resynthesized using a reversed source track that was shifted upwards in frequency by 20%.

The compression simulation was implemented in STRAIGHT in single-channel and multichannel versions. First, the intensities of the signal and noise (i.e., competing reversed speech) were summed at the appropriate signal-to-noise ratio (variable during threshold assessment) to estimate the level that would be detected by a hearing aid compressor. In the single-channel condition, this produced a single intensity estimate for each millisecond. For the multichannel condition, a similar intensity estimate was calculated in 18 one-third octave bands with nominal centre frequencies from 160 to 8000 Hz. The highest band included all information up to the upper limit of the stimulus recording (10 kHz). The band separation was achieved by mapping the 21.5 Hz spectral slices onto the approximate cut-off frequencies of the one-third octave band filters.

The 5-millisecond attack time and 100-millisecond release time were simulated by convolving the level estimate with a finite-impulse response filter. The filter was composed of four zeroes followed by a step to maximum value and a linear declination to zero over 100 samples. The filter coefficients summed to unity. The resultant filtered level estimate was thus delayed by 5 ms and slowly released over 100 ms. The threshold kneepoint was simulated by calculating the 30th percentile of the filtered level estimate in each frequency channel (for multichannel compression) or for the entire signal (for single channel compression). This estimate was selected based on the approximate correspondence between the 30th percentile of the speech level distribution for average speech and the kneepoint setting in modern hearing aids (e.g., 40-55 dB SPL in the high and low frequencies respectively). The level estimate was then used to calculate a gain adjustment value based on a 3:1 compression ratio, which was subtracted from the level of the signal and the noise. The minimum gain adjustment value was always zero for summed signal+noise levels at or below the threshold kneepoint. Note that the gain adjustment value for each channel (or for the single-channel) was always equal for the signal and the noise, since both would always be processed through the same hearing aid (i.e., the hearing aid must have one gain value at each point in time for each channel based on the total level estimate for that channel).

The smearing simulation, which simulates broadened auditory filters, was imposed after the simulation of compression because decreased auditory resolution must have an impact after a signal is processed through a hearing aid. The smearing was accomplished by multiplying the signal and noise log spectrograms (i.e., the 21.5 Hz / 1 millisecond 'filter' components obtained via decomposition in STRAIGHT) by a smearing matrix. The matrix was comprised of triangles centered at each 21.5 Hz frequency step, with a width equal to four times the width of the equivalent rectangular bandwidth at each frequency. This was calculated using the formula of Moore and Glasberg (1990):

$$ERB(f) = 24.7(4.37f + 1)$$

The loss of fine-structure was simulated by replacing the STRAIGHT source track with a series of zeroes, for both the signal and the competing speech.

In all cases, target and competing sentences were resynthesized and summed at the appropriate signal-to-noise ratio and then presented to the listener.

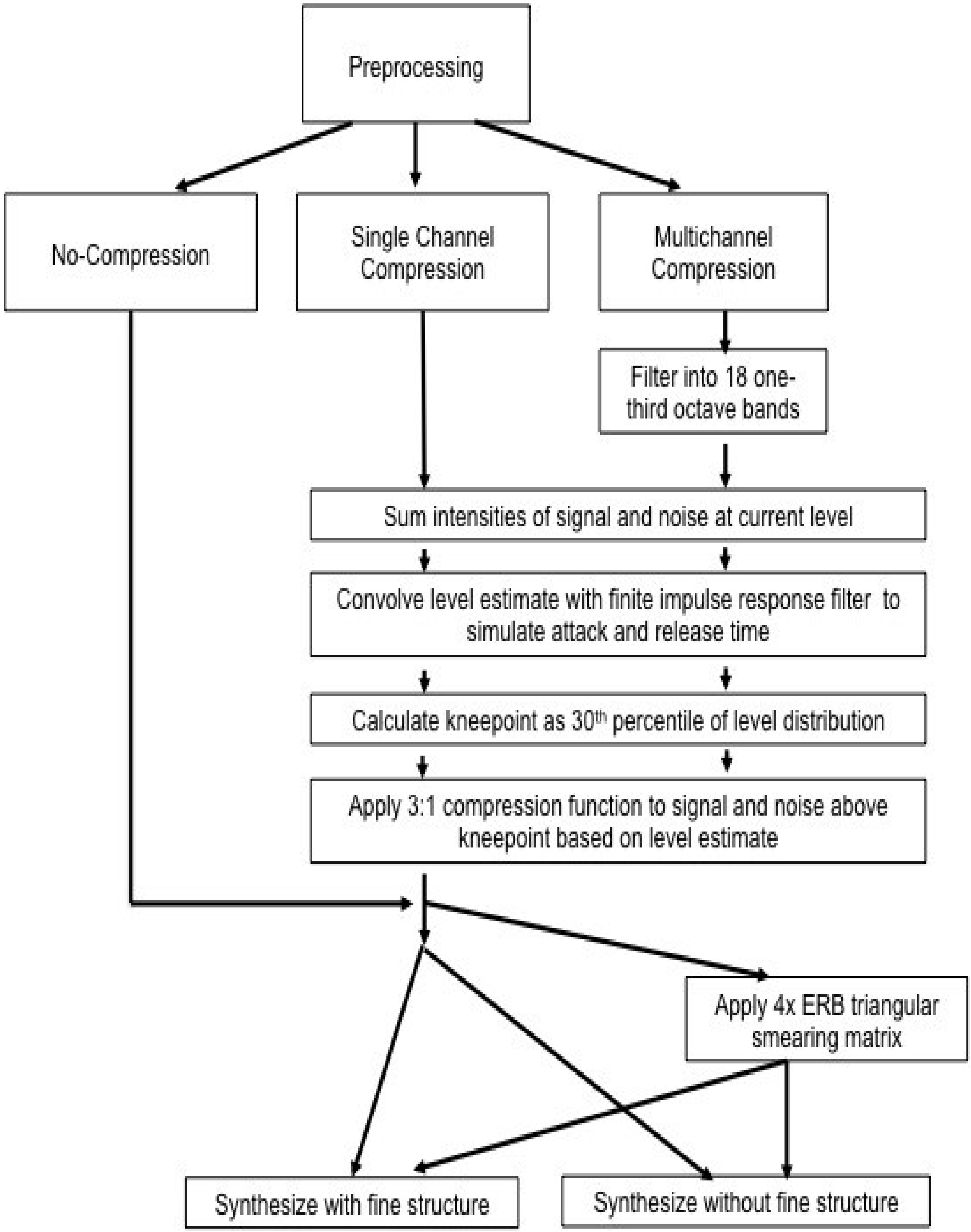

Figure 2.2 Flowchart illustrating processing of stimuli

2.3 Conditions

Twelve conditions were tested. There were four stimulus versions: spectrally smeared with and without fine-structure, and non-smeared with and without fine-structure. Within each stimulus version there were three compression schemes: fast multi-channel (18 channels), fast single-channel, and no compression.

2.4 Equipment and Procedure

The participants attended one session, lasting roughly one and a half hours. Audiometric thresholds were first measured for octave frequencies 250 – 8000 Hz using a GSI-61 audiometer and ER-3A-type insert earphones. The 12 test conditions were presented in randomized order.

All stimuli were presented using MATLAB (The Mathworks, Natick, MA), routed through a GSI61 audiometer to ER-3A-type insert earphones in a double-walled sound proof booth. Sentences were presented at a level of 65 dB SPL. Competing sentences were presented initially at a level that was 12 dB above the level of the sentences. This was reduced in 4 dB steps until the first sentence was repeated correctly. For the first 3 sentences, competing speech levels were adjusted in 4 dB steps, with an increase in the competing speech level for each correct answer and a decrease for each incorrect answer. For sentences 4 to 20, the levels were adjusted in 2 dB steps. The threshold was the mean value of the final 16 sentences as well as the level that would be used for the 21st sentence.

No training was provided for the task. Two 10-sentence lists were used for each of the twelve conditions, with each pair of lists randomly selected (without replacement) from a total set of 24 HINT lists.

2.4.1 Calculation Of ASMC And CSE

ASMC and CSE were calculated for all 240 test sentences at 4 SNRs, -12 dB, -6 dB, 0 dB, and 6 dB. Both indices were calculated for each of the following conditions: no compression, single-channel compression, multichannel compression, no compression smeared, single-channel compression smeared, and multichannel compression smeared. ASMC and CSE were not calculated for the conditions with removed fine-structure because the indices are primarily based on the distribution of levels across frequency and time (i.e., the speech envelope).

The calculation of ASMC is described in Stone and Moore (2007). This was calculated using the formula:

$$ASMC = \frac{1}{N}\sum_{i=1}^{N} C(a_i b_i)$$

Where a_i represents the log envelope of the target signal post-compression, $\text{Target}_{\text{postcomp}}$ and b_i represents the log envelope of the background signal post-compression, $\text{Background}_{\text{postcomp}}$, in the *i*th channel.

The procedure for calculating CSE is described in Stilp et al. (2010). The relative entropy of the spectrum across time was operationalized as changes in cochlea-scaled spectral Euclidean distance. The analysis consisted of normalizing with a root-mean-square (RMS) calculation, then dividing the signal into 16-ms slices. A 66-point FFT was used to capture the magnitude spectrum of each slice, and then rounded-exponential (ROEX) filters were used to simulate the cochlea's frequency distribution. Thirty-three ROEX filters were applied to the magnitude spectra and they were expressed as functions of ERB rate.. Euclidean distances were then calculated between each slice of the target sentence and all other slices, and then averaged for each sentence.

Chapter 3: Results

The main effects of compression, smearing and fine-structure were assessed with a within subject, repeated measures analysis of variance (ANOVA). The results are expressed in average SNR threshold.

3.1 Main Effects And Interactions

A main effect of compression was found [$F(2,18) = 12.747$, $p = 0.000$], and is shown in Figure 3.1. This shows the mean HINT thresholds for each condition expressed as signal-to-noise ratios in dB. Lower values indicate better performance. Post-hoc analyses indicate that thresholds for single channel compression were not significantly worse than for the no compression condition, but thresholds for multichannel compression were significantly poorer than thresholds for single-channel compression [$t(9) = -3.2521$, $p = 0.01$] and for the no compression condition [$t(9) = -5.1765$, $p = 0.0005821$].

There was a significant effect of removing the fine-structure [$F(1,9) = 82.875$, $p = 0.000$] as well as smearing [$F(1,9) = 86.584$, $p = 0.000$], and a significant interaction was found between fine-structure and smearing [$F(1,9) = 25.089$, $p = 0.001$], as illustrated in Figure 3.2. However, there were no interactions between compression and fine-structure or spectral smearing, nor was there a three-way interaction, indicating that the negative effects of multichannel compression on intelligibility were independent of both fine-structure and smearing.

Post-hoc analyses were used to investigate interactions between fine-structure and smearing, using paired t-tests with Bonferonni corrections to control family-wise error rate. The analyses showed that the differences between all ordered pairs were significant;

the smeared condition with absent fine-structure yielded worse performance than the condition with absent fine-structure but no smearing [$t(9) = 3.4189$, $p = 0.0076$], which was worse than the condition with smearing but intact fine-structure [$t(9) = -4.0692$, $p = 0.0028$], and this was worse than unprocessed speech [$t(9) = -12.95$, $p = 0.0000$].

3.2 ASMC And CSE

Across Signal Modulation Correlation (ASMC) and Cochlea-Scaled Spectral Entropy (CSE) were calculated for the 240 test sentences at four signal-to-noise ratios:
–12 dB, –6 dB, 0 dB, and +6 dB. The ASMC data are plotted in Figure 3.3 The scores are averaged across smeared and unprocessed sentences for each compression condition since there was no interaction between smearing and compression. The values of the ASMC scores are negative because there is a negative correlation between the modulations of the signal and the background noise; compression causes the level of the noise and the peaks to be reduced during the peaks in the signal, so the noise is at a low level during the peak levels in the signal. The negative correlation should be highest when the speech and noise are a similar level, at 0 dB SNR, and decrease with changes in either direction in SNR. This pattern was found with ASMC scores for the sentences in this study in the single channel condition but not in the multichannel condition. ASMC was also less negative for multichannel compression—a condition that was associated with significantly poorer intelligibility.

CSE data are plotted in Figure 3.4. CSE is not a correlation; it is a measure of spectral change and should be correlated with speech intelligibility. A higher value indicates more spectral change or information. As shown in Figure 3.4, there is a general pattern of higher CSE with better SNR up to 0 dB SNR. CSE scores are better for single-channel than multichannel compression, and best for the no compression condition, which parallels the pattern in speech intelligibility in the three conditions. Figure 3.5 shows CSE and ASMC scores at 0 dB SNR as well as average thresholds for the 3 compression conditions: no compression, single-channel, and multichannel compression. The y-axis of the average thresholds is expressed in descending order in order to make them more easily comparable to CSE scores. The two bottom panels of Figure 3.5 depicting CSE and SNR thresholds show that CSE declines with intelligibility.

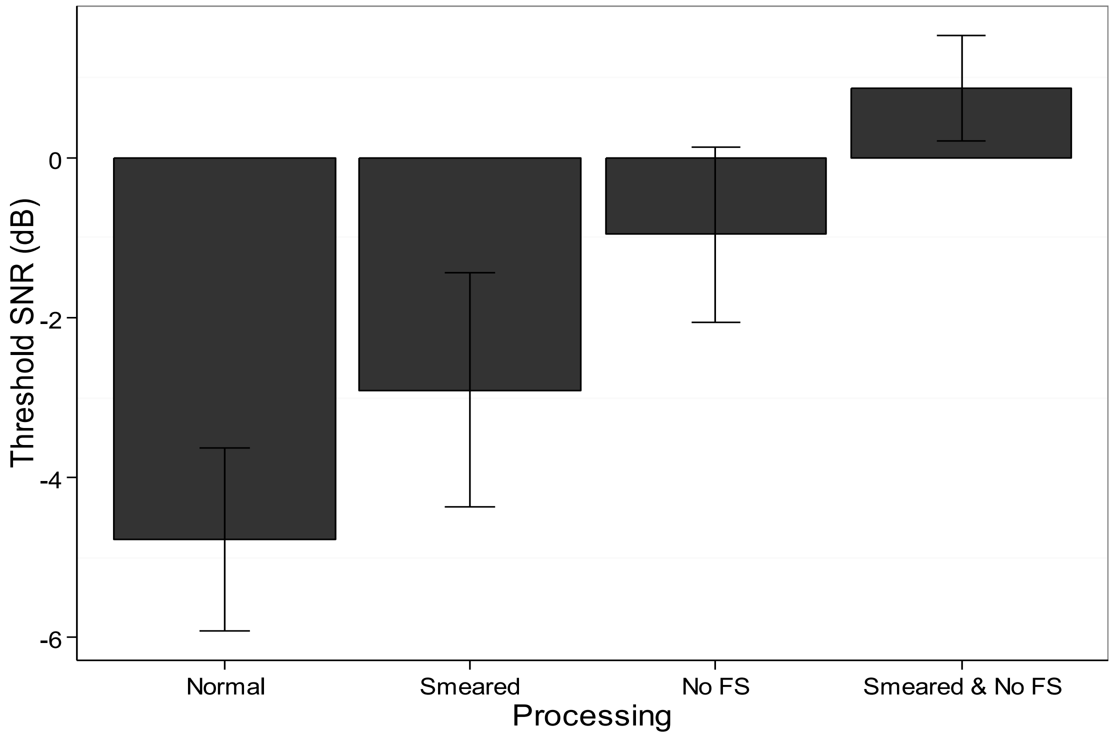

Figure 3.1. Main effect of compression on speech intelligibility, measured in signal to noise ratios (SNR). Error bars show 95% confidence intervals.

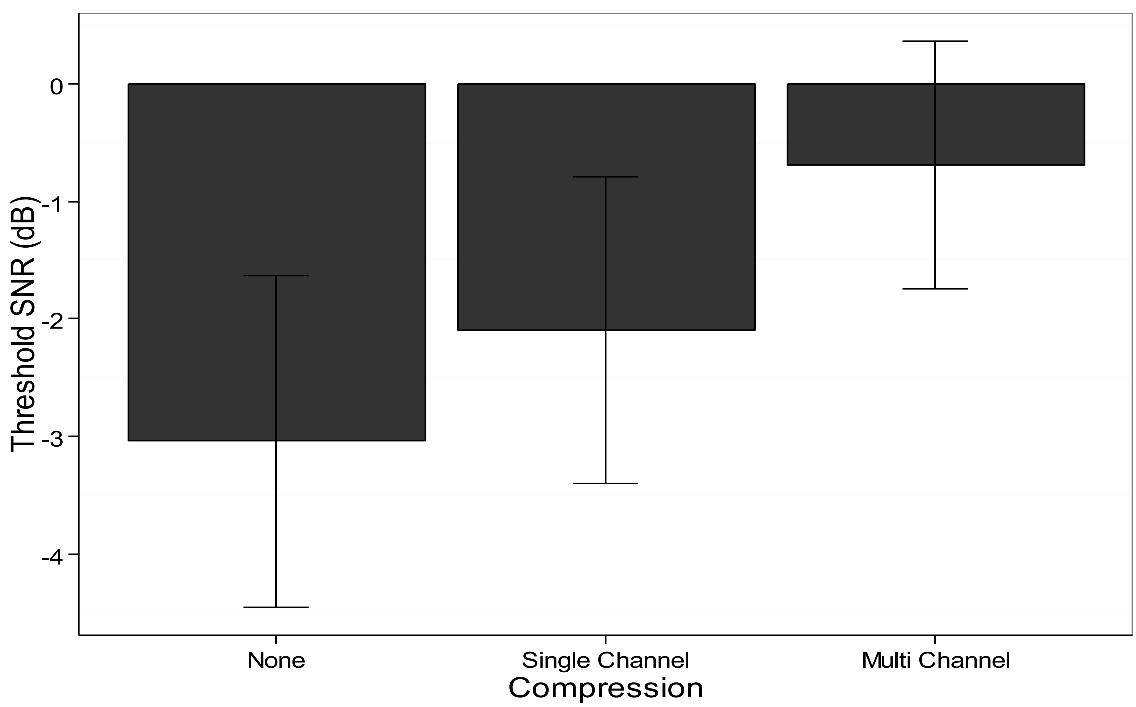

Figure 3.2. Interaction between smearing and removing fine-structure. Performance was worst in the condition with both smearing and removed fine-structure. Error bars show 95% confidence intervals.

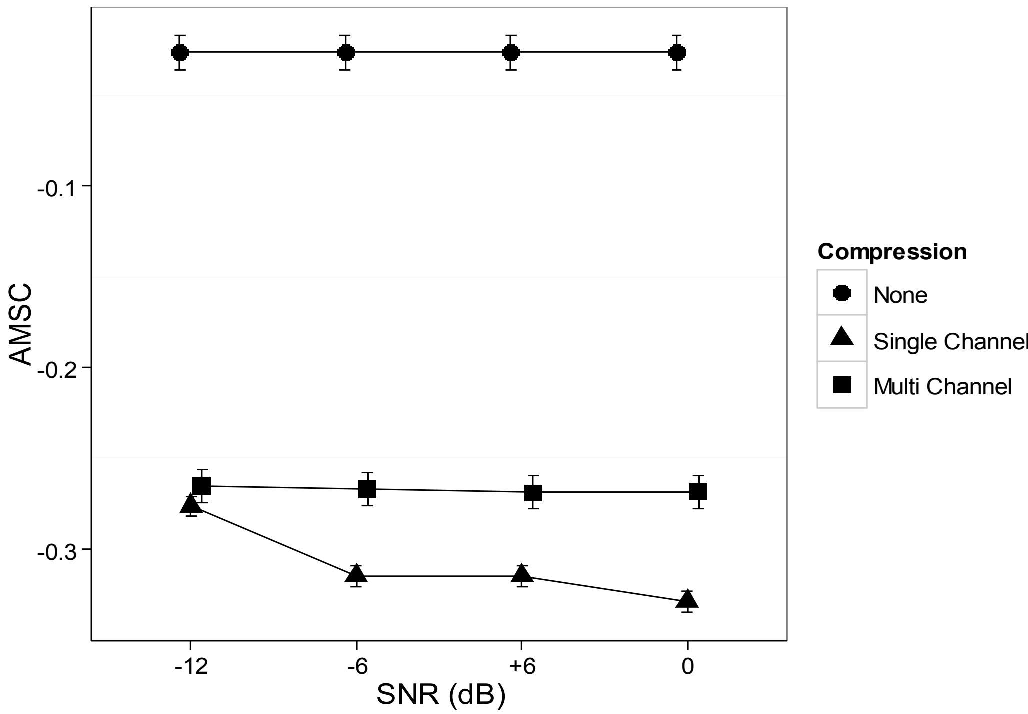

Figure 3.3. Average ASMC scores for all compression conditions, collapsed across smearing. Error bars show 95% confidence intervals.

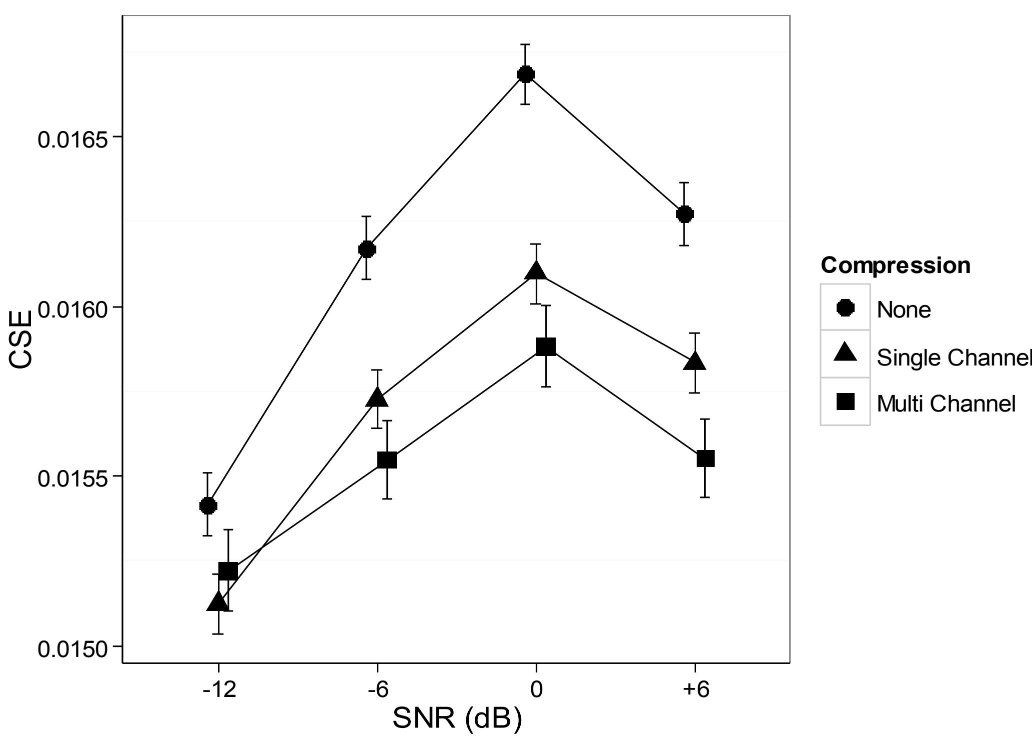

Figure 3.4. Average CSE scores for all compression conditions, collapsed across smearing. Error bars show 95% confidence intervals.

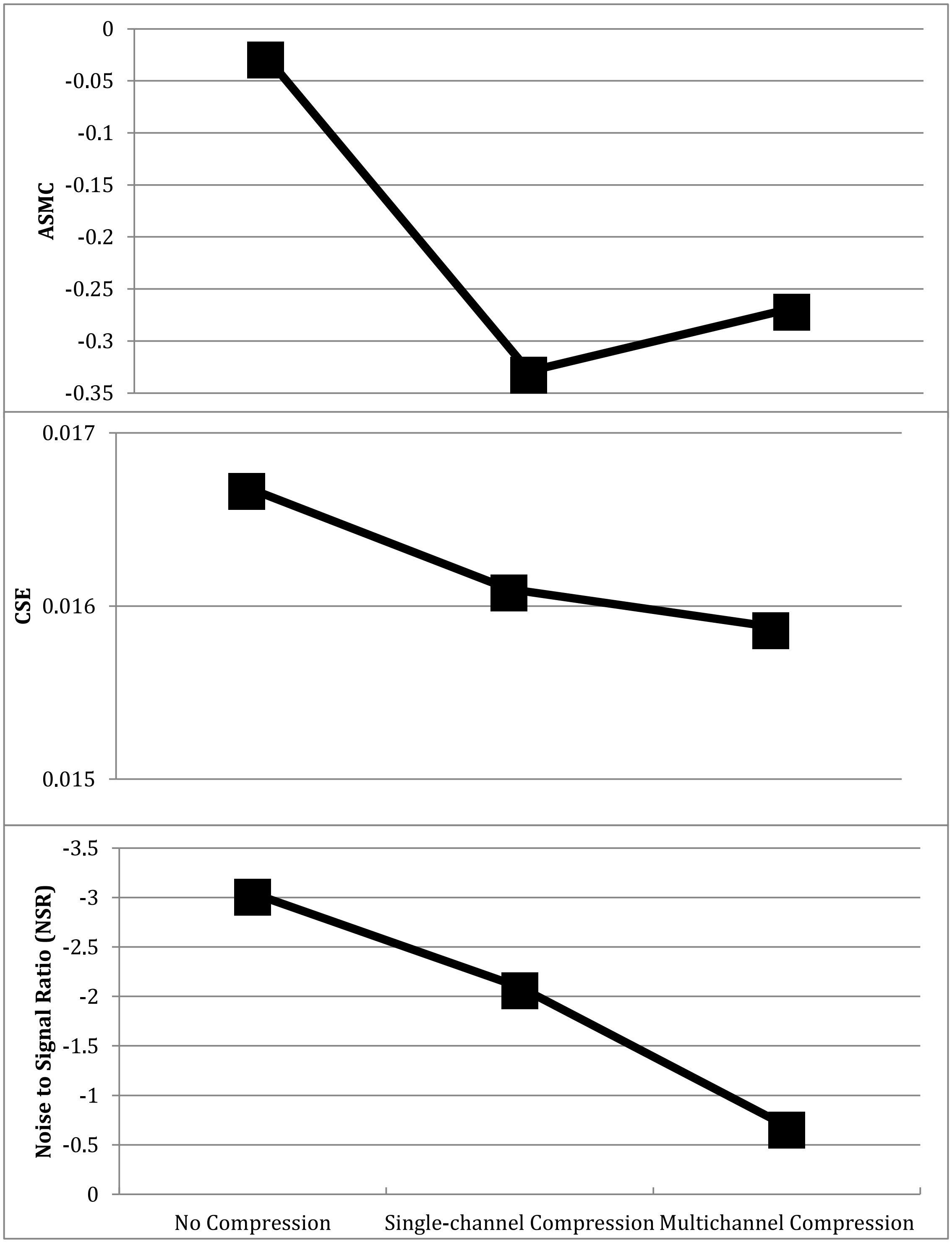

Figure 3.5. Panels from top to bottom showing ASMC at 0 dB SNR, CSE at 0 dB SNR, and threshold signal-to-noise ratios (SNR) for 3 compression conditions. Note that y axes differ for each panel. The y-axis for average thresholds in the bottom panel is presented in descending order so that it can be easily compared with CSE.

Chapter 4: Discussion

4.1 Summary Of Results

The results of this study show that fast, multi-channel compression has a negative impact on speech intelligibility in noise, while single channel compression does not. It was anticipated that the effect of compression would be worse when the stimuli were smeared and fine-structure removed, but the impact of compression was independent of spectral resolution and the presence of fine-structure, i.e., there was no significant interaction between compression and fine-structure. A significant interaction was found between fine-structure and smearing: performance was worst in conditions with smeared stimuli and absent fine-structure.

4.2 Comparing Results To Relevant Studies

The main goal of this study was to investigate the distortion associated with multichannel versus single-channel compression and to assess the importance of fine-structure and spectral resolution in coping with the distortion. The results of this study support those found by Hopkins et al. (2012), who reported that the availability of fine-structure was not related to the effects of compression. Hopkins et al. (2012) used tone vocoding, so some fine-structure information was still available in their stimuli (albeit artificial). This was considered a potential limiting factor in their study. However, noise vocoding was used in the present study to completely remove fine-structure from the signal, and similar results were found. Also, the simulation in Hopkins et al. (2012) included reduced audibility and abnormal growth of loudness, such that fast compression was found to be beneficial. Therefore, the study did not directly address the question of whether availability of fine-structure might be important for reducing the deleterious effects of fast-compression in conditions where fast compression does not improve audibility. The present study tested this and found that fine-structure availability did not affect the negative impact of fast multichannel compression on intelligibility. Therefore, the results of this study do not support Moore's (2008) contention that access to fine-structure mitigates the negative effects of fast compression, and that fine-structure processing tests might provide valuable information concerning optimum compression speed. The results of the current study suggest that fast, multichannel compression could have adverse effects on speech understanding in noise for any hearing aid user, regardless of fine-structure availability, just as the results of Hopkins et al., (2012) showed that fast multichannel compression can have beneficial effects on speech understanding regardless of fine-structure availability. Patterns in performance across conditions reveal that access to fine-structure cues does not help listeners compensate for the effects of fast, multichannel compression. The effect of smearing was also independent of compression, which is in agreement with the results of Hopkins et al. (2012).

These results are not in agreement with Stone et al. (2009), which found that intelligibility suffered with increasing compression channels and speed, but that the effect was smaller with increased number of processing channels (i.e., better resolution) and when some form of speech fine-structure was provided (i.e., access to high-rate envelope cues). The authors suggested that high spectral resolution in combination with high-rate envelope cues, which are due to interactions of fine-structure components (speech harmonics) in the vocoder channels, can improve intelligibility with multichannel compression. The reasons for the discrepancy are not clear. Stone et al. (2009) used a different method of reducing fine-structure than the current study, a tone vocoder with varying numbers of channels, and they varied the lower frequency of the low pass filtering on the envelope to control availability of pitch-related cues (i.e., the periodicity envelope). Thus, their study compared conditions where fine-structure was absent to conditions where a derivative of the fine-structure was present. This is less of a contrast than in the present study, which compared fine-structure completely absent to fine-structure completely present, so any effect found in Stone and Moore (2009) would be expected to be present, and perhaps larger, in the present study. Other aspects of the study were similar (e.g., they used a compression ratio just below 3:1, and had compression speeds that were faster and slower than in the present study) so cannot easily account for the differences.

One possible explanation for the discrepancy is that their listeners had more experience with the vocoded stimuli; they underwent an hour of training before testing, and were tested for two hours. In the present study there was no training and testing lasted for one and a half hours. It is possible that listeners can learn to make use of subtle fine-structure cues to reduce the negative effects of the fast compression and that this learning had not occurred in the present study.

A more likely possibility is that the results of Stone and Moore (2009) are due to a peculiar artifact of their methodology, because periodicity cues in a channel vocoder are cues that are not present in a natural speech signal. With natural speech, the periodicity envelope is introduced when speech fine-structure components interact on the basilar membrane and arise due to the asymmetric inner hair cell potentials and the rectification-like process of neural transduction, but this occurs only *after* mixing of the speech and noise. In contrast, a tone vocoder's periodicity envelope is introduced explicitly as an amplitude modulation of the carrier tones for the speech alone, in any vocoder channel wide enough to contain two or more harmonics. Therefore, the periodicity is artificially introduced without the presence of noise during the vocoding process. It is possible that comodulating the carrier tones of a channel vocoder at the precise fundamental frequency of the voice provides an artificial segregation cue that greatly facilitates separation of speech and noise, thereby counteracting the comodulation of speech and noise imposed by the compression. In other words, the benefit of the high-rate envelope cues found by Stone and Moore (2009) might be a simple artifact of their methodology. The present results cannot resolve this issue but suggest that Stone and Moore's (2009) findings should be interpreted cautiously, particularly in light of the similar findings in Hopkins et al. (2012).

Unlike multichannel compression, single channel compression did not result in a significant decrease in performance. Other studies have found similarly that fast, multichannel compression has more substantial negative effects on intelligibility in noise than single channel compression (Stone & Moore, 2008). The different types of distortion present in the speech signal with multichannel compression compared to single-channel compression might explain this. While single-channel compression can create some level of temporal distortion, multichannel affects the spectral shape of the signal as well as the temporal envelope.

Only fast-acting compression was used in this study. Although it can improve audibility, some studies suggest that fast-acting compression may have adverse effects on speech understanding in background noise compared to slow compression (Stone & Moore 2003; 2004; 2008; 2009). Similarly to our results, Stone and Moore (2003; 2004;

2008) found a negative effect of fast, multichannel compression with the use of a noise vocoder. However, Hopkins et al. (2012) found better performance with fast compression than slow compression. There is a key difference between the two studies that can explain this difference. Hopkins et al. simulated loudness recruitment and elevated thresholds, while Stone and Moore did not. The advantage that fast-acting compression can offer in improving audibility would not have been relevant in Stone and Moore's studies since no measures were taken to reduce audibility for listeners. In the present study, threshold elevation and recruitment were not simulated, so audibility was also not an issue. Similar to the findings of Stone and Moore, fast multichannel compression hindered performance compared to no compression and single-channel compression. In other words, fast multichannel compression can improve speech intelligibility by increasing audibility when audibility is limited, but it also creates distortion that can reduce speech intelligibility (Souza et al., 2005; Stone & Moore, 2008). The effect of distortion is best shown when audibility is equated across conditions. In the present study, there was no attempt to limit audibility and so no positive effect of fast multichannel compression was found, as expected. The goal of the present study was to measure and attempt to quantify the distortion of compression, not to assess the effects of improved audibility, which are well accounted for by the SII as implemented in hearing aid test equipment. Hopkins et al. (2012) simulated hearing loss and recruitment and thus showed some benefits associated with fast compression.

4.3 Simulating Hearing Loss

Smearing was used in the present study to simulate broadened auditory filters observed in cochlear hearing loss and a noise vocoder was used to completely remove the fine-structure from the signal. The reason for this was the expectation that the subtle effects of compression on the spectro-temporal envelope would be most apparent for individuals with impaired resolution and reduced access to speech fine-structure. It should be noted that the goal of the present study was not to provide a precise and accurate simulation of reduced access to fine-structure but rather to resolve the question in the literature about whether fine-structure plays any role in mitigating the effects of compression. While more subtle manipulations of fine-structure access would likely be more realistic, the absence of an effect would fail to rule out an effect for less subtle manipulations. Also, artificial introductions of fine-structure as in Stone and Moore (2009) could produce effects unrelated to the difference in fine-structure access that occur in real auditory systems.

The simulation of loss of spectral resolution via smearing was designed differently than in typical vocoder studies (e.g., Stone & Moore, 2003) to remove the natural confound between vocoder channel bandwidth for manipulations of fine-structure and for manipulation of spectral resolution, and to avoid the artificial discontinuities between neighbouring channels present in noise-vocoded speech, since these could produce spurious effects on the tested indices (ASMC and CSE). To this end, a smearing matrix was produced that simulated broadened cochlear filters by using a triangularly shaped filter that was four times the width of each equivalent rectangular bandwidth.

As with any simulation of hearing loss, the model may not accurately represent real SNHL and results may vary with different parameters. However, in cases where drastic manipulations fail to have any effect, it is unlikely that more subtle and realistic manipulations would have an effect. In other words, if the drastic removal of all fine-structure and the 4xERB smoothing in the present study failed to show any interaction with the effects of compression, it is unlikely that more realistic but limited reductions in fine-structure access and spectral resolution would interact with the effects of compression. Hearing loss simulations are often used in subjects with normal hearing because they allow researchers to isolate different aspects of hearing loss and understand the role of each more accurately, without confounding interactions. In SNHL, psychophysical variables are interrelated and difficult to test independently. A large amount of individual variability is often found in the performance of individuals with real SNHL (Gatehouse et al., 2006a,b).

Both manipulations impacted speech intelligibility as expected, but neither of the manipulations was found to mitigate the effects of fast multichannel compression.

4.4 ASMC And CSE

The results indicate that there is a greater level of spectral and temporal distortion introduced by multichannel than single-channel compression. The next phase of the research is to quantify this distortion. As discussed in the introduction, the SII as implemented in hearing aid test systems is not sensitive to all the relevant distortions introduced by compression. An index of spectro-temporal distortion that can be used to supplement this version of the SII would be clinically useful when determining optimal compression settings for an individual. Across-source modulation correlation (ASMC)

Stone & Moore, 2007) and cochlea-scaled entropy (CSE) (Stilp et al., 2010) were investigated for this purpose.

CSE is an alternative index for predicting speech intelligibility. It is based on the principle that changes in the acoustic signal over time (entropy) provide important information for speech perception. ASMC measures the (negatively) correlated modulation of two independent sources, such as a speaker and background noise. This occurs when a compressor is applied to a target and background talker with similar levels and can potentially lead to difficulty with perceptual separation of the independent sources. This is most relevant for speech understanding in the presence of a competing speaker. In the work of Stone and Moore (2009), for both single and multi-channel compression, ASMC appeared to be the only measure that corresponded to speech in noise performance.

CSE and ASMC were measured for all of the test sentences to see if they corresponded well to the speech intelligibility scores obtained. Unfortunately, neither of the indices provided a useful index the effects of signal-to-noise ratio on speech intelligibility for the signal and noise used in the present study. The likely reason for this is that the background noise used in this study was very acoustically similar to the target speech. The noise accompanying each sentence was the reversed version of the same sentence, so an intelligibility index based on the spectrotemporal envelope could not differentiate between conditions with positive and negative signal-to-noise ratios. Reversed speech was appealing to use as background noise because it provided deep modulations for listening in background noise, contained the same spectral content as the target signal, and was non-intelligible so did not have any informational masking. A similar problem could exist in real world applications, where an acoustic index would not be able to differentiate between situations where the target speaker is clearly above a competing speaker and where the competing speaker is above the target speaker. However, in clinical test equipment it would be possible to create an index based on the target signal to resolve this issue.

It is possible to compare the effectiveness of the indices at a given SNR. At a 0 dB SNR, CSE showed promise for predicting speech intelligibility in that its value correctly declined as speech intelligibility became worse (i.e., as the speech-in-noise threshold increased). Future research could be conducted to produce a version of the CSE that is sensitive to the differences between the target speech and a typical source of background noise, such as natural speech from a single competing speaker. This could involve measuring both positive and negative information based on whether the information pattern is maximally correlated with the target speech or with the competing speech. The correlations between measured CSE and measure intelligibility would also need to be tested for a wide variety of realistic stimuli.

Other indices could also be tested. One potentially useful index would be the version of the SII that incorporates the Speech Transmission Index. Since the SII is already incorporated into test systems, this might be a relatively simple modification that would improve the accuracy of the SII. One advantage associated with this approach is that audibility and distortion would both be reduced to a single value, so clinicians would not be required to separately maximize one value (an SII) while minimizing another (e.g., a measure of distortion). The disadvantage would be that the reason for a low SII could be more difficult to determine (e.g., a low SII could be because of poor audibility in an important spectral region or because of distortion introduced in the process of trying to maximize audibility).

4.5 Clinical Implications

The findings regarding multichannel compression in the current study have implications for compression settings used in modern hearing aids. For instance, a typical high-end hearing aid with fast-acting compression across 15-20 channels might hinder speech understanding when listening in noise, at least in cases where the aid does not provide adequate benefits in audibility. There is evidence suggesting that using fewer channels might reduce the adverse effect of multichannel compression. Speech-in-noise performance with a small number of channels (up to four channels) has been shown to be comparable to that achieved with single channel compression (Keidser & Grant, 2001).

A limited number of compression channels might also be related to Hopkins et al.'s (2012) finding that fast compression showed better performance. The authors used 6 compression channels, which might help improve audibility with minimum spectrotemporal distortion. In the current study, 18 channels were used in the multichannel condition, and Stone and Moore used up to 12 or 16 in their studies. Table 1 outlines the number of compression channels used by each of the Stone and Moore (2003; 2004; 2008) studies and Hopkins et al. (2012). One of the findings of Stone and Moore (2008) is the trend that speech intelligibility decreased as compression speed and the number of channels increased; fast compression appeared to have less of a negative impact on performance with fewer channels. Given that deleterious effects of fast compression tend to be found in studies using systems with many channels (12-18; e.g., Stone & Moore, 2003; 2004; 2008 and the present study) and benefits of fast compression have been shown in studies using systems with fewer channels (2-6, e.g., Hopkins et al., 2012; Gatehouse et al., 2006), it might make sense to favour hearing aids with smaller numbers of channels than with larger numbers of channels.

One of the goals of this study, as discussed in the introduction, was to quantify distortion introduced by fast multichannel compression in a meaningful way that could be used by clinicians. Future research with CSE could lead to the development of a clinical tool to aid the fitting process. If the index reliably depicts spectral distortion and corresponds to speech intelligibility, a step further would be to process speech through a sample of modern hearing aids and calculate the CSE of the output, and then to measure intelligibility through those same hearing aids. The longer-term goal is to determine whether changes in CSE can reliably be used to predict changes in speech intelligibility. If so, this measure might be a helpful addition to clinical test equipment.

4.6 Conclusions

The results of the current study show that fast multichannel compression negatively impacted speech intelligibility, whereas fast single channel compression did not, and this effect was independent of spectral resolution and access to speech fine-structure. The results also suggested that CSE, but not ASMC, might be sensitive to compression-related spectral distortion, although more research needs to be done to determine whether it can be used to reliably predict speech intelligibility and supplement the SII.

References

Acoustical Society of America., & American National Standards Institute. (1997). Methods for calculation of the speech intelligibility index. New York, N.Y.: Published by the American Institute of Physics for the Acoustical Society of America

AIcantara, J., & Moore, B. C. J. (1995). The identification of vowel-like harmonic complexes: Effects of component phase, level, and fundamental frequency. *Journal of Acoustical Society of America, 97*(6), 3813-3823.

Ardoint, M., Sheft, S., Fleuriot, P., Garnier, S., & Lorenzi, C. (2010). Perception of temporal fi ne-structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss. *International Journal of Audiology, 49*(11), 821-831.

Assman, P., & Summerfield, Q. (1989). Modeling the perception of concurrent vowels: Vowels with the same fundamental frequency. *Journal of Acoustical Society of America, 85*(1), 327-338.

Bacon, S. P., Opie, J. M., & Montoya, D. Y. (1998). The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *Journal of Speech, Language, and Hearing Research, 41*(3), 549-563.

Baer, T., & Moore, B. C. J. (1993). Effects of spectral smearing on the intelligibility of sentences in noise. *Journal of Acoustical Society of America, 94*(3), 1229-1241.

Baer, T., & Moore, B. C. J. (1994). Effects of spectral smearing on the intelligibility of in the presence of interfering speech. *Journal of Acoustical Society of America, 95*(4), 2277-2280.

Bernstein, J. G. W., Summers, V., Grassi, E., & Grant, K. W. (2013). Auditory models of suprathreshold distortion and speech intelligibility in persons with impaired hearing. *Journal of the American Academy of Audiology, 24*(4), 307-328.

Byrne, D., Dillon, H., Ching, T., Katsch, R., & Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. *Journal of the American Academy of Audiology, 12*(1), 37-51.

Crain, T. R., & Van Tasell, D. J. (1994). Effect of peak clipping on speech recognition threshold. *Ear & Hearing, 15*(6), 443-453.

Desloge, J. G., Reed, C. M., Braida, L. D., Perez, Z. D., & Delhorne, L. A. (2010). Speech reception by listeners with real and simulated hearing impairment: Effects of continuous and interrupted noise. *Journal of Acoustical Society of America, 128*(1), 342-359.

Dreisbach, L. E., Leek, M. R., & Lentz, J. J. (2005). Perception of spectral contrast by hearing-impaired listeners. *Journal of Speech, Language, and Hearing Research, 48*, 910-921.

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of reducing slow temporal modulations on speech reception. *Journal of Acoustical Society of America, 95*(5), 2670-2680.

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *Journal of Acoustical Society of America, 95*(2), 1053-1064.

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2002). Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *Journal of Acoustical Society of America, 111*(6), 2897-2907.

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2003). Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing. *Journal of Acoustical Society of America, 113*(4), 2084-2094.

Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *Journal of Acoustical Society of America, 74*(3), 739-743.

Eisenberg, L. S., Dirks, D. D., & Bell, T. S. (1995). Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *Journal of Speech and Hearing Research, 38*, 222-233.

Feng, Y., Yin, S., Kiefte, M., & Wang, J. (2009). Temporal resolution in regions of normal hearing and speech perception in noise for adults with sloping high-frequency hearing loss. *Ear & Hearing, 30*(6), 1-11.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speechreception threshold for impaired and normal hearing. *Journal of Acoustical Society of America, 88*(4), 1725-1736.

French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *Journal of Acoustical Society of America, 19*(1), 90-119.

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal of Acoustical Society of America, 110*(2), 1150-1163.

Gatehouse, S., Naylor, G., & Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Intertiational Journal of Atidiology, 42*(1), S77-S85.

Gatehouse, S., Naylor, G., & Elberling, C. (2006). Linear and nonlinear hearing aid fittings 1. patterns of benefit. *International Journal of Audiology, 45*(3), 130-152.

Gatehouse, S., Naylor, G., & Elberling, C. (2006). Linear and nonlinear hearing aid fittings 2. patterns of candidature. *International Journal of Audiology, 45*(3), 153-171.

George, E. L. J., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *Journal of Acoustical Society of America, 120*(4), 2295-2311.

Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *Journal of Acoustical Society of America, 79*(4), 1020-1033.

Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear & Hearing, 23*(4), 369-380.

Hideki Kawahara, Masanori Morise, Toru Takahashi, Ryuichi Nisimura, Toshio Irino, Hideki Banno, Tandem-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation, Proc. ICASSP 2008, Las Vegas,pp.3933-3936(2008).

Hopkins, K., King, A., & Moore, B. C. J. (2012). The effect of compression speed on intelligibility: simulated hearing-aid processing with and without original temporal fine structure information. *Journal of Acoustical Society of America, 132*(3), 1592-1601.

Hopkins, K., & Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *Journal of Acoustical Society of America, 125*(1), 442-446.

Hopkins, K., & Moore, B. C. J. (2010). Development of a fast method for measuring sensitivity to temporal fi ne structure information at low frequencies. *International Journal of Audiology, 49*(12), 940-946.

Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *Journal of Acoustical Society of America, 130*(1), 334-349.

Hopkins, K., Moore, B. C. J., & Stone, M. A. (2008). Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *Journal of Acoustical Society of America, 123*(2), 1140-1153.

Ito, M., Tsuchida, J., & Yano, M. (2001). On the effectiveness of whole spectral shape for vowel perception. *Journal of Acoustical Society of America, 110*(2), 1141-1149.

Keidser, G., & Grant, F. (2001). The preferred number of channels (one, two, or four) in NAL-NL1 prescribed wide dynamic range compression (WDRC) devices. *Ear & Hearing, 22*(6), 516-527.

Kiefte, M., Enright, T., & Marshall, L. (2010). The role of formant amplitude in the perception of õiÕ and õuÕ. *Journal of Acoustical Society of America, 127*(4), 2611-2621.

Leek, M. R., Dorman, M. F., & Summerfield, Q. (1987). Minimum spectral contrast for vowel identification by normalhearing and hearing-impaired listeners. *Journal of Acoustical Society of America, 81*(1), 148-154.

Leger, A. C., Moore, B. C. J., Gnansia, D., & Lorenzi, C. (2012). Effects of spectral smearing on the identification of speech in noise filtered into low- and mid-frequency regions. *Journal of Acoustical Society of America, 131*(5), 4114-4123.

Leger, A. C., Moore, B. C. J., & Lorenzi, C. (2012). Abnormal speech processing in frequency regions where absolute thresholds are normal for listeners with high-frequency hearing loss. *Hearing Research, 294*, 95-103.

Leger, A. C., Moore, B. C. J., & Lorenzi, C. (2012). Temporal and spectral masking release in low- and mid-frequency regions for normal-hearing and hearing-impaired listeners. *Journal of Acoustical Society of America, 131*(2), 1502-1514.

Liu, C., & Eddins, D. A. (2008). Effects of spectral modulation filtering on vowel identification. *Journal of Acoustical Society of America, 124*(3), 1704-1715.

Lorenzi, C., Debruille, L., Garnier, S., Fleuriot, P., & Moore, B. C. J. (2009). Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal (L). *Journal of Acoustical Society of America, 125*(1), 27-30.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences, 103*(49), 18866-18869.

Lorenzi, C., Husson, M., Ardoint, M., & Debruille, X. (2006). Speech masking release in listeners with flat hearing loss: Effects of masker fluctuation rate on identification scores and phonetic feature reception. *International Journal of Audiology, 45*(9), 487-495.

Lunner, T., Hietkamp, R. K., Andersen, M. R., Hopkins, K., & Moore, B. C. J. (2012). Effect of speech material on the benefit of temporal fine structure information in speech for young normal-hearing and older hearing-impaired participants. *Ear & Hearing, 33*(3), 377-388.

Middelweerd, M. J., Festen, J. M., & Plomp, R. (1990). Difficulties with speech intelligibility in noise in spite of a normal pure-tone audiogram. *Audiology, 29*(1), 1-7.

Moore, B. C. J. (2008). The choice of compression speed in hearing aids: Theoretical and practical considerations and the role of individual differences. *Trends in Amplification, 12*(2), 103-112.

Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research and Otolaryngology, 9*(4), 399-406.

Moore, B. C. J., & Glasberg, B. R. (1986). Comparisons of frequency selectivity in simultaneous and forward masking for subjects with unilateral cochlear impairments. *Journal of Acoustical Society of America, 80*(1), 93-107.

Moore, B. C. J., Glasberg, B. R., Flanagan, H. J., & Adams, J. (2006). Frequency discrimination of complex tones; assessing the role of component resolvability and temporal fine structure. *Journal of Acoustical Society of America, 119*(1), 480-490.

Moore, B. C. J., Glasberg, B. R., & Hopkins, K. (2006). Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hearing Research, 222*(1-2), 16-27.

Moore, B. C. J., Peters, R. W., & Stone, M. A. (1999). Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. *Journal of Acoustical Society of America, 1*(400), 411.

Nearey, T. M. (1989). Static, dynamic, and relational properties in vowel perception. *Journal of Acoustical Society of America, 85*(5), 2088-2113.

Nelson, P. B., Jin, S., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *Journal of Acoustical Society of America, 113*(2), 961-968.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *Journal of Acoustical Society of America, 95*(2), 1085-1099.

Peters, R. W., Moore, B. C. J., & Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *Journal of Acoustical Society of America, 103*(1), 577-587.

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *Journal of Acoustical Society of America, 83*(6), 2322-2327.

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function.*83*(6), 2322-2327.

Plyler, P. N., Reber, M. B., Kovach, A., Galloway, E., & Humphrey, E. (2013). Comparison of multichannel wide dynamic range compression and ChannelFree processing in open canal hearing instruments. *Journal of American Academy of Audiology, 24*(2), 126-137.

Schvartz, K. C., Chatterjee, M., & Gordon-Salant, S. (2008). Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. *Journal of Acoustical Society of America, 124*(6), 3972-3988.

Seewald, R., Moodie, S., Scollie, S., & Bagatto, M. (2005). The DSL method for pediatric hearing instrument fitting: historical perspective and current issues. *Trends in Amplification, 9*(4), 145-157.

Shannon, R. V., Zeng, F., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*(5234), 303-304.

Souza, P. E., & Boike, K. T. (2006). Combining temporal-envelope cues across channels: Effects of age and hearing loss. *Journal of Speech, Language, and Hearing Research, 49*(1), 138-149.

Souza, P. E., Jenstad, L. M., & Folino, R. (2005). Using multichannel wide-dynamic range compression in severely hearing-impaired listeners: Effects on speech recognition and quality. *Ear & Hearing, 26*(2), 120-131.

Stilp, C. E., Kiefte, M., Alexander, J. M., & Kluender, K. R. (2010). Cochlea-scaled spectral entropy predicts rate-invariant intelligibility of temporally distorted sentences. *Journal of Acoustical Society of America, 128*(4), 2112-2126.

Stone, M. A., Füllgrabe, C., & Moore, B. C. J. (2009). High-rate envelope information in many channels provides resistance to reduction of speech intelligibility produced by multi-channel fast-acting compression (L). *Journal of Acoustical Society of America, 126*(5), 2155-2158.

Stone, M. A., & Moore, B. C. J. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *Journal of Acoustical Society of America, 114*(2), 1023-1034.

Stone, M. A., & Moore, B. C. J. (2004). Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task. *Journal of Acoustical Society of America, 116*(4), 2311-2323.

Stone, M. A., & Moore, B. C. J. (2007). Quantifying the effects of fast-acting compression on the envelope of speech. *Journal of Acoustical Society of America, 121*(3), 1654-1664.

Stone, M. A., & Moore, B. C. J. (2008). Effects of spectro-temporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task. *Journal of Acoustical Society of America, 123*(2), 1063-1076.

Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *Journal of Acoustical Society of America, 125*(5), 3328-3345.

Syrdal, A. K., & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of american english vowels. *Journal of Acoustical Society of America, 79*(4), 1086-1100.

Takahashi, G. A., & Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech and Hearing Research, 35*(6), 1410-1421.

ter Keurs, M., Festen, J. M., & Plomp, R. (1993). Effect of spectral envelope smearing on speech reception. II. *Journal of Acoustical Society of America, 93*(3), 1547-1552.

Thibodeau, L. M. (1991). Exploration of factors beyond audibility that may influence speech recognition. *Ear & Hearing, 12*(6), 109S-115S.

Turner, C. W., Chi, S., & Flock, S. (1999). Limiting spectural resolution in speech for listeners with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research, 42*(4), 773-784.