Perceived Quality of Recorded Music Processed through Compression Hearing Aids

by

Naomi B.H. Croghan

B.A., University of Northern Colorado, 2005

Au.D., University of Colorado, 2010

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Speech, Language, and Hearing Sciences

2013

This thesis entitled:
Perceived Quality of Recorded Music Processed through Compression Hearing Aids
written by Naomi B.H. Croghan
has been approved for the Department of Speech, Language, and Hearing Sciences

Kathryn H. Arehart

James M. Kates

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB protocol #11-0133

Croghan, Naomi B.H. (Ph.D., Speech, Language, and Hearing Sciences)

Perceived Quality of Recorded Music Processed through Compression Hearing Aids

Thesis directed by Associate Professor Kathryn H. Arehart

Music perception with hearing aids has recently gained attention as a clinical and scientific problem. The goal of this dissertation was to examine the roles of input-signal properties, hearing-aid processing, and individual variability in the perception of recorded music. Two experiments considered music-industry compression limiting (CL) and hearing-aid wide dynamic-range compression (WDRC). The specific aims were: 1) to characterize the acoustic effects of CL on recorded music and to determine how listeners with normal hearing perceive recorded music processed by CL, 2) to quantify the acoustic and perceptual effects of CL, WDRC, and CL+WDRC on recorded music processed through simulated hearing aids, and 3) to examine the relationship between preferred compression parameters and individual perceptual characteristics of listeners with hearing loss.

Experiment 1 investigated the effects of CL on classical and rock music. An acoustic analysis indicated reduced fidelity with increasing compression. When loudness varied due to compression, normal-hearing listeners preferred mild CL, on average. When loudness was equalized, mild compression did not affect music quality, but heavier compression was less preferred. Responses varied among individual listeners.

Experiment 2 evaluated CL+WDRC using hearing-aid simulations and listeners with hearing loss. Acoustically, WDRC diminished amplitude contrasts and upset spectral balance, particularly for conditions with more channels (18 vs. 3) and a faster release time (50-msec vs. 1000-msec). Perceptually, linear processing and slow WDRC were equally preferred over fast WDRC for classical music. For rock music, linear processing was preferred to both slow and fast WDRC. The main effect of channels was not significant for classical music, but for rock, 3-channel processing was preferred. CL degraded music quality for classical but not for rock. For classical music, listeners

with broader estimated auditory filters preferred 3 channels and linear processing, while listeners with narrower filters preferred 18 channels and slow WDRC. For rock music, the degree of preference for linear processing was greater for listeners with a larger dynamic range.

The present findings have highlighted specific acoustic modifications that contribute to music quality and have identified compression parameters that are perceptually relevant for both music-industry recording and hearing-aid design.

Dedication

In memory of my Gramps, who sparked a love of science and showed me the value of education.

Acknowledgements

My heartfelt gratitude goes to my advisor and mentor, Kathy Arehart, who inspired me to push myself even in the most daunting of tasks, and whose unending support made it possible for me to do so. I also thank my mentor, Jim Kates, who challenged me to think in new ways, and my committee members, Lew Harvey, Phillip Gilley, and Anu Sharma. I value their efforts, encouragement, and diverse perspectives throughout this process. I am grateful to several outside researchers for providing helpful input, and I would especially like to acknowledge Pam Souza and Marshall Chasin for their constructive comments.

I would not have made it without my lab mates and fellow student researchers (current and former!), who were generous with their time, expertise, and friendship – especially Ramesh Muralimanohar, Melinda Anderson, Cory Portnuff, In-Ki Jin, Katie Dreher, Mallene Wiggin, and Paula Messamer. I also sincerely thank my colleagues at Cochlear who showed me incredible support and patience while finishing this project. I am fortunate to be a member of your team.

I acknowledge my participants, whose contributions of time and energy made this research possible, and the financial support provided by GN ReSound.

Above all, I am grateful to my family and friends, including my Au.D. cohort, who listened to my struggles and shared in my successes. Mom, Dad, Jeani, "G.G.," and Grandma Dee - thanks for believing in me. Kevin, I could not have asked for a better partner by my side. You have been my relief and my advocate, and we made it through this journey together.

Contents

Tables

Table

Figures

Figure

Chapter 1

Introduction

According to the National Institutes of Health, approximately 36 million Americans have some degree of hearing loss (NIDCD, 2010). The standard treatment for many types of hearing loss is hearing aids, yet only about 25% of people with hearing loss actually choose to adopt hearing aids (Kochkin, 2012). To identify factors that might encourage hearing-aid use, Kochkin (2012) recently outlined the key determinants that would most likely influence a person to get hearing aids for the first time. On the list of 53 choices, the factor "music sounds better through hearing aid" ranked above "job performance suffers," "recommendation from spouse," "wireless connection to cell phone," and "friend has good hearing aid experience." The more general factor "much better sound quality" was 10th on the list. The advancement of music perception with hearing aids is becoming increasingly recognized as both a clinical and scientific problem. Many hearing-aid manufacturers now include music programs for their hearing instruments, but the settings included in those programs are variable and are often proprietary. Recommendations for hearing-aid music fittings have been reported in the clinical literature (Chasin, 2003, 2006, 2010; Stender, 2012), but laboratory research to complement this clinical experience is largely lacking.

Music perception with hearing aids is a highly complex problem. When a person listens to music through hearing aids, at least five separate components are involved (Figure 1.1, A-E). First, the acoustic features of music pose a challenge to hearing-aid processing. Hearing aids are designed for speech, and acoustic differences between speech and music may require different approaches in hearing-aid design (Figure 1.1, A). Additionally, music may be either live or recorded. For recorded music, signal processing applied in the audio industry may affect the acoustic properties and sound quality of the music prior to the input stage of the hearing aid (Figure 1.1, B). During the hearing-aid processing stage (Figure 1.1, C), multiple signal-processing algorithms act upon the music in a complex manner. The optimal combination of hearing-aid parameters for music is presently unknown. Following hearing-aid processing, the modified music signal arrives at an abnormal auditory system (Figure 1.1, D). Sensorineural hearing loss leads to multiple perceptual changes, including reduced audibility as well as disrupted suprathreshold sound processing (Carney and Nelson, 1983; Moore and Glasberg, 1997). Finally, the signal is perceived by an individual with specific characteristics, such as age, gender, and musical background (Figure 1.1, E). Differences in individual factors might contribute to variability in perceived music quality.

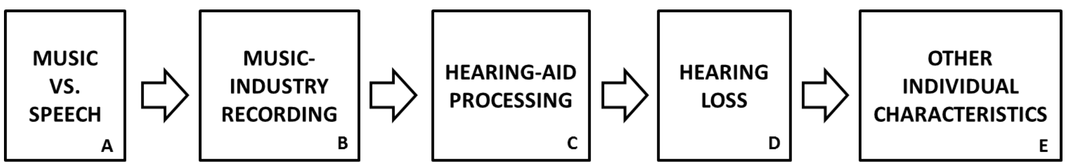

Figure 1.1: The perception of recorded music with hearing aids is a complex problem involving at least five components: the acoustic properties of music (and differences between music and speech), the effects of the music-industry recording process, the effects of hearing-aid digital signal processing, the effects of hearing loss, and the other individual characteristics of the listener.

To investigate the problem of music perception with hearing aids, research is needed not only to understand what hearing-aid parameters are best for music but also to understand the factors that influence individual preferences. The goal of this research was to help explain the contributions and interactions of the five factors illustrated in Figure 1.1, including the complexities that underlie individual variability and patterns of benefit. The methodology was to select one group of variables related to a specific form of hearing-aid processing that is important to current hearing-aid design, wide dynamic-range compression (WDRC). The goal was to perform an in-depth analysis of the effects of WDRC on music perception and to relate these effects to signal properties, hearing loss, and individual characteristics. In the following sections, the five elements of music perception with hearing aids will be described in the context of the literature related to WDRC.

1.0.1 Music perception with hearing aids

1.0.1.1 Component 1: Music vs. speech

Hearing-aid research has primarily focused on improving speech understanding, with less emphasis placed on music perception. Because hearing aids are constructed with speech in mind, hearing-aid processing is often mismatched to the acoustic properties of music. Compared to speech, music has a wider dynamic range, defined as the difference between the most intense and least intense points of the signal. The dynamic range for speech is 30-35 dB; the dynamic range for music can be up to 100 dB (Chasin, 2006). Dynamic-range on a shorter time scale is described as crest factor (the ratio between the most intense points of the signal and the average intensity). The crest factor for speech is 12 dB, and the crest factor for music is 18-20 dB (Chasin, 2006). Live music also has a greater intensity level than speech (Chasin, 2010, 2003). Yelled speech is 80-85 dB SPL, while both rock and classical music sometimes reach 100-110 dB SPL (Chasin, 2010, 2003). In temporal and spectral properties, music is more variable than speech, as the spectrum and modulation rate of music depend on the instrumentation (Chasin and Russo, 2004) and tempo (Scheirer, 1998; Todd et al., 1999). Due to these differences between speech and music, current hearing-aid design may not process music signals optimally.

1.0.1.2 Component 2: Music-industry recording

For hearing-aid users listening to recorded music, an additional factor to consider is the music-industry recording process. Commercial recording methods entail a number of signal-processing stages, including mixing and mastering. Mixing involves combining multiple tracks into a single stereo file, including balancing of different instruments and using digital effects to achieve a particular sound (Bregitzer, 2009). Mastering is the final step of audio production prior to distribution. It may involve technical enhancement of the audio as well as sequencing and editing of an album as a whole (Katz, 2007). One type of processing often used during mastering is compression limiting (CL), which produces a relatively constant output level at a high compression ratio for all sounds above the compression threshold (Kefauver, 2001). By increasing the root-mean-square (RMS) level, CL increases the loudness of music. A common perspective in the music industry has been that “louder is better” (Vickers, 2011). Industry professionals have released music with increasing compression over time, attempting to gain advantage over their competition by producing the loudest music on the market. This phenomenon has been called the “loudness war.” Despite the trend of increasing compression, many artists and consumers have argued that using large amounts of CL negatively affects the quality of music by removing the emotional impact of dynamics and by causing distortion. Limited evidence has been available to support either the contention that more heavily compressed music sells better or the contrary viewpoint that listeners prefer less compression. The current study examined perceptual aspects of CL in listeners with normal hearing to determine the extent to which music-industry CL affects music quality.

The use of music-industry CL may also influence how hearing-aid users perceive music. Like normal-hearing listeners, people with hearing loss may be sensitive to distortions caused by CL. Moreover, a recorded music signal that is processed by CL prior to the input stage of a hearing aid will be further modified by hearing-aid digital signal processing. An important processing feature included in nearly all modern hearing aids is WDRC. A person listening to recorded music through hearing aids could, therefore, experience two-fold effects of compression: the effects of CL on music quality and the effects of WDRC on music quality, along with potential interactions between these two types of processing. This study included several configurations of CL, WDRC, and CL+WDRC to measure the separate and combined effects of compression for listeners with hearing loss.

1.0.1.3 Component 3: Wide Dynamic Range Compression (WDRC)

WDRC is a nonlinear processing strategy that reduces the amount of gain with increasing signal levels, across a wide range of input levels. The primary motivation for using WDRC is the diminished dynamic range that accompanies cochlear hearing loss (Moore and Glasberg, 1997). The effects of WDRC have been studied widely in terms of speech recognition and speech quality. Substantially less research has been performed to determine optimal WDRC parameters for music.

Nonetheless, insights may be gained from the literature on WDRC and speech perception to identify the effects of WDRC that might be important for music. Therefore, the following sections will discuss some effects of WDRC on speech and then introduce preliminary evidence related to WDRC and music.

WDRC and speech perception

While WDRC may provide audibility for a wide range of sounds, it also causes acoustic modifications to the speech signal, and these distortions may be detrimental for some listeners (Davies-Venn et al., 2009). Acoustic effects of WDRC on speech have been demonstrated in both the spectral and temporal domains. For example, WDRC has been shown to flatten the spectrum of speech sounds. Reduced contrast between higher and lower spectral regions affects cues important for formant perception and leads to diminished vowel recognition (Bor et al., 2008; Souza et al., 2012). WDRC also causes temporal envelope changes to speech signals by decreasing the distinction between peaks and valleys (Souza and Turner, 1998; Jenstad and Souza, 2005). Additionally, WDRC affects the modulation spectra of consonant sounds. The modulation spectrum describes the depth of modulation for various modulation frequencies, providing a measurement of the primary rates of amplitude variation contained within a signal. Souza and Gallun (2010) showed that WDRC modified the modulation spectra of individual phonemes, causing the phonemes to be more dissimilar to one another. However, this increased heterogeneity did not improve consonant recognition for listeners with hearing loss (Souza and Gallun, 2010).

The acoustic modifications of a speech signal that are caused by WDRC are complex and, at different times, may be advantageous or deleterious to speech intelligibility and sound quality. Therefore, the literature shows mixed results regarding the benefits of WDRC for speech perception (Souza, 2002). One explanation for these mixed results is that WDRC algorithms produce a trade-off between audibility and distortion: the benefits derived from improved audibility must be balanced with detrimental effects of distortion of the speech signal (Kates, 2010).

WDRC and music perception

The relationship between audibility and distortion may also affect music perception with

WDRC. The few studies examining WDRC for music perception have yielded a range of results, possibly due to differences in the combinations of parameters being studied. For example, Davies-Venn et al. (2007) showed that WDRC gave higher perceived quality for classical and popular music, compared to two other strategies that only affected high-level sounds (peak clipping and limiting). However, Higgins et al. (2012) found that a modified linear strategy called adaptive dynamic range optimization (ADRO) was preferred to WDRC for classical, rock, and jazz music. For these two studies, differences in the implementation of WDRC settings make it difficult to draw conclusions about the benefits of WDRC for music. One approach to determine how WDRC should be used for music is to conduct a parametric analysis of WDRC settings. Using simulated hearing aids, the current study formed systematic combinations of WDRC parameters to determine the acoustic and perceptual outcomes associated with WDRC and music.

1.0.1.4 Component 4: Hearing loss

Responses to hearing-aid processed sounds are often highly variable. Individual patterns of benefit could be explained, in part, by the effects of hearing loss. Hearing loss may affect music perception with hearing aids in a multitude of ways. The most apparent consequence is a loss of audibility, leading to portions of the spectrum being imperceptible. The benefits of WDRC for speech intelligibility differ based on the degree (Souza and Bishop, 1999) and configuration (Souza and Bishop, 2000) of hearing loss. Likewise, listeners with different auditory thresholds might have different responses to WDRC for music. For example, the harmonic at 1500 Hz is important for the timbre of a clarinet, but the harmonic at 6000 Hz is important for a violin (Chasin and Russo, 2004). Listeners' auditory thresholds might affect how they perceive different instruments.

Hearing loss is also associated with suprathreshold processing deficits that influence several dimensions of music perception. In a study measuring fundamental frequency (F0) discrimination, Arehart (1994) showed that discrimination of complex-tone pitch was impaired for most listeners with hearing loss, although substantial variability was seen. As a group, subjects with hearing loss relied on more heavily on temporal cues for F0 discrimination than normal-hearing listeners. This finding is consistent with the idea that poorer frequency selectivity accounts for disrupted pitch perception in hearing loss, due to limited access to resolved harmonics (Oxenham, 2008). Abnormal pitch perception may affect how listeners with hearing loss perceive melody. Accordingly, de Laat (1985) found that hearing-impaired subjects had significantly more difficulty recognizing competing melodies than listeners with normal hearing.

Recent evidence also shows that hearing loss affects the perception of consonance and dissonance, which is important for the portrayal of tonality in music. Tufts et al. (2005) demonstrated that listeners with hearing loss perceived less of a distinction between consonant and dissonant musical intervals. Musical instrumentation is another important feature of music. Different musical instruments provide differences in timbre. In some cases, hearing loss affects timbre perception, although individual variability is present. In an experiment measuring just noticeable differences (JNDs) between musical instruments that were "morphed" along a continuum, Emiroglu and Kollmeier (2008) reported that some listeners with hearing loss had larger JNDs than normal-hearing listeners, whereas other hearing-impaired listeners performed closer to normal. The variability was explained in terms of hearing configuration; listeners with flat hearing loss tended to have low JNDs, while listeners with steeply sloping hearing loss had more difficulty distinguishing instruments. Based on this research relating hearing loss and perception of musical stimuli, the current study included measurements of hearing thresholds as well as several psychoacoustic tasks.

1.0.1.5 Component 5: Individual characteristics

Music perception with hearing aids might also depend on personal factors beyond the effects of hearing loss. Regardless of music-industry and hearing-aid processing, music is ultimately perceived by an individual with specific experiences and preferences. Several personal characteristics are potentially related to a listener's response to music. For example, musical training can enhance functional and neural auditory processing abilities (Parbery-Clark et al., 2011, 2012; Skoe and Kraus, 2012). This enhanced auditory processing may, in turn, influence sensitivity to alterations caused by dynamic-range compression. Other research has shown differences in music responses based on gender, with males preferring louder music to females (Kellaris and Rice, 1993). Age may also play a role in listeners' responses, with possible generational differences existing for preferred genres. To assess the contributions of individual characteristics to perceived sound quality, musical background, gender, and age were considered in this study.

1.0.2 Specific aims

The goal of this dissertation was to examine the five major components involved in the perception of recorded music with hearing aids: the acoustic properties of music, the effects of the music-industry recording process, the effects of hearing-aid digital signal processing, the effects of hearing loss, and the individual characteristics of the listener. The five factors were evaluated within the context of dynamic-range compression, an important type of processing that is relevant to both music-industry recording and hearing-aid design.

The specific aims of the study were:

(1) To characterize the acoustic effects of CL on recorded music and to determine how listeners with normal hearing perceive recorded music processed by CL.

(2) To quantify the acoustic and perceptual effects of CL, WDRC, and CL+WDRC on recorded music processed through simulated hearing aids.

(3) To examine the relationship between preferred compression parameters and individual perceptual characteristics of listeners with hearing loss.

The study was implemented through two experiments. Each experiment included an acoustic analysis, perceptual music quality judgments, and an investigation of individual variability. Experiment 1 addressed Aim 1 by examining the effects of varied amounts of CL on the perceived loudness and sound quality of rock and classical music with normally hearing listeners. The purpose of Experiment 1 was to determine baseline outcomes for CL and to develop a subset of conditions for the second experiment. The results of Experiment 1 are presented in Chapter 2. Experiment 2 addressed Aims 2 and 3 by evaluating the effects of various combinations of CL, WDRC, and

CL+WDRC for music preferences using simulated hearing aids and listeners with hearing loss. The results of Experiment 2 are presented in Chapter 3. Finally, Chapter 4 provides a general discussion of all experimental findings.

Chapter 2

Experiment 1: Music-Industry Dynamic-Range Compression

2.1 Introduction

2.1.1 Compression and music recording

[1] Dynamic-range compression is employed for practical and aesthetic purposes during the process of recording music (Kefauver, 2001; Bregitzer, 2009). Compression functions by reducing the short-term and long-term dynamic range of a signal. The long-term dynamic range of music refers to the difference between the lowest levels and the peak levels of the signal. The short-term dynamic range of music is called the crest factor, defined as the ratio of the peak amplitude to the root-mean-square (RMS) amplitude. Compression can alter both the dynamic range and the crest factor of recorded music by decreasing differences between the peaks and the average level and differences between the peaks and the lowest level in the music (Schmidt and Rutledge, 1996).

Mixing and mastering engineers typically apply compression at several stages and for a number of reasons during the recording process. For example, compression can function as an artistic effect to emphasize or de-emphasize particular aspects of a recording, such as the attack of an instrument or different levels between a vocalist's verse and chorus (Bregitzer, 2009). The reduction in dynamic range that results from compression may also be helpful when listening to music in background sounds, such as while in a car, due to increased audibility of low-level components of the music. Additionally, compression is used to fit recorded music into the limitations of the

[1] The work presented in Chapter 2 includes material previously published by the Acoustical Society of America. The original article was re-formatted to conform to the current specifications. See Croghan et al. (2012) for the full citation.

recorded storage medium (Kefauver, 2001). In today's popular music recording practices, one of the final stages of the mastering process is to apply a specific type of dynamic-range compression called compression limiting (CL) (Bregitzer, 2009), which produces a relatively constant output level at a high compression ratio for all sounds above the compression threshold (Kefauver, 2001). CL simultaneously reduces the dynamic variation of music while increasing its RMS level, with the intention of increasing the loudness of the recording. A common perspective in the music industry has been that "louder is better" (Vickers, 2011), and previous work shows that listeners prefer processing that increases loudness, under certain conditions (Maempel and Gawlik, 2009). As such, industry professionals have released music with increasing compression over time, attempting to gain advantage over their competition by producing the loudest music on the market. Referred to as the "loudness war," this trend of heightening compression is characterized by a reduction in dynamic range and an increase in RMS levels in commercial music (Katz, 2007; Vickers, 2011).

Despite the "loudness war" among members of the music industry, limited evidence is available to support the contention that more heavily compressed music sells better. Furthermore, many artists and consumers have argued that using large amounts of CL negatively affects the quality of music in several ways. First, dynamics help make music interesting, and a loss of dynamic range may diminish music's emotional impact. Bhatara et al. (2011) demonstrated that listeners rated emotionality as lower when piano performances were manipulated to have reduced amplitude variation compared to the original performance. These findings support the statements made by industry participants who feel that over-compression has caused music to sound cluttered and less exciting (Katz, 2007). Second, many music lovers argue that heavy degrees of CL have harmful effects on the clarity and sound quality of music (Katz, 2007), although direct experimental support of these arguments is limited. In summary, the use of CL has spurred opposing viewpoints: the assertion that more compressed, louder music sounds and sells better, and a contrary perspective that heavy compression is detrimental to the quality of commercially available music.

Regardless of the public forum of the compression debate, empirical research in this area is lacking. Several studies in the audio engineering literature have investigated how data compression (such as that used to generate MP3 files) affects music perception. For example, Pras et al. (2009) demonstrated that listeners preferred coding using higher bit rates to coding using lower bit rates. This effect depended on the genre of music and the expertise of the listener, such that sound engineers preferred the higher-quality version more often than musicians, and the higher-quality version was preferred more often for electronic music than for acoustic music (Pras et al., 2009). Similar types of studies should explore how CL affects listeners' perceptions of recorded music.

When discussing the possible ramifications of music-industry CL, insights may be gained from previous studies investigating the effects of dynamic-range compression on speech perception. For example, Stone et al. (2009) demonstrated that normal-hearing listeners expended increased cognitive effort with increasing compression when attempting to separate speech signals from two different talkers. Speech effects have also been shown for wide-dynamic range compression (WDRC) used in hearing aids. WDRC is intended to compensate for the reduced dynamic range and abnormal loudness growth that may accompany cochlear hearing loss (Moore and Glasberg, 1997). WDRC changes the acoustic properties of speech in number of ways, such as reducing the contrasts between peaks and valleys in the temporal envelope (Souza and Turner, 1998; Jenstad and Souza, 2005) and altering the modulation spectrum (Souza and Gallun, 2010). The acoustic modifications of a speech signal that are caused by WDRC are complex and, at different times, may be advantageous or deleterious to speech intelligibility and sound quality. Therefore, the literature shows mixed results regarding the benefits of WDRC (Souza, 2002). WDRC algorithms produce a trade-off between audibility and distortion: the benefits derived from improved audibility must be balanced with possibly detrimental effects of distortion of the speech signal (Kates, 2010).

The current study sought to examine a potentially similar tradeoff between loudness and distortion in recorded music processed by music-industry CL. The aims of the study were to characterize the acoustic effects of CL on recorded music and to determine how listeners with normal hearing perceive recorded music processed by CL, especially regarding patterns of loudness and sound quality.

2.1.2 Sound quality and loudness

The questions of interest were investigated by obtaining judgments of recorded music samples that were processed using different levels of CL. Within each block of judgments, the amount of CL was parametrically varied and applied to a single music sample. In one set of conditions, CL processing led to large variations in loudness across stimuli, with higher levels of CL resulting in louder music samples. This condition represented a listening environment similar to the radio or a randomized playlist on a portable listening device. In these types of real-world listening, some pieces of music may be more compressed than others, with the intention of maximizing the loudness of each musical selection.

Although loudness is the driving factor for the use of CL, loudness may not be the only consideration for a listener's music preferences. For example, compression also causes distortion (Kates, 2010), and this distortion may affect perceived quality, particularly when loudness differences are minimized. Hearing-aid compression has been shown to degrade music quality under certain conditions, such as high compression ratios (van Buuren et al., 1999; Arehart et al., 2011). Similarly, music-industry CL may introduce distortion that influences sound quality. In order to distinguish between the effects of loudness and distortion, this study included a second condition that minimized loudness differences to isolate the role of distortion in the perception of quality.

Sound quality for music is complex and multi-dimensional. Gabrielsson and Sjogren (1979) reviewed a large number of sound quality dimensions in a series of experiments using music and speech played over loudspeakers, headphones, and hearing aids. They identified eight main dimensions of sound quality: clearness/distinctness, sharpness/hardness-softness, brightness-darkness, fullness-thinness, feeling of space, nearness, disturbing sounds, and loudness. Clearness/distinctness was related to sounds that were rated as pleasant, natural/true-to-nature, and pure. Pleasantness had a complex pattern, in that sounds were sometimes rated as "too" clear, and therefore less pleasant. Although the Gabrielsson and Sjogren (1979) study was conducted using conditions that primarily altered the frequency spectrum of the music, the results suggest that a number of different aspects of sound quality may apply to dynamic-range compressed music. In the present study, listeners were asked to judge loudness, dynamic range, pleasantness, and preference for the music samples. Due to the nature of the "loudness war" debate, the loudness scale was included to directly examine the effects of CL processing on overall perceived loudness. The dynamic range scale was intended to measure whether listeners are sensitive to the changes in dynamic range caused by CL. In this study, dynamic range was defined as "loudness changes," as described previously by Neuhoff (2001) and Olsen et al. (2010). Pleasantness and preference scales were included separately due to the possibility that a listener may judge a particular musical sound as less pleasant, but may still prefer that sound (e.g. distortion on an electric guitar). Pleasantness and preference have both been used to investigate hearing-aid dynamic-range compression (Neuman et al., 1998; Davies-Venn et al., 2007; Moore et al., 2011). The preference scale was included to obtain an overall impression from the listener, taking into account the cumulative effect of all dimensions.

In addition to considering multiple dimensions of sound quality, music perception may depend on factors specific to each listener. For instance, previous experience with musical training or sound engineering may modulate listeners' perceptions of CL. Musical training has been associated with enhanced speech-in-noise perception (Parbery-Clark et al., 2009), auditory discrimination (Koelsch et al., 1999; Tervaniemi et al., 2005; Rammsayer and Altenmuller, 2006), auditory attention (Strait and Kraus, 2011) and analytic listening (Oxenham et al., 2003). Therefore, musically trained subjects may show different preferences or greater sensitivity to changes caused by compression than non-musicians. Additionally, listeners with sound engineering experience may be able to detect smaller differences in sound processing than musically trained listeners with no sound engineering experience (Pras et al., 2009). Many listening tests within the audio engineering literature include only expert listeners in order to reduce associated time and cost. However, other evidence suggests that although trained sound engineers provide more precise judgments of sound quality, broad preferences between trained listeners and the general public may be quite similar (Olive, 2003). To investigate possible differences based on listener experience, this study included subjects with musical training (some of whom also had audio recording training) and subjects with little to no experience in music or sound engineering.

2.2 Method

2.2.1 Subjects

Subjects included 24 normal-hearing adults (11 females and 13 males). Normal hearing was defined as thresholds of 20 dB HL or better across audiometric frequencies from 250-8000 Hz, bilaterally (except for two subjects who had a threshold of 25 dB HL at 8000 Hz in one ear only). All subjects were native speakers of American English. To consider possible effects of musical training, subjects were divided into musician and non-musician groups. Groups were determined by a survey that asked participants about their musical backgrounds, including current and past musical training and practice, participation in school and community musical ensembles, academic music classes and degrees awarded, music teaching experience, conducting experience, and the ability to read musical notation. The survey was adapted from a questionnaire previously used by Parbery-Clark et al. (2011) by adding questions about audio recording experience and listening habits (hours per week spent listening to live and recorded music, genres of music that were typical for listening). The musician group included 14 listeners (ages 18-40), who had at least seven years of musical training on their primary instrument or voice (range: 7 to 25 years) and were self-described musicians. Many of the musicians were trained on multiple instruments. At the time of the study, twelve of the musicians played an instrument at least once a week. Two of the musicians were not currently practicing weekly, but had been awarded a college degree in music performance. Two of the musicians reported having absolute pitch. Three listeners, all within the musician group, reported having audio recording experience. The non-musician group included 10 listeners (ages 23-60), who were self-described non-musicians. Nine of the ten non-musicians had five or fewer years of musical training (range: 0 to 5 years), which took place at least nine years prior to participation in the study. One of the non-musicians had seven years of training (group and self-taught), which began at age 49 and was suspended two years prior to participation.

2.2.2 Stimuli

2.2.2.1 Unprocessed recordings

Two samples of recorded stereo files were used, one rock recording and one classical recording. The rock sample was from "Anything At All," performed by Mere. The classical sample was from "Overture to the Magic Flute," written by W.A. Mozart and performed by the University Symphony Orchestra at the University of Colorado-Boulder. Both pieces were obtained directly from the recording engineers in their un-mastered forms so that no limiting had been applied prior to stimulus processing. Samples of approximately 13 seconds in duration were selected from the recordings, at a point consistent with musical phrasing.

2.2.2.2 Compression processing

The samples were peak-normalized to -4 dB full scale (dBFS) using Adobe Audition 3.0. Bregitzer (2009) recommends setting peak levels to between -6 dBFS and -3 dBFS prior to mastering to allow headroom for digital processing. The -4 dBFS peak normalization performed in this study is within the recommended range. Additionally, the peak normalization step ensured that the selected compression thresholds were below the peak of each waveform. Peak-normalized files were processed using the Massey L2007 Mastering Limiter plug-in for Pro Tools 9. The L2007 is a look-ahead brickwall digital limiter with an infinite compression ratio. The look-ahead feature allows the limiter to detect and compress peaks using very fast attack times. Attack times for peak-detector limiters may be on the order of 25 microseconds (Kefauver, 2001), and for look-ahead systems, attack times are often described as effectively instantaneous (close to 0 microseconds) (Katz, 2007). The threshold was varied, and default settings were used for the other parameters (mode, release, and max output). Specifically, the mode was set to "loud," which is the industry-standard setting for the envelope detector. The release was set to "normal," which is the fastest release setting and is recommended by the manufacturer for most applications. The max output was set to 0 dBFS (all stimuli were tested for clipping, and no clipping was present).

Six threshold conditions were created: no compression (unprocessed), -8 dBFS, -12 dBFS, -16 dBFS, -20 dBFS, and -24 dBFS. The -8 dBFS threshold represented the least compression, and the -24 dBFS threshold represented the most compression. In a digital limiter, makeup gain is automatically applied as the threshold is reduced (Bregitzer, 2009). Therefore, the stimuli with lower thresholds received more gain, resulting in higher RMS values. The RMS values for the rock stimuli ranged from -20.8 dBFS (unprocessed) to -5.4 dBFS (most compression). For comparison, an analysis of five commercial rock/pop recordings from 2006-2011 showed RMS values ranging from -8.2 to -5.7 dBFS (average -7.2 dBFS). The RMS values for the classical stimuli ranged from -20.6 dBFS (unprocessed) to -4.9 dBFS (most compression). Classical music is typically subjected to far less compression in the music industry than popular music. For comparison, an analysis of four commercial classical recordings from 2000-2008 showed RMS values ranging from -25.0 to -16.2 dBFS (average -19.8 dBFS). The range of compression conditions was intended to examine whether listeners prefer the amounts of compression currently used in the industry, or if more or less compression would be preferable, for both rock and classical music.

2.2.2.3 Un-equalized (UNEQ) condition

In the UNEQ condition, no loudness equalization was performed across stimuli following compression processing. This allowed loudness to vary among samples as a result of limiting.

2.2.2.4 Loudness-equalized (LEQ) condition

To assess the effects of CL on perceived quality when the role of loudness was reduced, a second condition was included in which loudness differences were minimized across varying levels of compression. Compressed music samples were first RMS-equalized to the value for the unprocessed sample using MATLAB software. Previous evidence suggests that amplitude-compressed speech is perceived as louder than uncompressed speech at equal RMS values (Moore et al., 2003). Therefore, further steps were taken to attempt to minimize loudness variation beyond the approximation provided by RMS-equalization. Stereo files were converted to mono. Using the time-varying loudness

DOS computer program developed by Glasberg and Moore (2002), the mono files were processed through the time-varying loudness model (Glasberg and Moore, 2002), calibrated to 65 dB SPL output with a diffuse-field response. Estimated loudness measurements were derived as minimum and maximum long-term loudness level values, measured in phons. Average overall loudness for each sample was calculated by taking the mean of all long-term loudness level values that were above two phons (corresponding to absolute threshold), in a manner similar to that described by Moore et al. (2003). For both genres, the -8 dBFS and -12 dBFS conditions had average overall loudness level values within one phon of that for the unprocessed file following RMS-equalization, so no further equalization was employed. For the remaining conditions (-16 dBFS, -20 dBFS, and -24 dBFS), compressed RMS-equalized stereo stimuli were attenuated using Adobe Audition 3.0 and then converted to mono and run through the Glasberg and Moore (2002) model. This process was repeated until all conditions for a single genre had average overall loudness level values within one phon of each other. A loudness-level step of 1 phon corresponds to a 1-dB change in level for a 1000-Hz tone, and listeners with normal hearing can typically detect level changes of approximately 0.5-1 dB for wideband noise (Moore, 2012). Therefore, the one-phon difference among stimuli was considered appropriately matched in loudness. Figure 2.1 shows the waveforms of the unprocessed samples and those with the highest level of compression (-24 dBFS threshold) in the UNEQ and LEQ conditions.

2.2.3 Instrumentation and playout

Stereo music samples were stored on a Dell Optiplex 990 computer as 24-bit, 44.1-kHz wave files. Stimuli were routed through a digital-to-analog converter (TDT RX8), an attenuator (TDT PA5), and a headphone buffer amplifier (TDT HB7) and were presented to listeners through Sennheiser HD580 earphones, which have a diffuse-field response. Listeners were seated in a sound-treated booth. The output of each earphone was calibrated using a 1000-Hz pure tone matched to the average RMS level of each group of six stimuli (rock UNEQ, rock LEQ, classical UNEQ, classical LEQ). The tone was measured using a Bruel & Kjaer sound level meter (model

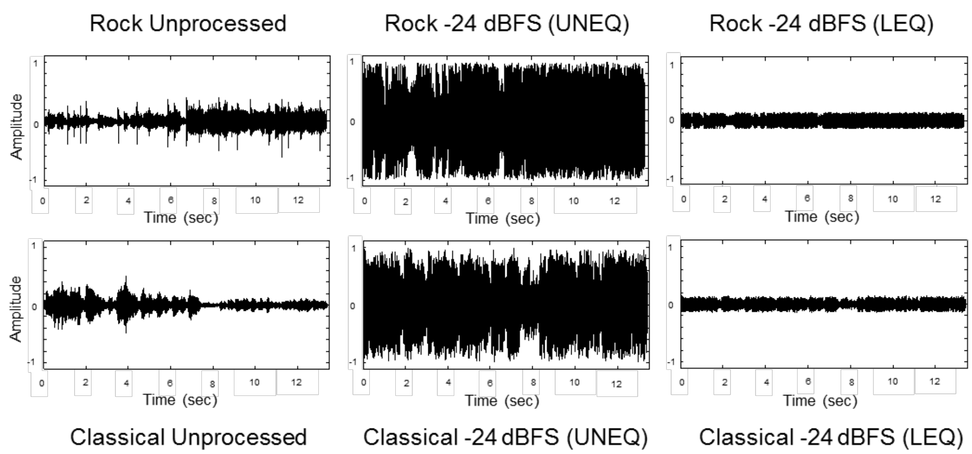

Figure 2.1: Waveforms of rock and classical music samples. Left panels: unprocessed stimuli. Middle panels: highest amount of compression (-24 dBFS threshold), shown for the UNEQ condition. Right panels: highest amount of compression (-24 dBFS threshold), shown for the LEQ condition.

2235) with an octave filter set (model 1624) and an artificial ear (model 4152) and was adjusted to 65 dB SPL using the TDT attenuator and headphone buffer. This calibration procedure resulted in average playout levels of 65 dB SPL for each block of trials.

2.2.4 Procedure

Listeners rated their perceptions of loudness, dynamic range, pleasantness, and preference using a six-point scaled paired-comparison method, which was modified from Dillon (1984) and was similar to the procedures used previously by Balfour and Hawkins (1992), Dahlquist and Leijon (2012), Munro and Lutman (2005), and Moore et al. (2011). The response interface is displayed in Figure 2.2. For each trial, one compression threshold condition was assigned to position A and another compression threshold condition was assigned to position B. Sample A was played through the headphones, followed by sample B. The task of the listener was to select one of the six possible options, choosing either sample A or sample B, along with a magnitude of the difference between the two samples (slightly, moderately, or much). All conditions were paired with all other conditions, and all conditions were presented in position A and in position B; therefore, each stimulus pair was presented twice, in randomized order, within one quality scale. Listeners rated all stimulus pairs on all four quality scales within one test block, in randomized order. The following four test blocks were completed, in randomized order, for a total of approximately six hours of testing spread out over several visits: rock UNEQ, rock LEQ, classical UNEQ, and classical LEQ.

2.3 Results

2.3.1 Acoustic outcomes

Several analyses were performed to quantify the acoustic effects of compression. The first analysis described the effects of compression on the wideband amplitude characteristics of the recordings, as illustrated in amplitude histograms. Figure 2.3 and Figure 2.4 plot the amplitude histograms for the compressed stimuli. In each plot, the solid line is for the unprocessed sample,

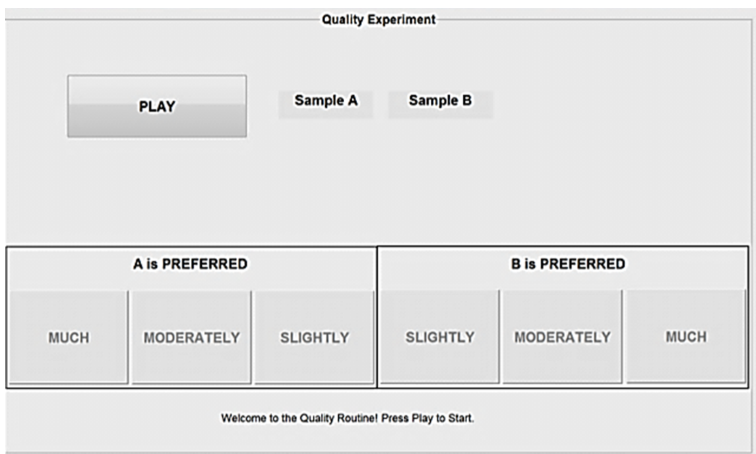

Figure 2.2: Example of the response interface. For the loudness scale, the text reads "A is louder," etc. For the dynamic range scale, the text reads "A has larger loudness changes," etc. For the pleasantness scale, the text reads "A is more pleasant," etc.

and the dashed line is for the compressed sample. The curves define the proportion of bins within a recording that were measured at a particular amplitude, relative to the RMS of the unprocessed stimulus. Measurements were taken at equal increments, using 51 histogram bins to span the range of -40 to +30 dB re: RMS. Figures 2.3 and 2.4 show that with increasing compression, the width of the curves is reduced, suggesting a decrease in crest factor with increasing compression (i.e. the distribution of samples is concentrated near the RMS, as opposed to a larger range of amplitudes). Additionally, in the UNEQ condition, the RMS of each compressed stimulus is higher than the RMS of the unprocessed recording, as demonstrated by an outward shift along the x-axis. In the LEQ condition, the RMS of each compressed stimulus is relatively close to the RMS of the unprocessed stimulus, while the shapes of the curves change similarly in the two conditions.

The second acoustic analysis was perceptually relevant, as it incorporated the effects of human auditory filters. Figures 2.5 and 2.6 are calculated from cumulative amplitude histograms within analysis bands equal to auditory filter bandwidths for normal hearing, as estimated using gammatone filters (Patterson et al., 1995). The signal envelope is computed in each gammatone filter band and converted to dB SPL, and the cumulative envelope level distributions are then computed from the envelope histograms. In Figure 2.5 , each line in the plot represents the level at which a certain percentage of samples is not exceeded within each auditory filter, plotted as a function of filter center frequency. For example, the top line for the uncompressed classical stimulus in the LEQ condition (solid black line, top left panel) indicates that 98% of the measured samples were at or below 40 dB SPL at 100 Hz, while 98% of the measured samples were at or below 20 dB SPL near 5000 Hz. The width of the distribution between the 98% line (top line, solid black) and the 30% line (bottom line, hash marks) is smaller for the -24 dBFS LEQ classical stimulus (top panel second from the left) than for the uncompressed classical stimulus (top left panel). This change in the difference between the 98% and 30% cumulative distributions indicates a reduction in dynamic range across frequencies. Similar effects are seen for classical music in the UNEQ condition and for rock music in the LEQ and UNEQ conditions.

Figure 2.6 shows difference plots, representing the difference in dynamic range between the

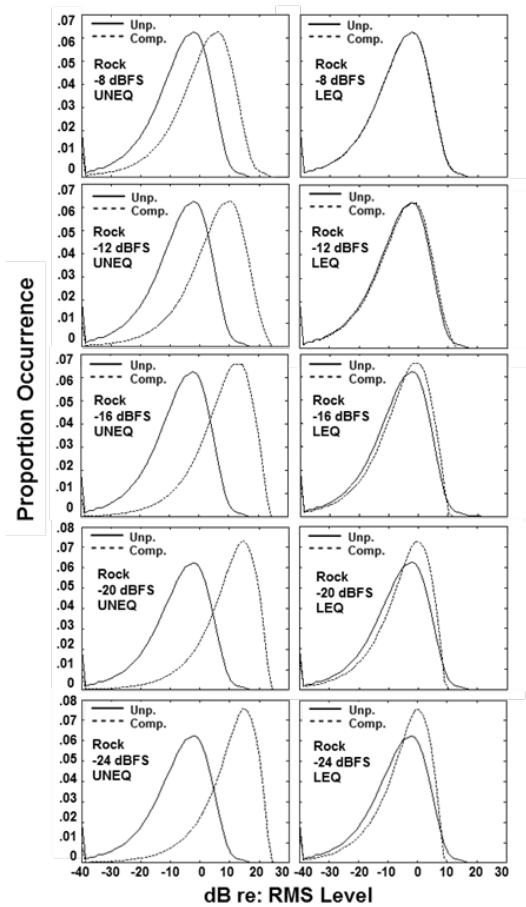

Figure 2.3: Wideband amplitude histograms comparing unprocessed rock stimuli to compressed rock stimuli. In each plot, the solid line is for the unprocessed music sample, and the dashed line is for the compressed music sample. The left panels show the UNEQ condition, and the right panels show the LEQ condition.

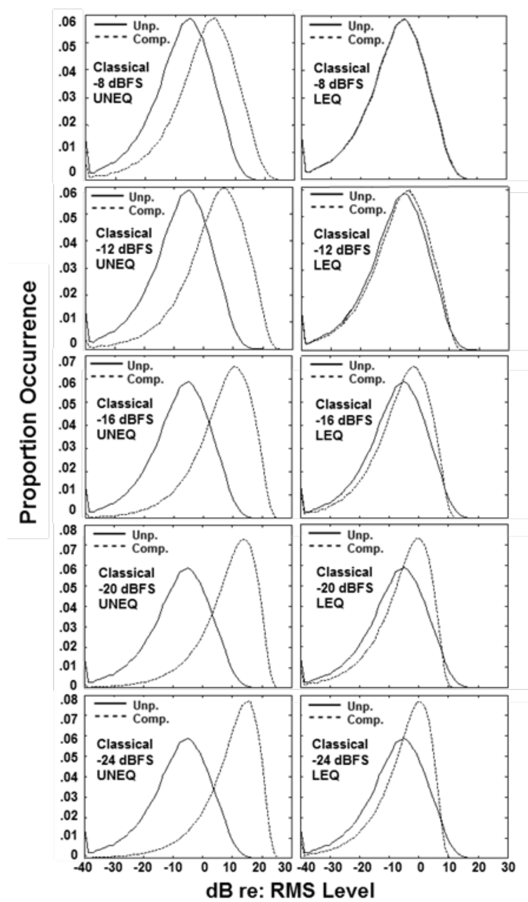

Figure 2.4: Wideband amplitude histograms comparing unprocessed classical stimuli to compressed classical stimuli. In each plot, the solid line is for the unprocessed music sample, and the dashed line is for the compressed music sample. The left panels show the UNEQ condition, and the right panels show the LEQ condition.

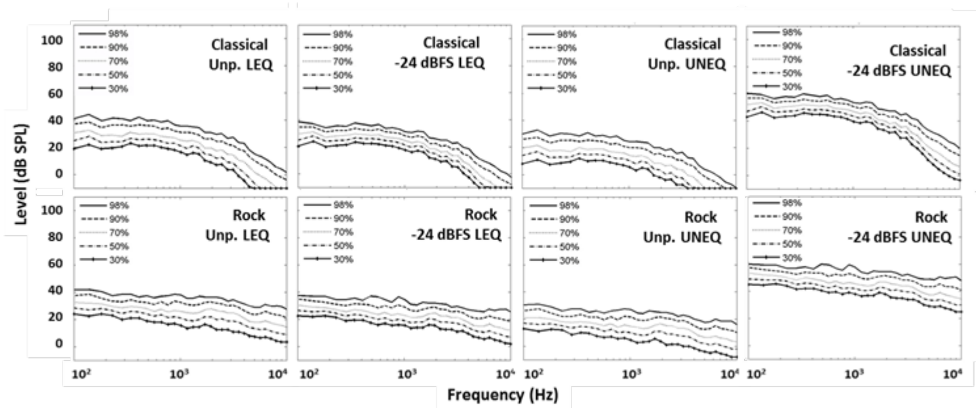

Figure 2.5: Dynamic-range histograms, calculated from cumulative amplitude histograms within analysis bands equal to auditory filter bandwidths. Each line in the plot represents the level at which a certain percentage of samples over the duration of the stimulus is not exceeded, across the frequency range from 100 to 10,000 Hz. From top to bottom, the lines represent the following percentages: 98%, 90%, 70%, 50%, and 30%. Both genres are represented in the LEQ and UNEQ conditions, for unprocessed stimuli ("Unp.") and highly compressed stimuli (-24 dBFS threshold).

compressed and uncompressed stimuli in the LEQ condition, calculated in auditory filter bandwidths as described above. The dynamic range of each signal was calculated as the difference between the 98% and the 30% cumulative distributions. The dynamic-range difference between two signals was then calculated by subtracting the dynamic range of the compressed signal from that of the unprocessed signal, with the compressed stimulus being the -24 dBFS condition. The difference plots show that the uncompressed stimuli have approximately a 5-dB greater dynamic range than the compressed stimuli, across auditory filter bandwidths from 100 to 10,000 Hz.

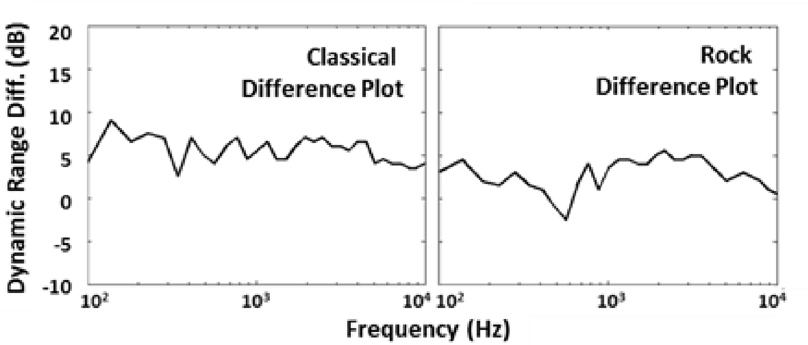

Figure 2.6: Difference plots showing the dynamic-range difference between the unprocessed stimulus and the -24 dBFS condition, analyzed within auditory filter bandwidths and represented across frequencies from 100 to 10,000 Hz. Dynamic range is defined as the difference between the 30% distribution and the 98% distribution for each stimulus.

Taken together, the wideband amplitude histograms (Figures 2.3 and 2.4) and the dynamic-range plots (Figures 2.5 and 2.6) demonstrate that CL reduces the dynamic range and crest factor of recorded music, and in the UNEQ condition, CL has the additional effect of increasing the level of the signal. The two types of figures show two different perspectives of the effects of compression. Figures 2.3 and 2.4 represent changes to the overall amplitude envelope, with low levels of compression causing little envelope distortion and high levels of compression causing substantial envelope distortion. From another viewpoint, Figures 2.5 and 2.6 show that the dynamic-range changes caused by compression are relatively consistent across a wide range of frequencies important for human hearing.

In the final acoustic analysis, Table 2.1 outlines several measured changes to the acoustic properties of the stimuli due to compression. In the UNEQ condition, the RMS levels of both genres of music increased with increasing compression, as expected. The crest factor – measured as the ratio of the amplitude exceeded 99% of the time and the RMS, and expressed in dB – decreased with increasing compression. Average predicted loudness level and maximum predicted loudness level in phons increased with increasing compression, as measured using the Glasberg and Moore (2002) model. The stimuli were also analyzed using a modified version of the Hearing Aid Speech Quality Index (HASQI), developed by Kates and Arehart (2010). HASQI provides a measure of the amount of signal modification caused by a processing scheme by comparing the amplitude envelope and frequency spectrum of a processed signal relative to those for the unprocessed signal. The measure has a value between 0 (very low fidelity) and 1 (perfect fidelity) (Kates and Arehart, 2010). The current version of HASQI was modified for music by computing new weights to fit the quality ratings reported by Arehart et al. (2011). Table 2.1 shows that HASQI values decreased with increasing compression, indicating a larger amount of signal modification (i.e. reduced fidelity) as more compression was applied. In the LEQ condition, RMS levels remained relatively stable but were slightly reduced with increasing compression. Average predicted loudness levels remained stable (within one phon), but maximum predicted loudness levels decreased with increasing compression. HASQI values decreased with increasing compression, in a manner similar to that for the UNEQ condition.

2.3.2 Quality ratings

Listener ratings were transformed into an overall score for each condition, using a procedure similar to that used by Moore et al. (2011). For each pair of excerpts, the stimulus that was chosen by the listener was assigned a score from +1 to +3, while the stimulus that was not chosen was assigned the same score with the opposite sign. For example, if the subject rating was "A is much preferred," the condition in position A received a score of +3, and the condition in position B received a score of -3. A score of +3 on the preference scale would mean that a particular condition was much preferred to all other conditions, a score of +2 would mean that a condition

Table 2.1: Acoustic properties of unprocessed and compressed stimuli.

	Rock Un-Equalized						Rock Loudness-Equalized					
Threshold	RMS (dBFS)	RMS (dB SPL)	Crest Factor	Max Predicted Loudness (phons)	Avg Predicted Loudness (phons)	HASQI	RMS (dBFS)	RMS (dB SPL)	Crest Factor	Max Predicted Loudness (phons)	Avg Predicted Loudness (phons)	HASQI
No Comp	-20.83	54.31	9.93	67.70	58.70	1.00	-20.83	65.59	9.93	67.70	58.7	1.00
-8 dB FS	-12.89	62.24	9.99	76.20	68.61	0.77	-20.81	65.61	9.99	67.70	58.8	0.86
-12 dB FS	-9.24	65.90	9.64	79.70	72.86	0.73	-20.74	65.67	9.64	67.90	59.19	0.85
-16 dBFS	-6.75	68.39	8.43	81.00	75.56	0.68	-21.80	64.62	8.43	65.40	58.43	0.81
-20 dBFS	-5.69	69.45	7.76	81.00	77.30	0.64	-21.92	64.50	7.76	64.00	58.81	0.76
-24 dBFS	-5.43	69.71	7.58	81.00	77.68	0.62	-22.40	64.02	7.58	63.10	58.34	0.74
	Classical Un-Equalized						Classical Loudness-Equalized					
Threshold	RMS (dBFS)	RMS (dB SPL)	Crest Factor	Max Predicted Loudness (phons)	Avg Predicted Loudness (phons)	HASQI	RMS (dBFS)	RMS (dB SPL)	Crest Factor	Max Predicted Loudness (phons)	Avg Predicted Loudness (phons)	HASQI
No Comp	-20.63	54.42	10.68	66.10	53.60	1.00	-20.63	65.44	10.68	66.10	53.60	1.00
-8 dB FS	-12.69	62.36	10.65	74.60	63.69	0.80	-20.57	65.50	10.65	66.20	53.72	0.87
-12 dB FS	-9.32	65.73	9.86	77.30	70.80	0.76	-20.28	65.79	9.86	65.80	54.55	0.84
-16 dBFS	-7.12	67.93	8.82	77.30	70.90	0.73	-21.61	64.46	8.82	61.80	53.67	0.78
-20 dBFS	-5.62	69.43	8.01	77.30	72.94	0.69	-21.68	64.39	8.01	59.90	54.22	0.70
-24 dBFS	-4.91	70.14	7.71	77.30	73.83	0.65	-21.64	64.43	7.72	55.90	54.51	0.64

was moderately preferred to all other conditions, a score of -1 would mean all other conditions were slightly preferred to that condition, etc. All subscores for a particular condition were averaged to obtain an overall score for that condition.

2.3.2.1 UNEQ and LEQ conditions

Figure 2.7 shows average rating scores across all subjects for rock and classical music in the UNEQ condition. The x-axis plots the limiting threshold, with the amount of compression ranging from no compression to the most compression. The y-axis plots the overall score for each compression condition, where positive values indicate higher scores on each scale, and negative values indicate lower scores on each scale. Perceived loudness increased with increasing compression for both genres. For rock music, ratings increased with increasing compression for dynamic range, pleasantness, and preference, up to about the mid-point of compression (-12 dBFS threshold). Higher levels of compression (beyond the -12 dBFS threshold) led to little effect for thresholds from -16 to -24 dBFS. For classical music, low levels of compression were again rated higher on dynamic range, pleasantness, and preference, compared to no compression. However, above the -8 dBFS threshold, higher levels of compression led to a decrease in perceived dynamic range and pleasantness for classical music. A repeated-measure mixed analysis of variance (ANOVA) was performed for each UNEQ test block, using compression threshold as the within-subjects variable and musician vs. non-musician as the between-subjects variable. Greenhouse-Geisser corrections were used in each analysis due to a significant Mauchley's test of sphericity. The effect of compression was significant for rock music for loudness [$F(2.3, 51.1)=558.77$; $p<0.001$], dynamic range [$F(2.6, 56.2)=17.71$; $p<0.001$], pleasantness [$F(1.6, 36.1)=12.87$; $p<0.001$], and preference [$F(1.6, 35.8)=19.50$; $p<0.001$]. The effect of compression was also significant for classical music for loudness [$F(2.7, 69.0)=433.39$; $p<0.001$], dynamic range [$F(1.8, 39.9)=7.46$; $p=0.002$], pleasantness [$F(1.6, 36.0)=9.08$; $p=0.001$], and preference [$F(1.7, 38.5)=5.26$; $p=0.012$].

No significant differences were found between groups in any analysis for the UNEQ condition. However, a significant compression by group interaction was observed for rock music for loudness

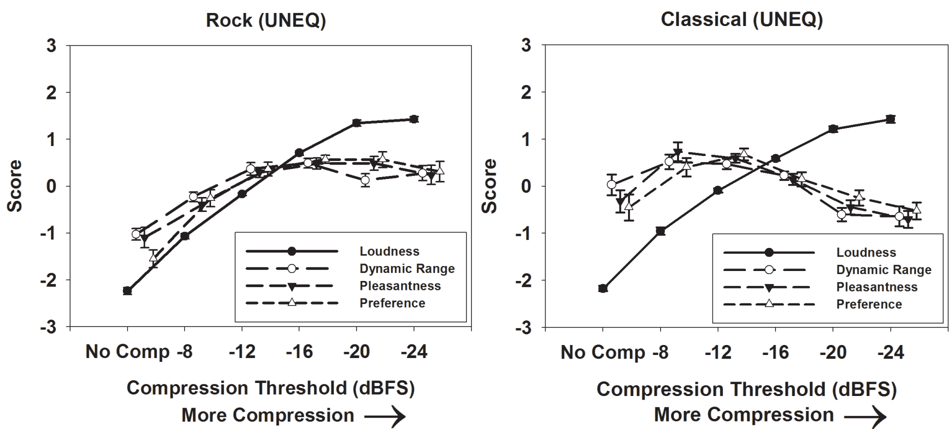

Figure 2.7: Average scores across all subjects for rock music (left panel) and classical music (right panel) in the UNEQ condition. Scores for each of the four quality scales are plotted against the compression threshold conditions. Error bars represent ±1 standard error (S.E.).

[F(2.3, 51.1)=4.68; p=0.010] and dynamic range [F(2.5, 56.2)=3.61; p=0.024]. Figure 2.8 shows these interactions. The left panel of Figure 2.8 demonstrates that the slope for perceived loudness was shallower for the musician group than for the non-musician group. This finding is consistent with previous work indicating that musicians may estimate smaller ranges of loudness than non-musicians, possibly due to the experience of the former using dynamic markings in music (Geringer, 1995). For dynamic range (Figure 2.8 , right panel), non-musicians perceived a larger increase in dynamic range with increasing compression, which may be due to an overpowering effect of overall loudness. The musician group also demonstrated an increase in perceived dynamic range for low levels of compression, up to the -12 dBFS condition. However, with further increases in compression, the musician group seemed to have a greater ability to separate overall loudness from dynamic range, as shown by their decreasing dynamic-range judgments above the -12 dBFS threshold. Disregarding the judgments for the unprocessed and -8 dBFS conditions, the effect of compression for the -12 dBFS through -24dBFS conditions was not significant for perceived dynamic range in the non-musician group [F(1.5, 13.4)=1.16; p=0.328] but was marginally significant in the musician group [F(1.5, 19.9)=3.57; p=0.058].

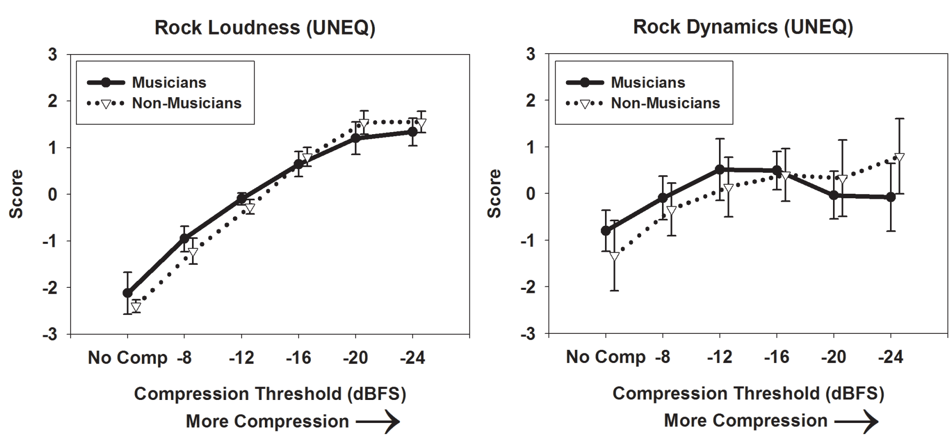

Figure 2.8: Significant group by compression interactions comparing musicians to non-musicians. Left panel: rock UNEQ condition, with loudness plotted against compression threshold. Right panel: rock UNEQ condition, with dynamic range plotted against compression threshold. Error bars represent ±1 S.E.

Figure 2.9 plots average rating scores across all subjects for rock and classical music in the LEQ condition. All quality scales showed a similar trend: low levels of compression had little effect. For example, the slopes of the lines for scales of loudness, dynamic range, pleasantness, and preference are all close to zero for rock music over the range from no compression to -12 dBFS. Above the -12 dBFS threshold, however, higher levels of compression were rated lower on all four scales. These patterns were observed for both rock and classical music. However, the drop in ratings for pleasantness and preference occurred earlier for classical music, after only the -8 dBFS threshold. This finding suggests that listeners may be more sensitive to changes caused by compression in classical music. An ANOVA with the same design as before was conducted. The effect of compression was significant for rock music for loudness [$F(2.2, 48.7)=64.17$; $p<0.001$], dynamic range [$F(2.1, 46.5)=17.03$; $p<0.001$], pleasantness [$F(2.2, 47.9)=16.77$; $p<0.001$], and preference [$F(2.1, 45.9)=28.16$; $p<0.001$]. The effect of compression was also significant for classical music for loudness [$F(1.6, 36.0)=21.99$; $p<0.001$], dynamic range [$F(1.9, 42.5)=25.44$; $p<0.001$], pleasantness [$F(1.7, 36.8)=34.34$; $p<0.001$], and preference [$F(2.1, 45.4)=33.21$; $p<0.001$]. No significant between-group differences and no significant group by compression interactions were found.

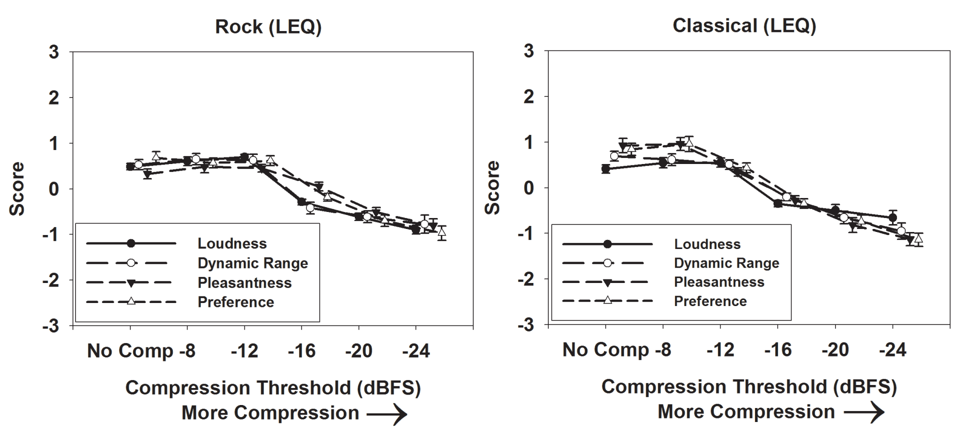

Figure 2.9: Average scores across all subjects for rock music (left panel) and classical music (right panel) in the LEQ condition. Scores for each of the four quality scales are plotted against the compression threshold conditions. Error bars represent ±1 S.E.

Figure 2.10 summarizes preference scores for rock music (left panel) and classical music (right panel) in the LEQ and UNEQ conditions. The open circles demonstrate that preference increased with increasing compression when loudness varied (UNEQ), up to about the mid-point of compression (-12 dBFS threshold), for both rock and classical music. As shown by the closed circles, low levels of compression did not affect preferences when loudness differences were minimized (LEQ). This finding was true for both rock and classical music but occurred at a lower limiting threshold for the classical music than for the rock music. In contrast, high levels of compression were often detrimental, as shown by the negative slopes occurring after -12 dBFS for rock LEQ, after -12 dBFS for classical UNEQ, and after -8 dBFS for classical LEQ. For the rock UNEQ condition, high levels of compression were neither detrimental nor beneficial, as shown by the relatively flat slope beyond the -12 dBFS threshold.

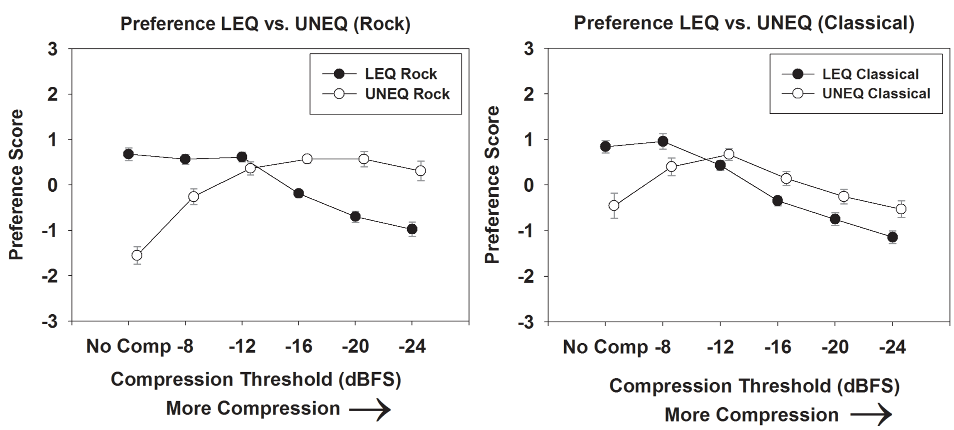

Figure 2.10: Average scores on the preference scale across all subjects, comparing scores in the LEQ condition to scores in the UNEQ condition. Preference scores are plotted against compression threshold. Left panel: rock music. Right panel: classical music. Error bars represent ±1 S.E.

2.3.2.2 Listener differences

Substantial individual variability was observed, particularly in the UNEQ condition. In general, listeners tended to fall within three different profiles. As shown in the left panel of Figure 2.11, Profile A listeners preferred low levels of compression compared to no compression. The exact compression threshold that was most preferred varied among Profile A listeners. Profile A may reflect a positive relationship between loudness and preference, up to a point, after which distortion caused by compression becomes detrimental. This was the most common profile, demonstrated by 14 listeners for the rock stimulus and 16 listeners for the classical stimulus. For Profile B (Figure 2.11, middle panel), preference increased monotonically with increasing compression. This pattern may be due to loudness trumping distortion, so that louder samples were always preferred. Nine listeners showed Profile B for rock music, and three listeners showed Profile B for classical music. Finally, for Profile C (Figure 2.11, right panel), preference decreased monotonically with increasing compression. These listeners may have been particularly sensitive to distortion caused by compression. Five subjects showed Profile C for classical music, and only one listener showed Profile C for rock music. Listeners' preferences were not significantly predicted by group (musician vs. non-musician), audio recording experience, hours per week spent listening to music, gender, or age.

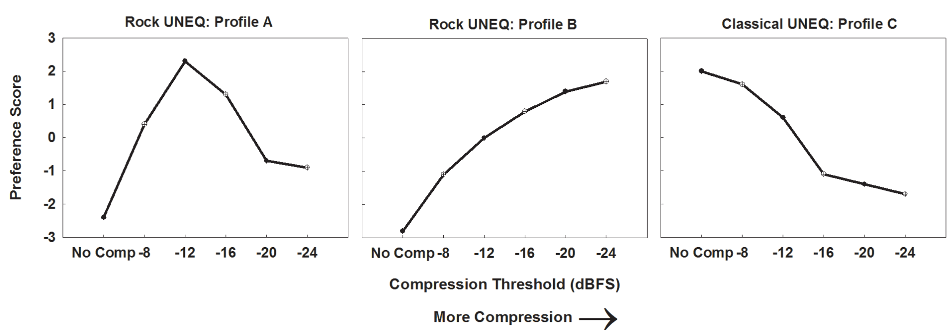

Figure 2.11: Three listener profiles that emerged on the preference scale in the UNEQ condition. Each panel shows a different profile, and each profile is represented by an individual listener. The left and middle panels (profiles A and B, respectively) are shown for rock music. The right panel (profile C) is shown for classical music.

2.3.2.3 Acoustic correlates of loudness

Although LEQ stimuli underwent a loudness equalization procedure, listeners' judgments revealed a decrease in perceived loudness with increasing compression above the -12 dBFS threshold for rock and classical music, as shown in Figure 2.9 and discussed in the results section above. This finding was surprising due to the expectation that LEQ stimuli would show very little variation in perceived loudness. Note that the loudness equalization process used for the LEQ stimuli was based on average predicted loudness derived from the Glasberg and Moore (2002) time-varying loudness model. In order to equate average predicted loudness, the RMS levels of several stimuli were attenuated. Additionally, crest factor and maximum predicted loudness both decreased with increasing compression in the LEQ condition. One possible explanation for the decrease in loudness judgments may be that listeners felt compelled to make decisions in terms of some other subjective effect, due to the small variation in loudness across stimuli. For example, loudness judgments may have been influenced more by the peak levels and RMS levels of the recordings than by the predicted average loudness. This explanation is supported by significant correlations in the LEQ condition between loudness ratings and RMS (rock: $r=0.993$, $p<0.001$; classical: $r=0.973$, $p=0.001$), loudness ratings and crest factor (rock: $r=0.978$, $p=0.001$; classical: $r=0.951$, $p=0.004$), and loudness ratings and maximum predicted loudness (rock: $r=0.995$, $p<0.001$; classical: $r=0.950$, $p=0.004$). Despite the trend for decreasing loudness with increasing compression observed in the LEQ condition, the range of loudness scores was far smaller than in the UNEQ condition.

2.4 Discussion

The present findings have characterized the acoustic and perceptual effects of CL for recorded music. Acoustic measurements revealed a decrease in crest factor and increased modification of the amplitude envelope with increasing compression. Changes in crest factor and amplitude envelope are consistent with the purpose of CL (i.e. reducing the dynamic range of a music signal). Additionally, these effects are similar to the smoothing of the amplitude envelope that has been reported previously for hearing-aid WDRC (Souza and Turner, 1998; Jenstad and Souza, 2005). Perceptually, listeners were able to detect a reduction in dynamic range with increasing compression, especially musician listeners. Combined changes to the amplitude envelope and spectrum of each signal, as measured by HASQI (Kates and Arehart, 2010), indicated a loss of fidelity to the original recording with increasing compression. This acoustic distortion of the music may partially explain the general decrease in pleasantness reported by listeners at the highest levels of compression.

When loudness varied as a result of compression (UNEQ), listeners – on average – preferred a small amount of compression compared to no compression for classical music and a moderate amount of compression compared to no compression for rock music. Therefore, CL may be advantageous when applied modestly to rock and classical music. This finding could be due, in part, to the increase in loudness associated with CL, consistent with previous research showing that listeners prefer louder music, all other things being equal (Maempel and Gawlik, 2009). However, judgments in the UNEQ condition also indicated that applying compression beyond the optimal amount was detrimental for classical music and led to no further improvement for rock music. Several factors may explain this outcome. One possible explanation is that the effects of distortion became more bothersome with increasing amounts of compression, leading to a decrease in preference. A second possibility is that the loudness of the more heavily compressed samples exceeded the preferred loudness level. However, the highest level of the UNEQ stimuli was approximately 70 dB SPL, which is no higher than preferred listening levels for music, as reported by Smeds (2004) and Muchnik et al. (2012). A third factor that may have influenced ratings is the average level of the stimuli (65 dB SPL). Preferences may have differed with higher or lower presentation levels, due to the relationship between preference and loudness. For example, at a higher average presentation level, less compression may have been preferred to that observed here.

Listeners' ratings in the LEQ condition provided insight into the effects of distortion caused by CL when the role of loudness was diminished. The low levels of compression that were preferred in the UNEQ condition were neither harmful nor beneficial to music of both genres when the samples were loudness-equalized. However, higher amounts of compression led to lower ratings on the pleasantness and preference scales. These findings are essential to understanding the effects of CL in real-world listening because the loudness of recorded music is ultimately controlled by the end user, who may adjust the volume control. Additionally, technology is currently available to normalize the loudness of music files as they are played out from a portable music device (Vickers, 2011). The combined results from UNEQ and LEQ conditions suggest that using heavy compression to maximize the loudness of recorded music may not lead to the listener preferences expected by some members of the music industry. Moreover, the distortion resulting from high levels of compression could reduce perceived music quality for most listeners.

Despite the average tendencies, judgments of compressed music differed across listeners. In the UNEQ condition, three main listener profiles emerged on the preference scale. The most common profile was characterized by an ideal amount of compression; below and above this point, preference decreased. In the second profile, listeners preferred louder, more compressed, music to less compressed music. The third profile showed a preference for no compression compared to any amount of compression. Musicians and non-musicians demonstrated generally similar patterns of judgments, with the exceptions of perceived loudness and dynamic range for rock music in the UNEQ condition. Musicians showed less change in loudness and a greater reduction in perceived dynamic range with increasing compression, compared to non-musicians. This finding is consistent with previous research demonstrating smaller ranges of perceived loudness in experienced musicians (Geringer, 1995).

The results of this study may have implications beyond optimizing sound quality in the commercial music industry. For instance, concerns have been raised about over-compression causing listener fatigue (Stone et al., 2009) and elevating risk for music-induced hearing loss (Vickers, 2011). Additionally, listeners with hearing loss who use compression hearing aids may experience combined effects of music-industry and hearing-aid compression. Future research should include the risks and treatment of hearing loss when examining the effects of dynamic-range compression on music.

2.5 Conclusions

The present study provides evidence to contribute to the "loudness war" debate taking place among members and consumers of the music industry. Contrary to a "louder is better" mentality, the current findings suggest that "louder is better . . . to a point." Louder, more compressed, music samples were preferred when dynamic-range compression was applied to a moderate extent.

However, high levels of compression were either detrimental to music quality or had little effect, suggesting that a shift to more restrained use of dynamic-range compression may improve listeners' perceptions of the quality of commercial music.

Chapter 3

Experiment 2: Combined Music-Industry and Hearing-Aid Dynamic-Range Compression

3.1 Introduction

Hearing aids have been shown to improve quality of life by decreasing the psychological, social, and emotional effects of hearing loss (Chisolm et al., 2007). However, a recent study by Kochkin (2012) revealed that only about 25% of people with hearing loss actually own hearing aids. Kochkin (2012) also reports that sound quality, including the quality of music, is an important consideration in a person's decision to adopt hearing aids. Unfortunately, 43% of hearing-aid users feel that hearing aids either make no difference in their music enjoyment or make music less enjoyable (Leek et al., 2008). These findings suggest that traditional hearing-aid fittings may not fulfill the needs of many hearing-aid users who listen to music or are musicians themselves.

A number of variables might contribute to this lack of satisfaction. The perception of music with hearing aids depends on several factors, including the properties of the external music signal and any pre-processing applied, the effects of hearing-aid digital signal processing, and listener-specific auditory characteristics.

Hearing-aid processing depends on the acoustic properties of the input. Music and speech differ acoustically, as music is often higher in intensity and more dynamic than speech (Chasin, 2003, 2006, 2010). Music and speech also have different spectral and temporal characteristics. The spectrum of speech is relatively similar across talkers, while the spectrum of music varies widely depending on the instrument (Chasin and Russo, 2004). The modulation rate for speech centers around 4 Hz and is related to syllabic structure. In contrast, the modulation rate for music depends on the tempo. For example, if a musical piece is performed at a tempo of 120 beats per minute (bpm), the primary modulation rate will emerge around 2 Hz, and the rhythmic structure will be shown in harmonic patterns around the 2-Hz modulation (Scheirer, 1998; Todd et al., 1999).

Beyond the distinction between speech and music, different types of music signals have different acoustic features. For example, music may be either live or recorded. Commercial music often undergoes signal processing during the audio-industry recording process, including a specific type of dynamic-range compression called compression limiting (CL). CL simultaneously reduces the dynamic variation of music while increasing its root-mean-square (RMS) level, with the intention of increasing the loudness of the recording. The results presented in Chapter 2 demonstrated that most normal-hearing participants preferred a small amount of CL when it increased loudness; however, when loudness was held constant, most listeners rated quality lower for higher levels of CL. This finding suggests that certain genres of music that are often heavily compressed within the music industry, such as pop/rock music, may have low sound quality before being processed by a hearing aid. Many hearing-aid manufacturers now include music programs, which are often coupled with wireless streaming tools that facilitate listening to recorded music through personal audio devices. Thus, an important factor to consider when studying music and hearing aids is the potential relationship between CL and hearing-aid processing.

Most modern hearing aids use wide dynamic-range compression (WDRC), which provides more gain to low-level sounds and less gain to high-level sounds to ensure audibility, while maintaining a comfortable range of loudness for listeners with hearing loss. WDRC has been shown to restore loudness growth for many types of sounds (Jenstad et al., 2000). However, WDRC also causes acoustic modifications, and signal processing parameters selected in hearing aids must balance a trade-off between audibility and distortion (Kates, 2010).

Two important WDRC parameters that affect the audibility-distortion relationship are release time and number of channels. The release time describes how quickly or slowly the system responds to reductions in signal levels (Kates, 2008). Fast release times restore audibility quickly for weak sounds following intense sounds, maximizing moment-to-moment audibility (Gatehouse et al., 2006; Moore, 2008). Fast compression can also compensate for loudness changes in specific frequency ranges and is effective at approximating normal loudness perception (Moore, 2008). However, fast compression also leads to distortion of the temporal envelope (Jenstad and Souza, 2005), flattens the spectrum (Plomp, 1988), and can cause nonlinear distortion (Kates, 2010). Slow release times cause less distortion to the temporal envelope (Jenstad and Souza, 2005) and short-term spectrum (Moore, 2008). However, with slow release times, gain takes longer to adjust to a new environment or a weak sound following an intense sound.

Hearing aids typically separate the incoming signal into a number of frequency bands, and several bands can be combined to form a processing channel. The use of a single channel or a small number of channels preserves spectral relationships in a signal (Bor et al., 2008). However, small numbers of channels may not provide sufficient resolution for frequency-specific amplification. Therefore, larger numbers of channels can improve audibility across frequencies, but more channels may also lead to greater amounts of distortion (Kates, 2010). Additionally, the channel number interacts with the release time, such that fast time constants have greater effects when more channels are used (Plomp, 1988; Henning and Bentler, 2008; Moore, 2008). The desired combination of WDRC parameters depends on the goal of the system. Compression systems that emphasize audibility use fast time constants and a greater number of channels, while slow time constants and fewer channels are used when minimizing distortion is the primary concern (Kates, 2010; Moore, 2008).

Several studies have compared WDRC to other processing schemes for music perception. Davies-Venn et al. (2007) measured quality ratings for classical and popular vocal music using WDRC, peak clipping, and compression limiting. Quality ratings revealed that WDRC was rated significantly more pleasant than peak clipping and compression limiting. In another study, Higgins et al. (2012) measured quality ratings for classical, rock, and jazz music, comparing adaptive dynamic range optimization (ADRO) to WDRC. ADRO is a modified linear strategy that places sounds in a range between audible and comfortably loud. ADRO is based on a histogram of the signal intensity, and the histogram levels are updated relatively slowly. Thus, ADRO has slower effective time constants than most WDRC implementations. In their study, Higgins et al. found that ADRO was preferred to WDRC. While it might seem that these two studies support different conclusions about WDRC and music, comparing the results is challenging due to differences in the numbers of channels and release times used. Davies-Venn et al. used a 15-channel compression system with a dual fast and slow detector. Therefore, the effective release time may have varied across channels at any given moment. Higgins et al. used a dual fast (4-channel) and slow (13-channel) system for the WDRC hearing aid but used a 32-channel system for the ADRO hearing aid. In another study that compared linear processing to fast compression with varying numbers of channels between 1 and 16, van Buuren et al. (1999) found that linear amplification was preferred to WDRC for three different instrumental pieces and one pop music selection.

Other experiments have compared fast and slow release times for music quality within the same hearing aid. Hansen (2002) investigated release times using WDRC within a 15-channel simulated hearing aid, for classical and pop music. Both normal-hearing listeners and listeners with hearing loss preferred the longest release time for music quality. Similarly, Moore et al. (2011) used jazz, classical, and solo percussion music to quantify the effect of compression speed on pleasantness in a five-channel simulated hearing aid. The authors found that slow compression was more pleasant than fast compression for classical music at a lower input level and for both classical and jazz music at a higher input level. Arehart et al. (2011) measured sound quality for unprocessed music compared to a simulated 18-channel hearing aid and found similar results for slow and fast release times. Fewer studies have directly examined the preferred number of channels for music. Van Buuren et al. (1999) found that increasing the number of channels was detrimental to music quality, such that a 16-channel condition was rated lower quality than 4-channel processing and 1-channel processing. Taken together, these recent studies provide some insights into the effects of several WDRC parameters on music perception. Yet, there is currently a lack of information related to systematic combinations of different release times with different numbers of channels.

An additional factor that might affect music perception with hearing aids is the auditory characteristics of the individual listener. In addition to decreased audibility, hearing loss leads to abnormal suprathreshold processing, including reduced frequency selectivity (Carney and Nelson, 1983), loss of temporal resolution (Glasberg et al., 1987), and reduced dynamic range (Moore and Glasberg, 1997). Suprathreshold processing deficits have been linked to perceptual consequences of WDRC processing. Souza et al. (2012) showed that frequency resolution predicted benefit from multichannel WDRC for subjects with hearing loss. Namely, listeners with broader auditory filters were less able to make use of multichannel compressed signals for vowel identification than listeners with narrower filters. Neuman et al. (1995a) found differences between listeners with smaller and larger dynamic ranges in their preferred listening levels for compressed speech. Although these studies were performed using speech rather than music, they indicate relationships between WDRC benefit and various perceptual abilities. Therefore, the current study included several psychoacoustic measures to determine whether individual hearing-loss characteristics affect the perception of hearing-aid processed music.

The purpose of this study was to quantify the acoustic and perceptual effects of dynamic-range compression on recorded music processed through simulated hearing aids. The experimental design was to vary the effects of WDRC on the acoustic structure of music stimuli by constructing a set of conditions that combined slow and fast release times with small and large numbers of channels. To consider interactions that might take place between CL and WDRC, the conditions included several levels of CL, WDRC, and CL+WDRC. The current study also sought to understand individual variability in listeners with hearing loss by examining the relationships between preferred compression parameters and individual perceptual characteristics.

3.2 Method

3.2.1 Subjects

Listeners included 18 adult hearing-aid users (12 females and 6 males), who were ages 49-87 (mean age 70 years). The participants were recruited from the Denver/Boulder, Colorado metropolitan area according to IRB-approved advertising and recruiting procedures. Research sessions took place at the Hearing Research Laboratory on the University of Colorado-Boulder campus. Listeners received a complimentary hearing evaluation prior to their participation and were compensated at an hourly rate for the time they spent in the research sessions. All listeners had acquired, bilateral sensorineural hearing loss that was symmetric (difference in pure-tone average of 500, 1000, 2000, 4000 Hz less than 15 dB between ears) and ranged from mild to severe (except for one listener who had thresholds in the profound range from 4000-6000 Hz in one ear only). Subjects' audiometric thresholds are plotted in Figure 3.1. Listeners also passed the Mini-Mental State Exam (Folstein et al., 1975) with a score of 27 or better to verify the ability to complete everyday thinking tasks and were free of neurological impairments that may impact cognition or auditory perception.

To determine whether differences in musical training affected perceived quality of hearing-aid processed music, participants' musical backgrounds were assessed using a survey. The survey was the same as that used in Experiment 1 (Chapter 2) and was adapted from a questionnaire previously used by Parbery-Clark et al. (2011). The survey asked participants about their current and past musical training and practice, participation in school and community musical ensembles, academic music classes and degrees awarded, music teaching experience, conducting experience, and ability to read musical notation. Questions were also asked about audio recording experience and listening habits (hours per week spent listening to live and recorded music, genres of music that were typical for listening). Based on responses to the survey, participants were divided into musician and non-musician groups. The musician group included nine listeners (ages 59-80), who had at least fifteen years of musical training on their primary instrument or voice (range 15-58 years). All of the musicians were trained on two or more instruments; therefore, the total amount of musical experience was actually greater than the number of years reported on the primary instrument. Musicians reported self-rated proficiency ranging from 4 to 10, on a scale of 1 to 10, on their primary instrument. One of the musicians had been awarded a bachelor's degree in music, and one of the musicians reported significant audio editing experience. The non-musician group

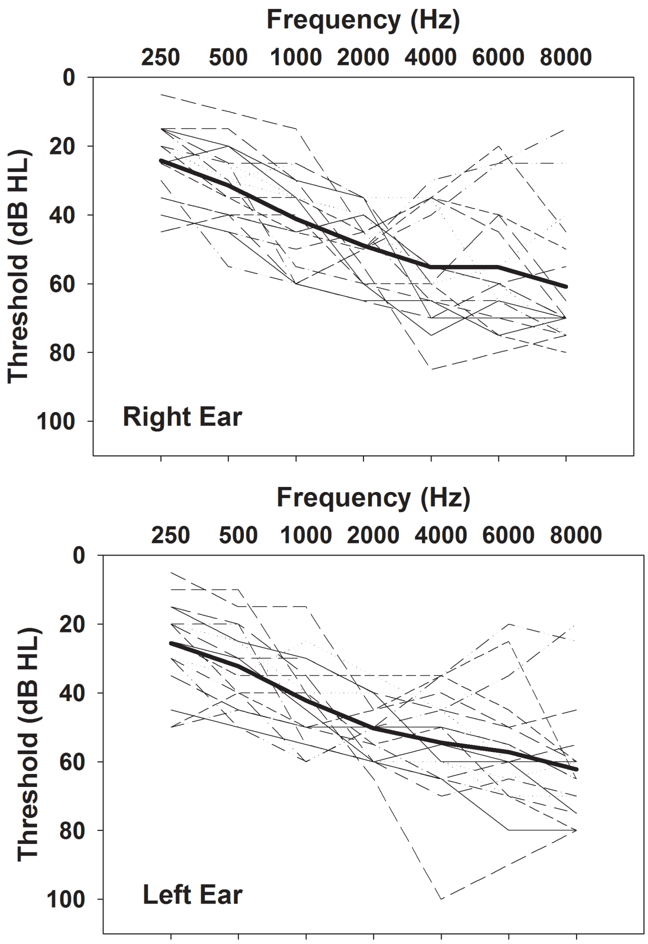

Figure 3.1: Audiometric thresholds of the 18 listeners. Average thresholds are shown in the solid black lines.

included nine listeners (ages 49-87). One of the nine non-musicians reported no musical training. For the other eight non-musicians, six reported some musical experience (2 to 12 years total) which was suspended at least 30 years prior to the study, and two participants reported current playing of the guitar, which was self-taught and self-rated at a proficiency level of 1.

3.2.2 Stimuli

3.2.2.1 Unprocessed recordings

The music stimuli considered here included two recordings: one classical sample and one rock sample. The classical sample was taken from “Overture to the Magic Flute” by W.A. Mozart and performed by the University Symphony Orchestra at the University of Colorado-Boulder. The rock sample was taken from “Anything At All” by Mere. The recorded stereo files were obtained directly from the recording engineers in their final mixes prior to the mastering stage, with no compression limiting applied. Samples of approximately 13 seconds in duration were selected from the recordings, at a point consistent with musical phrasing. Figure 3.2 shows the long-term spectra of the unprocessed music samples.

3.2.2.2 CL processing

A subset of conditions from Experiment 1 (Chapter 2) were included in this study. For CL processing, the samples were first normalized to have the same absolute peak level at -4 dB full scale (dBFS) and then were compressed using the Massey L2007 Mastering Limiter plug-in for Pro Tools 9. Six compression threshold conditions were originally created for each genre. See Chapter 2 for further details on CL processing. The current study included a subset of three conditions that are representative of the range of the original six compression thresholds: no compression (unprocessed), mild CL (-8 dBFS threshold), and heavy CL (-20 dBFS threshold). The stimuli were RMS-equalized following CL processing to reduce possibly confounding differences in input level. This step ensured that each compression-limited stimulus engaged the compressor in the simulated hearing aid to approximately the same extent.

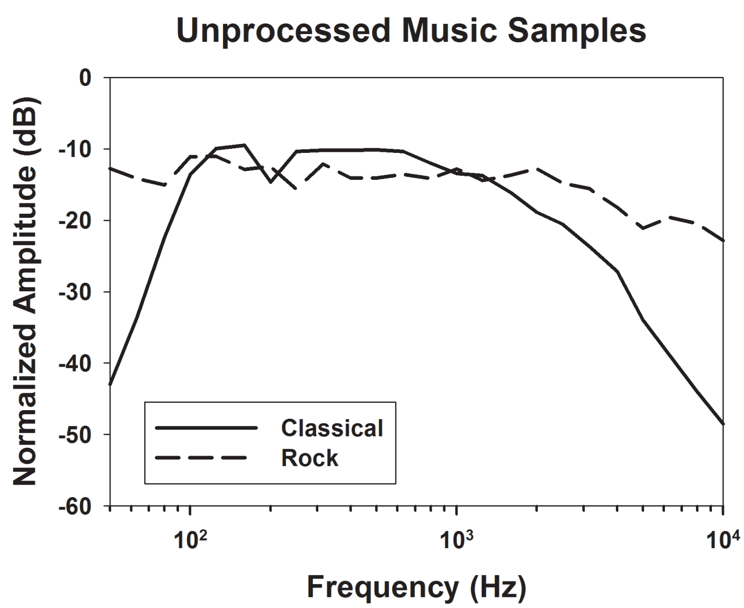

Figure 3.2: Spectra of music samples prior to processing.

3.2.2.3 WDRC processing

WDRC conditions were programmed using NAL-NL1 gain prescriptions (Byrne et al., 2001). Custom gains were calculated for each listener using the NAL-NL1 software included in the Audioscan Verifit hearing-aid analyzer. The WDRC parameters under investigation were the number of processing channels and release time. The compressor attack time was always 5 msec. Two processing-channel conditions were included: 3 channels and 18 channels. Two release-time conditions were included: 50 msec (fast) and 1000 msec (slow). The fast release time of 50 msec is comparable to release times implemented in hearing aids using syllabic compression, while 1000 msec was selected as an extreme value that has been used in a number of previous studies investigating the effects of a slow compressor (King and Martin, 1984; Neuman et al., 1995b, 1998). The combination of two processing-channel conditions and two release-time conditions resulted in four WDRC conditions. To examine the possibility that linear gain would be preferred to any combination of WDRC processing, two linear conditions were also included, using the NAL-R gain prescription (Byrne and Dillon, 1986) with 3 and 18 channels. Table 3.1 lists the conditions, numbered 1-18, that were used for each of the two genres.

The stimuli were presented at a nominal level of 65 dB SPL prior to amplification. The 65-dB SPL input level was determined based on pilot listening tests with eight hearing-aid users. The six input stimuli – classical unprocessed, classical with mild CL, classical with heavy CL, rock unprocessed, rock with mild CL, and rock with heavy CL – were played in the sound field through a GSI-61 audiometer. The hearing-aid users listened to each stimulus while wearing their personal hearing aids and adjusted the level to a chosen listening level. The mean chosen listening level was 65 dB SPL, with a standard deviation of 8.9 dB. Therefore, a 65-dB SPL input level was selected to represent an average chosen listening level that was likely to be within the range of comfortable listening for most subjects. See Appendix A for further details on the chosen listening levels procedure.

Hearing-aid simulations were implemented using MATLAB. The six stimuli resulting from

Table 3.1: Processing conditions numbered 1-18.

Condition	CL	Processing Type	Channels
1	None	Linear	3
2	Mild	Linear	3
3	Heavy	Linear	3
4	None	Linear	18
5	Mild	Linear	18
6	Heavy	Linear	18
7	None	Fast WDRC	3
8	Mild	Fast WDRC	3
9	Heavy	Fast WDRC	3
10	None	Fast WDRC	18
11	Mild	Fast WDRC	18
12	Heavy	Fast WDRC	18
13	None	Slow WDRC	3
14	Mild	Slow WDRC	3
15	Heavy	Slow WDRC	3
16	None	Slow WDRC	18
17	Mild	Slow WDRC	18
18	Heavy	Slow WDRC	18

the CL processing stage served as inputs to the simulated hearing aid. For the WDRC conditions, compression was performed using the frequency-warped compressor described by Kates and Arehart (2005), with a sampling rate of 22.05 kHz. A block size of 24 samples was used for processing, with 33 first-order all-pass filter sections and a 34-point FFT resulting in 18 frequency analysis bands. The center frequencies of the 18 bands are shown in Table 3.2. To create the 3-channel condition, subsets of the 18 channels were combined by summing the power in bands 1-6, 7-12, and 13-18. The signal power was summed across the 6 bands comprising each frequency group and then peak detected to set the signal level for the compressor. The peak-detected signal powers used to control the compression were, therefore, higher in the 3-channel processing than in the 18-channel processing, meaning that more of the signal was above the lower compression kneepoint in the 3-channel system than in the 18-channel system. The lower kneepoint of the compressor was 45 dB SPL, and the upper kneepoint was 100 dB SPL. Below 45 dB SPL, the simulated hearing aid provided linear gain. Above 100 dB SPL, the compression ratio was infinite. The compression ratio for sounds between 45 dB SPL and 100 dB SPL was individually determined by the NAL-NL1 gain prescription. WDRC was implemented between the two ears independently. The final stage of processing was to match the loudness of the amplified signals. The reference stimulus was specified as the 18-channel NAL-R condition. The loudness computed over the music selection was adjusted to match the estimated loudness of the reference condition, as calculated by the model of Moore and Glasberg (1997). The left and right channels were averaged for loudness estimation. The loudness-equalization step was intended to reduce the possible contribution of loudness to overall quality judgments.

3.2.3 Instrumentation and playout

Processed music samples were stored on a Dell Optiplex 990 computer as 24-bit, 22.05-kHz wave files. Stimuli were routed through a digital-to-analog converter (TDT RX8), an attenuator (TDT PA5), and a headphone buffer amplifier (TDT HB7) and were presented to listeners binaurally through Sennheiser HD-25 earphones. The output of each earphone was calibrated to an

Table 3.2: Center frequencies of analysis bands for warped compressor.

Band number	Center frequency, Hz
1	0
2	140
3	282
4	429
5	583
6	748
7	929
8	1129
9	1358
10	1625
11	1946
12	2345
13	2860
14	3555
15	4541
16	6001
17	8186
18	11025

average playout level of 65 dB SPL prior to the above described NAL-NL1 amplification. Listeners were seated in a sound-treated booth. For one listener (S11), the amplified signals were uncomfortably loud and were attenuated by 9 dB prior to playout using the headphone buffer.

3.2.4 Procedure

Listeners made forced-choice paired-comparison preference judgments within each genre using a customized computer interface. For each trial, listeners heard Sample A, followed by Sample B, and they were asked to select which of the two samples they preferred. Each sample represented one of the 18 processing conditions, in randomized order, and listeners rated all possible combinations of conditions (except for any condition against itself), for a total set of 153 comparisons. Each set of 153 comparisons was performed twice, resulting in four total sets of comparisons, including rock and classical. Each condition was presented once in position A and once in position B for a given comparison. At each appointment, listeners completed a practice block of trials to become familiarized with the task and some of the conditions. Paired comparison judgments were completed in 24 test blocks, spread over four or five appointments for a total of approximately eight hours of testing.

3.2.5 Psychophysical measures

In addition to preference judgments, listeners completed three psychophysical tasks. Each psychophysical task had a primary purpose in relation to a specific aspect of suprathreshold processing, including loudness perception, temporal processing, and spectral processing. Because listeners had symmetric hearing loss, the psychophysical tasks were performed for the right ear only. Each measure is described below.

3.2.5.1 Contour Test of Loudness Perception

The first psychophysical task was the Contour Test of Loudness Perception (Cox et al., 1997). During the Contour Test, subjects provided categorical ratings of loudness for warble tones. The categories were 1 (very soft), 2 (soft), 3 (comfortable, but slightly soft), 4 (comfortable), 5 (comfortable, but slightly loud), 6 (loud, but o.k.) and 7 (uncomfortably loud). The test was performed through a GSI-61 audiometer at test frequencies of 500 Hz and 2000 Hz and was conducted using ER-3A insert earphones according to the instructions provided by Cox et al. (1997). Four ascending runs were completed at each test frequency, and the median of the four runs was calculated for each loudness category. The results of the Contour Test provided information about a given listener's dynamic range, defined here as the difference in dB between the levels for category 1 (very soft) and category 7 (uncomfortably loud).

3.2.5.2 Amplitude modulation depth discrimination

The second psychophysical task was amplitude modulation depth discrimination, which was performed using the procedure applied by Sabin et al. (2012), similar to the methods described by Bacon and Viemeister (1985). The task was implemented using a MATLAB program. The signal was a speech-shaped noise carrier, which was modulated by a cosine wave at a 4-Hz modulation rate. The depth of modulation was varied by adjusting the amplitude of the cosine modulator. The modulation discrimination test was performed using a three-interval forced choice paradigm, in which listeners selected which interval was different from the reference signal. The program used a 2-down, 1-up adaptive procedure to track the 70.7% correct point on the psychometric function (Levitt, 1971). Sounds were played using a Dell Optiplex 990 computer and TDT RX8, PA5, and HB7 and were presented to listeners monaurally through Sennheiser HD-25 earphones at a level of 86 dB SPL. Sixty trials were presented, and the threshold was the mean of all reversals except for the first four (if an even number of reversals) or the first three (if an odd number of reversals). The threshold is described in terms of dB, calculated as 20 log (m), where m is the threshold modulation index.

3.2.5.3 Sweeping Psychophysical Tuning Curve (SWPTC)

The third psychophysical task, the SWPTC (Sek et al., 2005; Sek and Moore, 2011), measured frequency selectivity by deriving a psychophysical tuning curve. Sounds were played through the headphone jack of a Dell Optiplex 990 computer, calibrated according to the instructions provided by Sek and Moore (2011), and were presented monaurally through Sennheiser HD-25 earphones. A pulsed pure tone signal was presented in the presence of a continuous masker. The signal was 200 msec in duration with a 200-msec inter-stimulus interval. The pure tone signal was presented at 10 dB sensation level (SL), as measured using the threshold estimation procedure provided in the SWPTC software. The masker was a narrowband noise that was swept either upward (forward sweep) or downward (reverse sweep) in frequency. The level of the masker increased when the signal was audible and decreased when the signal was inaudible. The listener indicated whether or not the signal could be heard using a Békésy-style procedure, by holding down the space bar when the tone was audible and releasing the spacebar when the tone was inaudible. SWPTC measurements were performed for test frequencies of 500 Hz and 2000 Hz, including two runs of forward sweeping noise and two runs of reverse sweeping noise at each frequency. Default settings were used for the starting and ending frequencies of the masker and the masker bandwidth. Specifically, for the 500-Hz signal, the lowest center frequency of the masking noise was 250 Hz, the highest center frequency was 750 Hz, and the bandwidth was 100 Hz. For the 2000-Hz signal, the lowest center frequency of the masking noise was 1000 Hz, the highest center frequency was 3000 Hz, and the bandwidth was 320 Hz. A 1-dB/sec rate of change was used for the noise level at both frequencies. For listeners with steeply sloping losses, a continuous low-pass noise was presented to prevent the detection of combination cues. In the results files provided by the SWPTC software, the Q10 value for the auditory filter corresponding to the tuning curve was given using five different calculation methods (double-linear regression, moving average, quadratic function, lowpass filtering, and a rounded-exponent function). Q10 is measured by dividing the center frequency of the auditory filter by the bandwidth of the filter, measured at 10 dB below the peak. Therefore, higher Q10 values indicate sharper tuning.

3.3 Results

3.3.1 Acoustic outcomes

Several analyses were performed to quantify the acoustic effects of CL, WDRC, and CL+WDRC. These measurements were taken for two reasons: 1) to describe the specific outcomes of the processing algorithms on the acoustic structure of the music stimuli, and 2) to provide a basis for the interpretation of preference ratings in the context of acoustic features. To include the WDRC processing, stimuli were generated using the simulated hearing-aid described above. The hearing thresholds for generating the stimuli were based on the average hearing loss of the 18 listeners. Refer to Figure 3.1, solid line, for the average audiogram. Stimuli were converted from stereo to mono for the acoustic analysis.

3.3.1.1 Amplitude histograms

Amplitude histograms measure the wideband amplitude characteristics across the entire duration of a signal. Amplitude histograms provide a "snapshot" of the distribution of levels across all frequencies present in the stimulus. Figure 3.3 shows amplitude histograms for four representative conditions, using the rock music as an example. The y-axis of each histogram illustrates the proportion of samples that occur at a particular level, with level represented on the x-axis and expressed in relation to RMS. In each panel of Figure 3.3, the reference (solid line) is condition 4 (linear, 18-channel processing with no CL), and the dashed line shows the comparison condition. The top two panels compare condition 4 to slow WDRC, with 3-channel processing on the left and 18-channel processing on the right. The bottom two panels show fast WDRC with 18 channels, with the bottom left panel showing no CL and the bottom right panel showing heavy CL. Each curve defines the proportion of samples that occur within a specific bin, using 51 histogram bins at equal increments to span the range of -40 to +30 dB re: RMS of the reference stimulus.

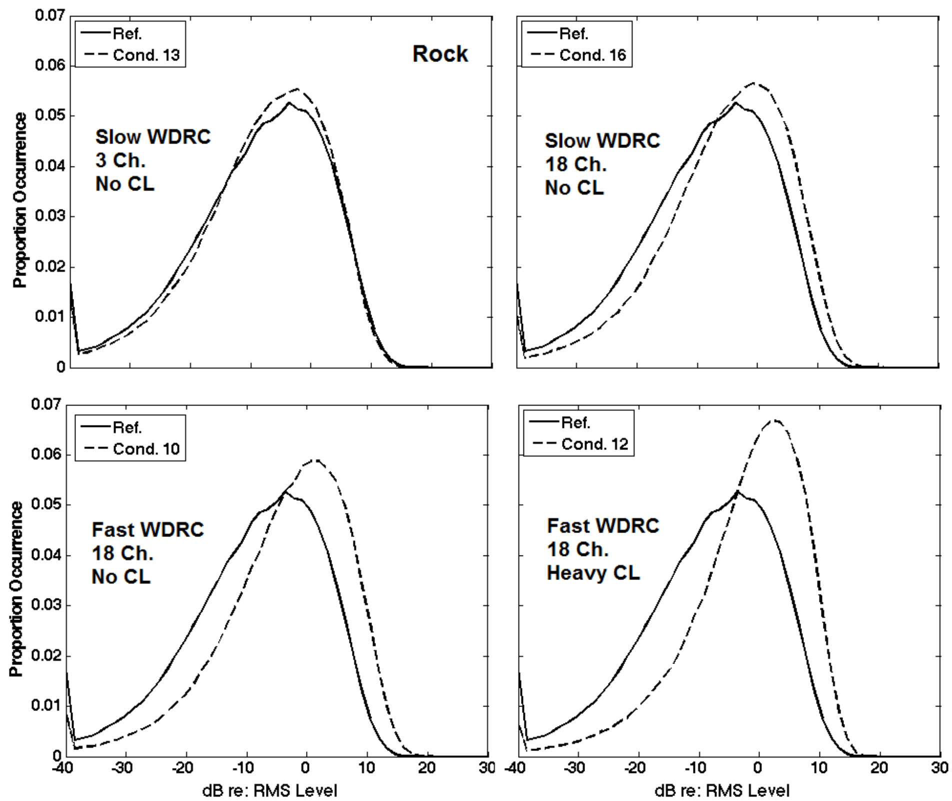

Figure 3.3: Amplitude histograms comparing the reference (condition 4: linear, 18-channel processing with no CL) to four representative conditions with the rock stimulus. In each panel, the reference stimulus is the solid line, and the comparison stimulus is the dashed line. The comparison conditions are slow WDRC with 3 channels and no CL (condition 13; top left), slow WDRC with 18 channels and no CL (condition 16; top right), fast WDRC with 18 channels and no CL (condition 10; bottom left), and fast WDRC with 18 channels and heavy CL (condition 12; bottom right).

The amplitude histograms of these four representative conditions indicate that with increasing amounts of compression, the amplitude variation within each stimulus is reduced. The decrease in dynamic range can be observed in the narrowing shape of the curve. Each layer of complexity (going from 3 to 18 channels, slow to fast WDRC, and no CL to heavy CL) further reduces the dynamic range of the stimulus. To quantify the extent of the dynamic-range reduction, the crest factor for each stimulus was measured, defined as the ratio of the amplitude exceeded 99% of the time to the RMS level, and expressed in dB. The measured crest factors for the compressed stimuli in Figure 3.3 (dashed lines) were 9.99 and 9.93 dB for the top two panels (left and right, respectively) and 9.70 and 8.98 dB for the bottom two panels (left and right). The crest factor for the reference condition (solid line) was 10.25 dB. As expected, the crest factors of the music samples decreased due to compression, with increasing amounts of compression causing a further decrease in crest factor. The trends for classical music were the same as for rock music. The amplitude histograms for all conditions and both genres are shown in Appendix B.

3.3.1.2 Dynamic-range difference plots

Although the amplitude histograms plot the overall dynamic range of a stimulus, they do not adjust for differences that might occur in various frequency regions. Therefore, a further analysis of the dynamic range of the music signals was performed in a perceptually relevant manner, taking into account auditory filter bandwidths for the hearing loss shown in Figure 3.1, solid line. The bandwidths were estimated with gammatone filters (Patterson et al., 1995) and were used as separate analysis bands for calculating cumulative envelope level distributions. For example, the 98% cumulative level distribution represents the level that is not exceeded 98% of the time within the output of the estimated auditory filter. The 30% cumulative level distribution represents the level that is not exceeded 30% of the time. The difference between the 98% level distribution and the 30% level distribution constitutes the dynamic range across frequencies for a given stimulus.

Differences in dynamic range between stimuli are shown in Figure 3.4. The dynamic-range difference between two signals was calculated by subtracting the dynamic range of the comparison stimulus from that of the reference stimulus (condition 4). Dynamic-range difference values were plotted as a function of frequency, so that positive values indicate a greater dynamic range for the reference stimulus and negative values indicate a greater dynamic range for the compressed stimulus. In the top left panel of Figure 3.4, the comparison stimulus is condition 13 for classical music (slow WDRC with 3 channels and no CL). In the top right panel, the comparison stimulus is condition 10 for classical music (fast WDRC with 18 channels and no CL). Both panels show that the compressed music samples have a reduced dynamic range in the higher frequencies, compared to the reference linear condition, as indicated by the positive values in the higher-frequency regions. The bottom two panels of Figure 3.4 demonstrate the effects of CL on the dynamic range of the stimuli, for both classical (bottom left) and rock (bottom right) music. In both cases, heavy CL causes a substantial decrease in dynamic range across frequencies. For classical music, the effect is relatively flat. For rock music, the high- and low-frequency regions show greater changes than the mid-frequency region. This effect is likely due to the flatter spectrum of the unprocessed rock stimulus, which has a large dynamic range in the low and high frequencies, compared to the classical stimulus. Because these frequency regions start with a larger dynamic range, the changes they incur due to CL are subsequently greater. The dynamic-range difference plots for all conditions and both genres are shown in Appendix B.

3.3.1.3 Modulation spectra

Another acoustic property affected by compression is the modulation spectrum. To understand the modulation spectrum, one must differentiate modulation frequencies from acoustic frequencies. A modulation spectrum represents dynamic level variations over time (i.e. modulation frequencies) within an acoustic frequency region. A helpful example is provided by Souza and Gallun (2010). Consider a 1000-Hz carrier tone modulated by a sine wave with a frequency of 4 Hz. The acoustic spectrum would show a component centered at 1000 Hz with modulation sidebands at 996 and 1004 Hz. The modulation spectrum would have a single component at 4 Hz. If the carrier tone were fully modulated by the sine wave so that its envelope has a periodic minimum of 0, it would

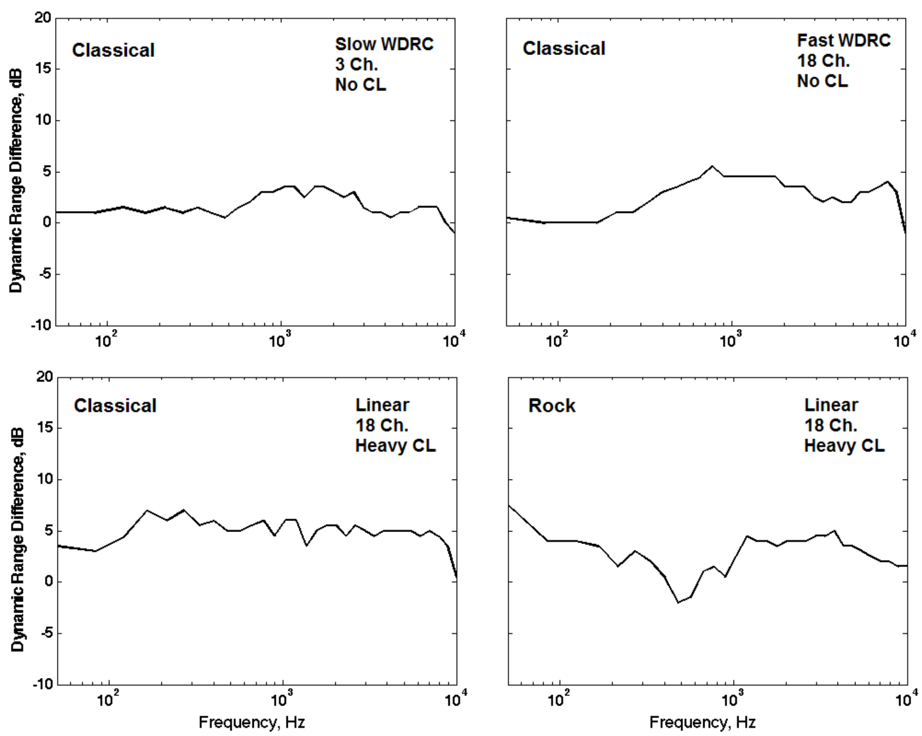

Figure 3.4: Dynamic-range difference plots comparing the reference (condition 4: linear, 18-channel processing with no CL) to four representative conditions. Envelope distributions are analyzed within auditory filter bandwidths and represented across frequencies from 50 to 10,000 Hz. Dynamic range is defined as the difference between the 30% distribution and the 98% distribution for each stimulus. Curves are derived by subtracting the dynamic range of the comparison stimulus from the dynamic range of the condition 4 stimulus.

be described as 100% modulated (with a modulation index of 1). Dynamic-range compression reduces differences between peaks and valleys in the temporal envelope. Therefore, compression may reduce the magnitude of modulation at the modulation frequencies that are present within a signal. Because the modulation spectrum of music corresponds to the tempo and rhythmic structure, the distinctiveness of the rhythm could potentially be reduced if the magnitude of the modulation is flattened.

To conduct a modulation spectrum analysis for the current study, the signal envelope was divided into 8-msec segments having a 50% overlap. Each segment was windowed using a raised cosine window, summed, and then divided by the window sum to produce the average. The smoothed envelope was then passed through a modulation filter bank with filter center frequencies from 1 to 32 Hz, using 1/3-octave spacing. The filter output was divided by the filter bandwidth to give the power spectral density. The resulting modulation spectra show a small effect of compression. Specifically, with increasing amounts of compression, the rhythmic structure of the music samples becomes less distinguishable. As shown in Figure 3.5 (top left panel), the peak of the modulation spectrum for the classical stimulus is at about 2.67 Hz, corresponding to 160 bpm. The surrounding peaks are at half and double this rate, demonstrating the rhythm of the music (e.g. half notes, eighth notes). When the classical stimulus is heavily compressed (top right panel), the primary tempo can still be observed in the peak at 2.67 Hz, but the difference between the primary peak and surrounding peaks is less pronounced. Similarly, the peak of the modulation spectrum for the linear rock stimulus is at about 1.67 Hz (Figure 3.5, bottom left panel), corresponding to 100 bpm. With a heavy amount of compression (bottom right panel), the peak at double the tempo (about 3.34 Hz) becomes blurred. The modulation spectra for all conditions and both genres are shown in Appendix B.

3.3.1.4 HASQI

The stimuli were also analyzed using a modified version of the Hearing Aid Speech Quality Index (HASQI), developed by Kates and Arehart (2010). HASQI uses an auditory model including

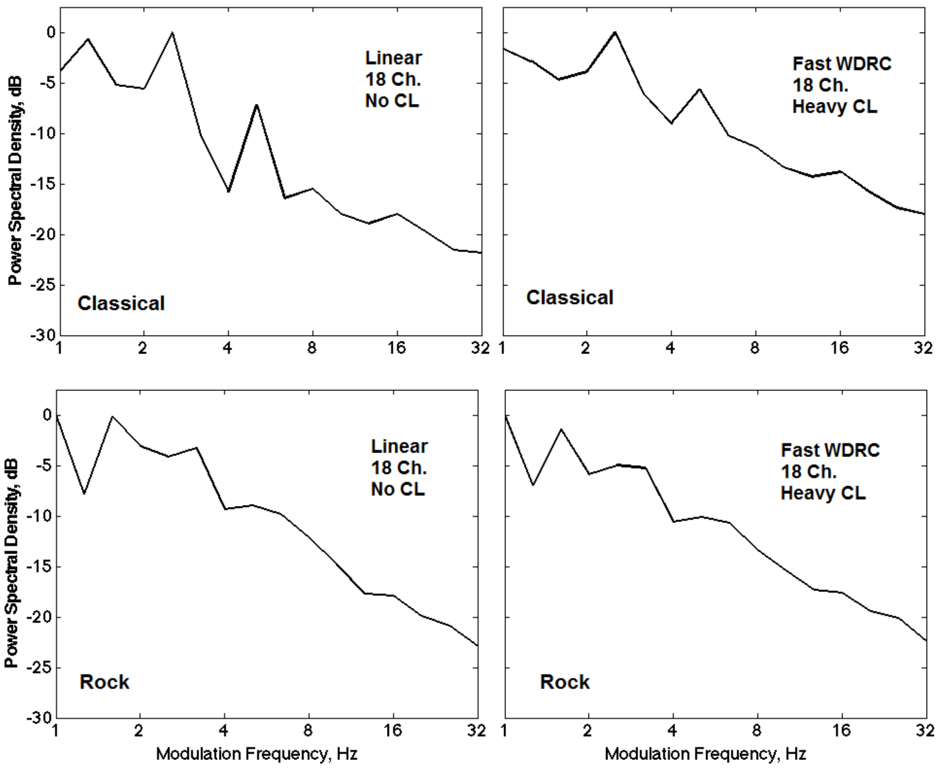

Figure 3.5: Modulation spectrum analysis for four representative stimuli. The left panels show the linear 18-channel condition with no CL (condition 4) for classical (top) and rock (bottom). The right panels show the fast WDRC condition with 18 channels and heavy CL (condition 12) for classical (top) and rock (bottom). Peaks of the modulation spectrum represent the rhythmic structure of the music. For the heavily compressed stimuli (right panels), the rhythmic structure is less distinct.

cochlear hearing loss to give a minimum mean-squared error fit to listener quality judgments. The metric was originally developed using speech stimuli, although the current version of HASQI was adapted for music by computing new weights to fit the quality ratings reported by Arehart et al. (2011). HASQI contains three terms: nonlinear, linear, and combined. The nonlinear component of HASQI compares the envelope time-frequency modulation of a processed stimulus to an unprocessed stimulus when the long-term spectra of the two stimuli are matched. The linear component of HASQI compares the long-term spectrum of a processed signal relative to that for the unprocessed signal. Each term provides a value between 0 (very low fidelity) and 1 (perfect fidelity). The combined HASQI term quantifies the total amount of signal modification by calculating the product of the nonlinear and linear components. Table 3.3 reports HASQI figures for all conditions for the classical and rock stimuli. In general, the nonlinear HASQI terms are more greatly affected by compression than the linear HASQI terms. The lowest HASQI values are given for fast WDRC combined with heavy CL.

Table 3.3: HASQI values for all conditions with classical and rock music.

	Classical			Rock		
Condition	HASQI Nonlinear	HASQI Linear	HASQI Combined	HASQI Nonlinear	HASQI Linear	HASQI Combined
1	0.96	0.82	0.79	0.97	0.84	0.82
2	0.96	0.82	0.79	0.97	0.84	0.81
3	0.53	0.82	0.44	0.64	0.85	0.55
4	1.00	1.00	1.00	1.00	1.00	1.00
5	1.00	1.00	0.99	1.00	1.00	0.99
6	0.53	0.96	0.51	0.64	0.95	0.61
7	0.77	0.77	0.59	0.82	0.81	0.66
8	0.76	0.77	0.59	0.82	0.81	0.66
9	0.42	0.77	0.32	0.55	0.81	0.45
10	0.79	0.83	0.65	0.84	0.87	0.74
11	0.78	0.82	0.64	0.84	0.87	0.73
12	0.40	0.81	0.33	0.54	0.87	0.47
13	0.83	0.78	0.65	0.87	0.81	0.71
14	0.83	0.78	0.65	0.87	0.81	0.70
15	0.47	0.78	0.37	0.59	0.83	0.49
16	0.82	0.84	0.70	0.87	0.89	0.78
17	0.82	0.84	0.69	0.86	0.89	0.77
18	0.46	0.84	0.38	0.58	0.89	0.52

3.3.2 Paired comparisons: average data

Paired comparisons were transformed into preference scores using the methods described in Rabiner et al. (1969) and applied by Neuman et al. (1995b) and Arehart et al. (2007). The total number of times a condition was preferred was divided by the total number of times that condition was presented, for each participant. For example, the number of presentations for each condition was 34 per subject (17 presentations x 2 repeats). If a particular condition was selected 20 times total, its preference score would be $20/34 = 0.59$. Preference scores ranged from 0 to 1 and represented the proportion of times each condition was preferred.

The first experimental question examined whether CL and WDRC affected preference scores, across all listeners included in the sample. To answer this question, preference scores were arcsin transformed and subjected to repeated-measures analysis of variance (ANOVA) using the general linear model procedure in SPSS 21.0. A separate repeated-measures ANOVA was performed for each genre with three within-subjects variables: WDRC processing type (linear, slow WDRC, and fast WDRC), number of processing channels (3 and 18), and amount of input CL (none, mild, and heavy). Greenhouse-Geisser corrections were used when Mauchly's test of sphericity was significant. The preference scores for classical music are shown in the top row of Figure 3.6. From left to right, the panels illustrate the effects of the WDRC processing conditions within each level of the CL variable (no CL, mild CL, and heavy CL, respectively). The three lines represent the preference scores for linear, slow WDRC, and fast WDRC processing, as a function of the number of channels. Figure 3.6 shows similar preference scores for linear processing and slow WDRC, with fast WDRC clearly rated lower in all conditions. For classical music, the effect of WDRC was significant [$F(1.4, 23.2)=11.278$; $p=0.001$; partial $\eta^2=0.399$]. Specifically, pairwise comparisons with Bonferroni corrections showed that slow WDRC was not different from linear processing ($p=1.0$), but fast WDRC was significantly less preferred than both linear ($p=0.012$) and slow WDRC ($p<0.001$). Additionally, the effect of CL was significant [$F(1.2, 19.8)=55.125$; $p<0.001$; partial $\eta^2=0.764$]. Heavy CL was rated lower than both mild CL ($p<0.001$) and no CL ($p<0.001$), but no difference was found between no CL and mild CL ($p=1.0$). The main effect of number of channels was not significant [$F(1, 17)=3.966$; $p=0.063$; partial $\eta^2=0.189$]. None of the interactions among processing conditions were significant for classical music.

The preference scores for rock music are shown in the bottom row of Figure 3.6. Similarly

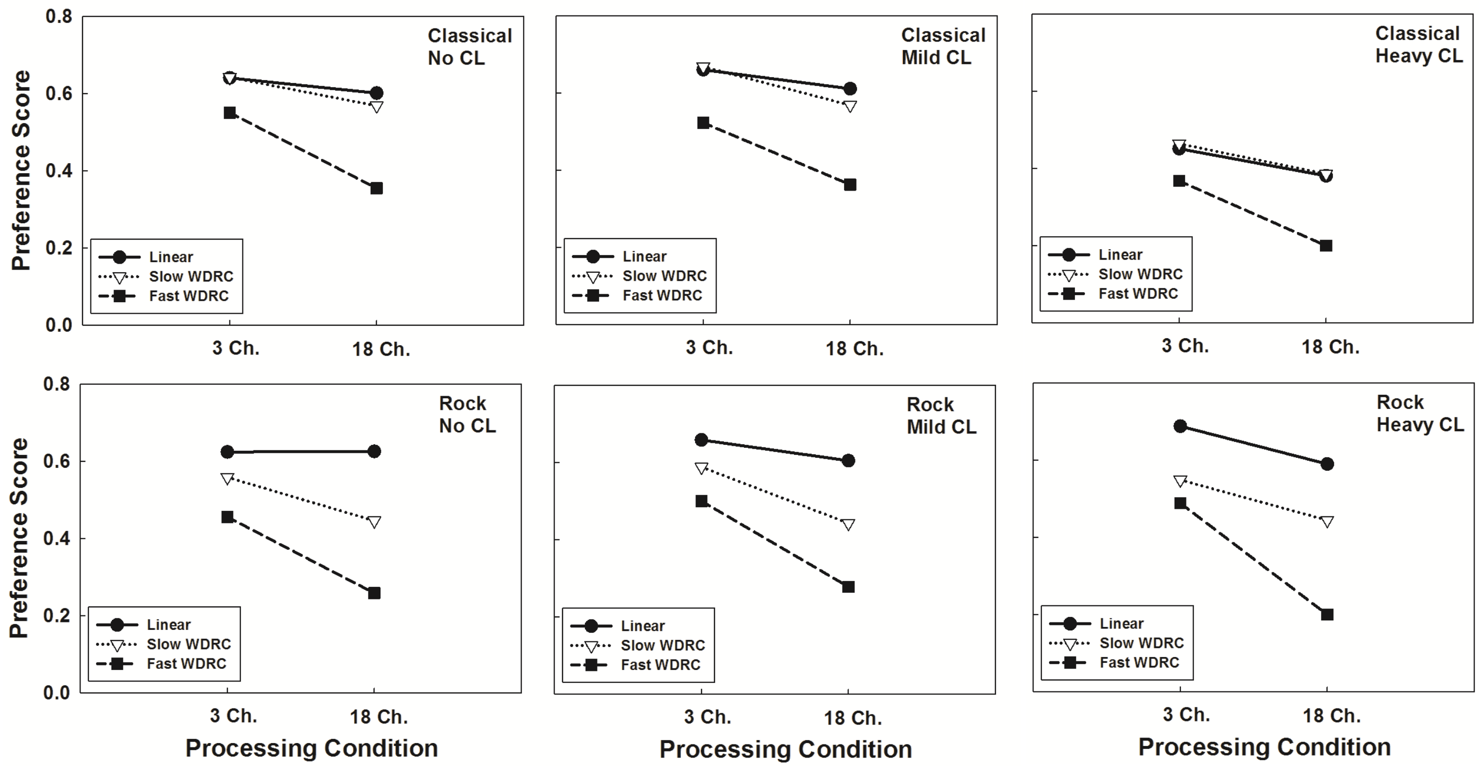

Figure 3.6: Scores for paired comparison preference ratings of the 18 processing conditions. Preference scores indicate the proportion of times each condition was preferred across all subjects. The top row shows classical music, and the bottom row shows rock music. From left to right, each panel represents no CL, mild CL, and heavy CL. Linear processing is represented by the solid lines, slow WDRC by the dotted lines, and fast WDRC by the dashed lines, each as a function of the number of processing channels. Standard errors were very small for each condition (0.004 or below) and are not shown on the graphs.

to classical music, the main effect of WDRC was significant [$F(1.3, 22.2)=26.471$; $p<0.001$; partial $\eta^2=0.609$]. However, unlike for classical music, the linear processing condition was preferred to both slow WDRC ($p=.008$) and fast WDRC ($p<0.001$), while fast WDRC was again rated lower than slow WDRC ($p<0.001$). Another difference for the rock music analysis was that the main effect of channels was also significant, indicating that the 3-channel condition was preferred overall to the 18-channel condition [$F(1, 17)=9.678$; $p=0.006$; partial $\eta^2=0.363$]. The interaction between WDRC condition and channels was significant, due to the fact that the difference between channels was largest for fast WDRC [$F(1.5, 25.5)=10.523$; $p=0.001$; partial $\eta^2=0.382$]. The interaction between channels and CL was also significant [$F(2, 34)=5.771$; $p=0.008$; partial $\eta^2=0.253$], indicating that the difference between channels was largest for heavy CL. The main effect of CL was not significant for rock music [$F(1.1, 19.4)=0.252$; $p=0.652$; partial $\eta^2=0.015$].

3.3.3 Individual variability

3.3.3.1 Paired comparisons: interactions

A second experimental goal was to identify listener-specific factors related to music preference ratings. Listeners were divided into several groups based on the psychophysical measures described above and a number of personal characteristics. The between-subjects factors included gender; musician vs. non-musician; average hearing loss at 500, 1000, 2000, and 4000 Hz; dynamic range at 500 Hz; dynamic range at 2000 Hz; psychophysical tuning curve bandwidth (Q10) at 500 Hz; psychophysical tuning curve bandwidth (Q10) at 2000 Hz; and modulation depth discrimination threshold in dB = 20 log (m). Subjects were divided into groups as follows: PTA <45 dB HL (9 subjects) and PTA ≥ 45 dB HL (9 subjects); dynamic range at 500 Hz ≤ 45 dB (7 subjects) and dynamic range at 500 Hz ≥ 50 dB (11 subjects); dynamic range at 2000 Hz ≤ 30 dB (6 subjects), between 31 and 38 dB (8 subjects), and ≥ 45 dB (4 subjects); Q10 value at 500 Hz <3.0 (12 subjects) and ≥ 3.0 (6 subjects); Q10 value at 2000 Hz <3.0 (13 subjects) and ≥ 3.0 (5 subjects); modulation depth discrimination threshold >-10 dB (10 subjects) and modulation depth discrimination threshold <-10 dB (7 subjects). One subject did not complete the modulation depth task. For three subjects, it was not possible to derive a Q10 value from either the 500-Hz or 2000-Hz SWPTC measurement because the shape of the curve was too broad or abnormal; therefore, these subjects were placed in the group with broader tuning curve measurements. Figure 3.7 shows the individual data for the psychophysical measurements.

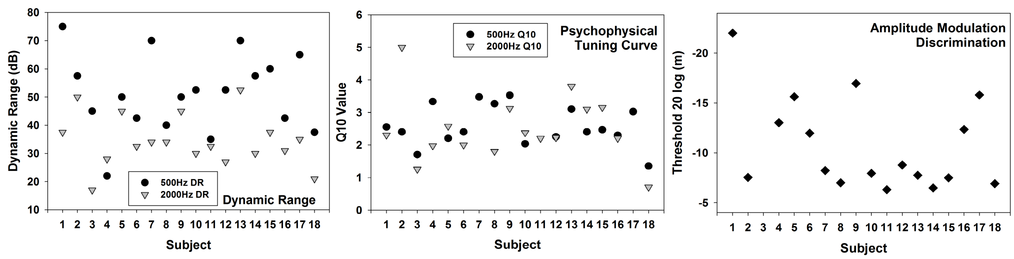

Figure 3.7: Scatterplots showing individual data for dynamic-range measurements (left panel), psychophysical tuning curve Q10 values (middle panel), and modulation depth discrimination thresholds (right panel).

To examine individual variability in preference ratings, separate mixed-model ANOVAs were conducted for each genre, using the three within-subjects variables (WDRC processing type, number of processing channels, amount of input CL) and each of the between-subjects factors described above. Two of the mixed-model analyses showed significant processing-by-group interactions. For classical music, Q10 value at 500 Hz had significant interactions with WDRC processing [F(2, 32)=5.138; p=0.022; partial η^2=0.243] and the number of processing channels [F(1, 16)=6.291; p=0.023; partial η^2=0.282]. As illustrated in the left panel of Figure 3.8, listeners with broader tuning curves at 500 Hz preferred 3-channel processing to 18-channel processing, while the opposite was true for listeners with narrower tuning curves. Additionally, listeners with broader tuning curves preferred linear processing to slow WDRC, but listeners with narrower curves preferred slow WDRC to both linear processing and fast WDRC (middle panel of Figure 3.8).

For rock music, the interaction between dynamic range at 500 Hz and WDRC processing was significant [F(2, 32)=3.714; p=0.035; partial η^2=0.188]. Specifically, the group with a larger

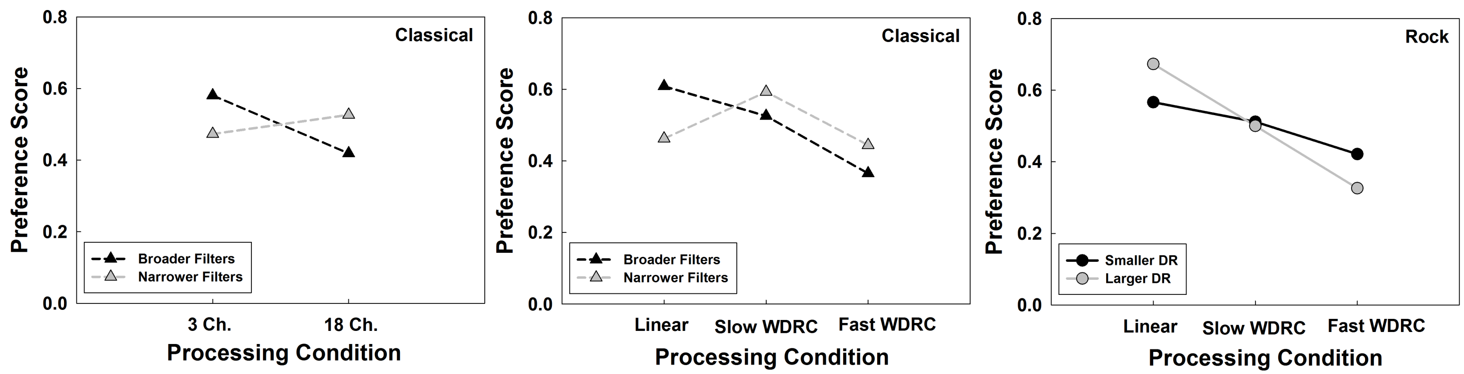

Figure 3.8: Preference scores by group for the significant interactions resulting from the repeated-measures ANOVAs. The left and middle panels show the interaction between Q10 value at 500 Hz and the number of processing channels (left) and WDRC processing condition (middle) for classical music. Listeners with broader auditory filters are represented by the black dashed lines and listeners with narrower auditory filters are represented by the gray dashed lines. The right panel illustrates the interaction between dynamic range at 500 Hz and WDRC processing condition for rock music. Listeners with a smaller dynamic range are represented by the solid black line and listeners with a larger dynamic range are represented by the solid gray line. Standard errors were very small for each condition (0.004 or below) and are not shown on the graphs.

dynamic range had larger differences in their preference for linear processing over fast WDRC (see Figure 3.8, right panel). Individual preference data are shown in Figures C.1 through C.6 in Appendix C.

3.3.3.2 Multidimensional unfolding

Multidimensional unfolding (Borg and Groenen, 2005; Coombs, 1950) is a special case of multidimensional scaling that considers preferential choice data. It represents distances among choice objects and individuals. In the current study, the choice objects are processing conditions. All individuals are assumed to operate within the same psychological space. In the psychological space, multiple dimensions may drive preferences, and stimuli are plotted in locations computed from their preference scores. Each individual has an ideal point in the space that represents the best combination of attributes for that individual. The proximity between an individual's ideal point and any given stimulus location equates to the preference score given by the individual for that stimulus. Multidimensional unfolding is included here as a way of uncovering individual variability and exploring perceptual dimensions that may underlie preference ratings on a group and individual level.

For each listener, the preference score for each condition, given as the proportion of times that condition was chosen chosen between 0 and 1, was used as a basis for conducting the analysis. Preference scores were converted to dissimilarities using the method described by Borg and Groenen (2005). A value of 0.5 was subtracted from each preference score, and the absolute value of the difference served as the dissimilarity value. The PREFSCAL algorithm for SPSS 21.0 was used to conduct the analysis from a listeners-by-condition matrix containing the dissimilarities. PREFSCAL was performed with a classical scaling start to complete the multidimensional scaling matrix using Spearman distances. No transformations were applied to the dissimilarities. Default values were used for the strength ($\lambda = 0.5$) and range ($\omega = 1.0$) of the penalty, which penalizes equal distances to reduce the possibility of a degenerate result (a solution that is trivial, although mathematically correct).

For both rock and classical music, a two-dimensional solution was selected based on several measures. Badness-of-fit is described using normalized stress. Figure 3.9 plots normalized stress as a function of the number of dimensions in the solution. As shown in Figure 3.9, the badness of fit improved going from two dimensions to three and four dimensions, but then no further improvement was seen with additional dimensions. Although normalized stress decreased with more dimensions, the amount of change by adding more dimensions was small because the stress values for the two-dimensional solutions were already rather small. For the two-dimensional solutions, DAF was high for both classical (0.84) and rock (0.89). DAF is analagous to variance and is used in conjunction with normalized stress. The ease of interpretation of the two dimensions over a greater number of dimensions was a final reason for selecting the two-dimensional solutions.

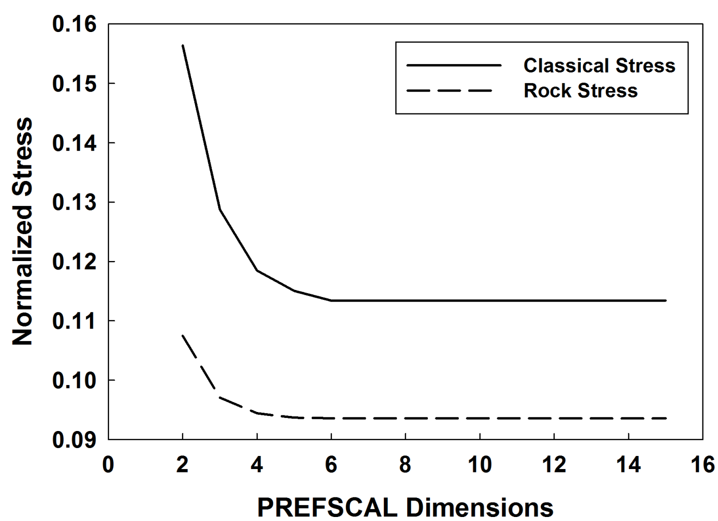

Figure 3.9: Normalized stress plotted as a function of the number of dimensions included in the PREFSCAL algorithm.

Several aspects of the unfolding analysis shed insight on listeners' preference judgments. The first aspect has to do with the positioning of the stimuli in the psychological space. Interpretation of the two dimensions was performed by conducting a correlational analysis between the stimulus coordinates and several known acoustic features of the stimuli (crest factor, average sensation level, nonlinear HASQI, linear HASQI, combined HASQI, amount of CL, number of WDRC channels, and WDRC release time). Sensation level was estimated using excitation patterns, as described by Moore and Glasberg (1997). Each listener's excitation pattern was calculated for the reference condition (no CL, linear with 18 channels). Excitation patterns indicated that all listeners had audibility of both classical and rock stimuli through at least 2100 Hz for the reference stimulus (see Appendix C, Figures C.7 and C.8 for individual listeners' excitation patterns). Other stimuli provided higher sensation levels than the reference. An additional excitation pattern analysis was conducted using the average audiogram (see Figure 3.1, solid line) in order to see differences in the mean sensation level for the different processing conditions. Sensation level was averaged across all frequencies. The results of this excitation pattern analysis showed that the fast WDRC condition with 18 channels had almost a 3-dB higher average sensation level than the reference condition.

Pearson's and Spearman's correlations were measured between each acoustic feature and the stimulus coordinates for the two psychological dimensions. For Dimension 1, the strongest relationship was seen for average sensation level for both classical and rock, using Pearson's correlations ($r=0.438$ for classical; $r=0.537$ for rock). For Dimension 2, both classical and rock music showed the strongest relationships using Spearman's correlations. The highest correlation for classical was seen for the linear HASQI value (ρ=-0.544). For rock music, the strongest relationship with Dimension 2 was seen for average sensation level (ρ=-0.356), and the next strongest relationship was seen for linear HASQI (ρ=-0.304). Figure 3.10 plots the stimulus locations within the two-dimensional space, including vectors representing the best fitting stimulus properties (interpreted across both genres). The stimulus coordinates were multiplied by a factor of 3 to better illustrate the relationships betwen stimulus locations and stimulus properties (Harvey and Gervais, 1981). Each vector illustrates the value of the correlation for that stimulus property along the two axes. In the left panel of Figure 3.10 (classical), the vector for average sensation level follows Dimension 1, and the vector for linear HASQI follows Dimension 2. The same two vectors are represented for rock music in the right panel.

Figure 3.11 shows the two-dimensional spaces for classical music (left panel) and rock music

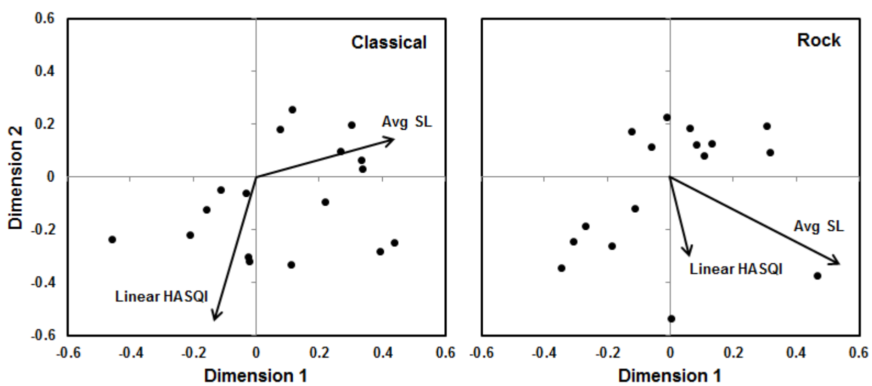

Figure 3.10: Results of a correlational analysis to interpret the two-dimensional psychological space. The strongest correlations between stimulus properties and stimulus coordinates are represented as vectors superimposed on the space, with the coordinate of the vector representing the value of the correlation. Dimension 1 correlated best with average sensation level, and Dimension 2 correlated best with the linear term of HASQI.

(right panel). The stimuli, shown in black circles, are positioned relative to their perceived attributes along the two dimensions. As can be seen in Figure 3.11, the stimuli generally follow a pattern along Dimension 1 moving from a lower sensation level to a higher sensation level, from left to right. From the bottom to the top of Dimension 2, the stimuli approximately follow a pattern moving from greater spectral flattening to less spectral flattening. One outlier stimulus is seen in the top-right area of the Classical space (left panel, Figure 3.11). This stimulus was the least preferred processing – Condition 12 (heavy CL, fast WDRC with 18 channels), and its stimulus location is the farthest from the ideal points of all of the listeners. However, the location of this stimulus point along Dimension 2 does not correspond with the other stimuli, which seem to follow a pattern of less spectral smearing toward the top of the space. The explanation for this outlying stimulus location is unclear. Nonetheless, the "spectral modification" interpretation of Dimension 2 accounts for the preferential data for most processing conditions, particularly for the classical music. The classical sample was dominated by strings in its instrumentation. For stringed instruments, the balance between low-frequency and high-frequency harmonics is an important factor in sound quality. Processing conditions that upset this harmonic content might cause stringed instruments to sound artificial or unpleasant. Consistent with this explanation, several listeners subjectively reported that certain conditions made the violins sound "tinny," which influenced their preference choices. While Dimensions 1 and 2 seem to have similar meaning for the rock music, Dimension 1 is more strongly related to the stimulus positions than Dimension 2. This finding indicates that spectral content might contribute more to preferences for classical music in the current listener group, while sensation level might contribute more to preferences for rock music.

A similar correlational analysis as before was used to determine potential explanations for the individual variability of the listeners. Pearson's and Spearman's correlations were measured between the individual ideal-point coordinates and each listener's gender, age, pure-tone average, musician group, dynamic range, psychophysical tuning curve, and modulation detection threshold. The factor having the strongest relationship with listeners' ideal points was estimated auditory filter bandwidth (Q10 value), which corresponded to Dimension 2. The ideal points of the listeners

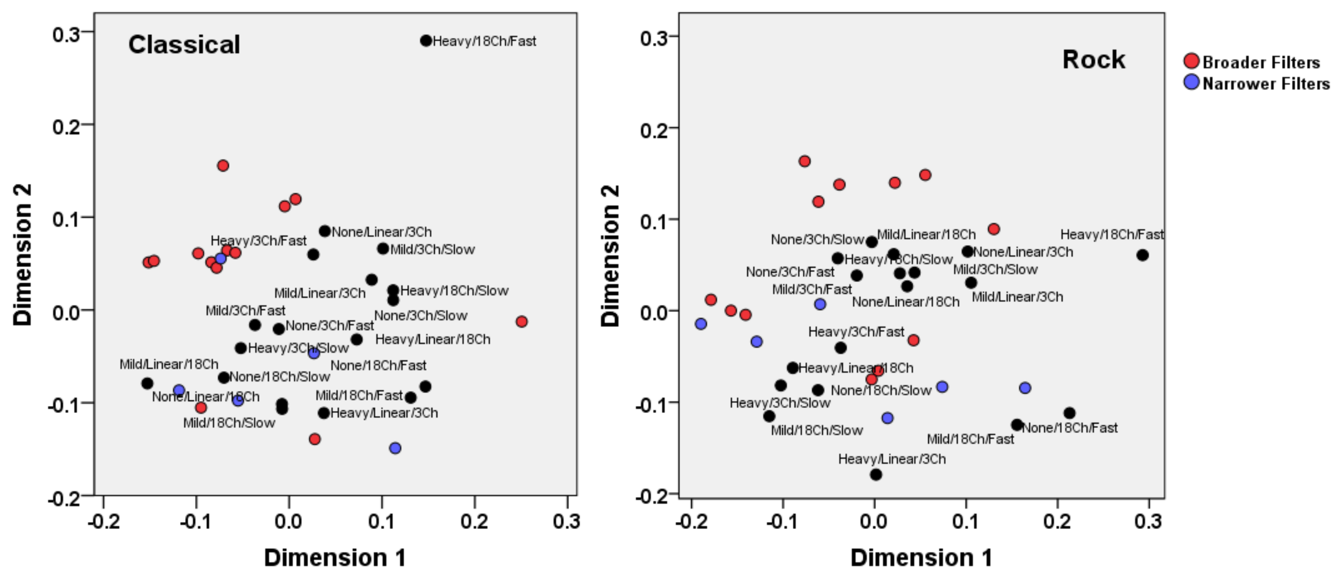

Figure 3.11: Two-dimensional solutions for multidimensional unfolding analysis for classical music (left) and rock music (right). Stimulus locations are represented in black circles and labeled based on the stimulus condition. Listener ideal points are shown in colored circles, with red circles representing listeners with broader tuning curves (Q10 <3.0) for the 2000-Hz measurement for classical and for the 500-Hz measurement for rock. Blue circles represent the ideal points of listeners with narrower tuning curves (Q10 $\geq$3.0).

are shown in the red and blue circles. Listeners with broader tuning curves are shown in red, and listeners with narrower tuning curves are shown in blue. Dimension 2 for classical music correlated more strongly with the 2000-Hz tuning curve measurements (r=-0.522,; ρ=-0.624), and rock music correlated best with the 500-Hz tuning curve measurements (r=-0.562,; ρ=-0.558). Therefore, the listeners are divided based on their 2000-Hz Q10 values for classical music and 500-Hz Q10 values for rock music. The locations of the listener ideal points generally show listeners with broader tuning curves falling toward the upper end of Dimension 2 (less spectral modification), while listeners with narrower tuning curves tend to be on the lower end of Dimension 2 (more spectral modification). This finding may be due to an increased ability of listeners with good frequency selectivity to resolve harmonics in spectrally degraded signals, thus giving higher preference ratings to such stimuli than their counterparts with broader tuning curves.

3.4 Discussion

Music perception with hearing aids is a complex problem involving multiple factors. This study sought to address several of the factors important for music and hearing aids by quantifying the acoustic and perceptual effects of music processed with different levels of CL, WDRC, and CL+WDRC. An acoustic analysis of the stimuli showed similar effects of dynamic-range compression for music as have been shown previously for speech. Both CL and WDRC caused reduced temporal envelope contrasts and a decrease in crest factor, which was more pronounced for WDRC conditions that had 18 channels compared to 3 channels and was more pronounced for a fast release time (50 msec) compared to a slow release time (1000 msec). Additionally, WDRC upset the spectral balance for the music stimuli, such that higher-frequency sounds showed a greater reduction in dynamic range than lower-frequency sounds (for a high-frequency hearing loss), and this effect was greatest for the 18-channel fast WDRC. These findings are consistent with previous work showing temporal envelope distortion and flattened spectral contrast in response to WDRC for speech signals (Bor et al., 2008; Jenstad and Souza, 2005). In the current study, compression also affected the modulation spectra of the music samples – specifically, heavy CL smoothed the peaks of the modulation spectrum such that the rhythmic structure of the music was less distinct, while WDRC showed little effect on the modulation spectrum. In a previous study using speech, Souza and Gallun (2010) showed that CL and WDRC both affected the modulation spectra of consonant sounds. The authors reported that CL caused some phonemes to become more similar to one another and other phonemes to become more dissimilar, and fast-acting multichannel WDRC actually increased heterogeneity among the phonemes (Souza and Gallun, 2010). It should be noted that the type of CL used by Souza and Gallun (2010) was implemented in hearing aids, and it was, therefore, a different type of CL than the music-industry CL used in this study (the hearing-aid CL was used to compress only high-level sounds, and it did not increase the RMS level of the signals).

The perceptual results for music processed by different levels of compression were generally consistent with a preference for the uncompressed dynamic range. For classical music, linear processing and slow WDRC were equally preferred, while fast WDRC gave significantly poorer quality. For rock music, linear processing was preferred to both WDRC conditions, with slow WDRC being the second most preferred and fast WDRC again being the least preferred condition. Follow-up analyses showed that these effects were significantly predicted by the nonlinear term in HASQI, which measures changes in the envelope time-frequency modulation (Kates and Arehart, 2010). For classical music, preferences were significantly correlated with nonlinear HASQI for both parametric and non-parametric analyses ($r=0.789$, $p<0.001$; $\rho=0.787$, $p<0.001$), while for rock music, preferences were only correlated with nonlinear HASQI using non-parametric statistics ($r=0.314$, $p=0.204$; $\rho=0.472$, $p=0.048$). The reduced quality caused by fast-acting WDRC in this study is consistent with previous research showing preferences for ADRO over WDRC for classical, jazz, and rock music (Higgins et al., 2012) and with preferences for slow release times over fast release times for classical, jazz (Moore et al., 2011), and pop music (Hansen, 2002).

The effect of the number of processing channels was slightly more complex. For classical music, the main effect of channels was not significant. However, the interaction between channels and psychophysical tuning curve bandwidth at 500 Hz was significant, which will be discussed in more detail below. For rock music, the 3-channel processing condition was significantly preferred to the 18-channel processing condition. This result is consistent with the idea that maintaining the spectral balance in music is important for perceived quality, which supports previous research showing that spectral changes degrade music quality for listeners with normal hearing and with hearing loss (Arehart et al., 2011; Moore and Tan, 2003; Gabrielsson and Sjogren, 1979). An additional effect seen for the rock music was that the number of channels significantly interacted with the compression release time. The release time variable had a larger effect in the 18-channel condition compared to the 3-channel condition. This finding is consistent with interactions reported by van Buuren et al. (1999), who showed that the more the signal envelope was affected in music stimuli, the worse the quality judgments were.

An additional factor examined here was the use of music-industry CL on music stimuli prior to the input stage of the hearing aid. Due to the common use of CL in the music industry, especially for pop music genres, a hearing-aid user could potentially experience twofold effects of compression: the results of CL prior to hearing-aid processing, and the results of WDRC processing within the hearing aid. The current findings demonstrated that heavy CL was detrimental to quality for classical music, but the main effect of CL was not significant for rock music. The classical results are consistent with ratings performed by normal-hearing listeners, reported in Chapter 2. Namely, a small amount of CL did not affect quality, but heavy CL was significantly less preferred. However, the Chapter 2 results showed similar preferences for rock music, while the current study did not show an overall effect of CL with rock. One explanation for this finding may be that the current listeners with hearing loss were less sensitive to distortion caused by CL than listeners with normal hearing. Another possibility could be that most participants in this study were not regular listeners of rock music. Only 6 of the 18 listeners reported listening to rock music on a regular basis. Despite the lack of a main effect for CL in rock music, the interaction between CL and the number of channels was significant, in that the difference between 3 channels and 18 channels was largest for the heavy CL condition. This result suggests that the listeners with hearing loss included in the current study were sensitive to the total amount of distortion caused by CL+WDRC processing, even in rock music.

We also considered the extent to which threshold and suprathreshold auditory characteristics contribute to individual variability in preference ratings. In addition to audiometric thresholds, several psychoacoustic measurements were taken to assess the auditory processing capabilities of listeners in terms of frequency resolution, temporal resolution, and dynamic range. Frequency resolution was measured with psychophysical tuning curves using the SWPTC test (Sek et al., 2005; Sek and Moore, 2011), which quantified tuning curve bandwidth at 500 Hz and 2000 Hz using a Q10 value. Temporal resolution was measured using amplitude modulation depth discrimination with an adaptive procedure that varied the modulation depth of a 4-Hz modulated carrier noise. Dynamic range was measured using the Contour Test of Loudness Perception (Cox et al., 1997). Of these variables, tuning curve bandwidth and dynamic range proved to be predictive of listeners' preferences. For classical music, listeners with broader filters at 500 Hz preferred 3-channel processing, while listeners with narrower filters preferred 18-channel processing. This finding is consistent with a recent study by Souza et al. (2012), who reported that participants with better frequency resolution were more able to make use of reduced spectral contrast in multi-channel compressed vowels. The authors suggested that listeners with narrower auditory filters may have been able to resolve closely spaced vowel formants, thereby showing better vowel recognition than listeners with broader filters who could not as easily perform this task. Interestingly, a similar explanation may be available for the current results with music. In orchestral pieces such as the classical music included in this study, the instrumentation is dominated by strings. Perceived quality for stringed instruments relies heavily on harmonic structure (Chasin and Russo, 2004). Like the listeners in Souza et al. (2012), the current participants with narrower tuning curves may have had a greater ability to utilize the spectral details present in the multi-channel compressed music stimuli than listeners with broader auditory filters. The group of listeners with better frequency resolution also showed a preference for slow WDRC over linear processing for classical music, whereas the listeners with poorer frequency resolution preferred the linear stimulus. This finding might also be due to a greater resistance to acoustic degradations for the listeners with narrower tuning curves. The dynamic-range measurement at 500 Hz predicted listeners' preferences for the rock music stimulus. Although listeners with small and large dynamic ranges preferred linear processing over slow WDRC and fast WDRC, the degree of preference was greater for listeners who had a wider dynamic range. This result is consistent with Gatehouse et al. (2006), who reported that better performance with linear processing for speech was associated with a wider dynamic range.

Preference ratings were also subjected to a multidimensional unfolding analysis. The unfolding solution revealed that preference ratings varied along two dimensions, one that corresponded to the sensation level of the stimuli, and a second dimension that corresponded to the amount of spectral modification to the stimuli. The sensation level dimension was more strongly correlated with the rock music, and the spectral dimension was more strongly correlated with the classical music. The frequency resolution of the listeners again seemed to play a role in their preference decisions. With a few exceptions, the ideal points of listeners with narrower tuning curves were closer to stimuli that fell lower on Dimension 2, corresponding to greater amounts of spectral modification. These findings were consistent with the ANOVA results showing the importance of frequency resolution and dynamic range in listeners' judgments of music quality.

3.5 Clinical Implications

The results of this study have provided insight into the extent to which specific acoustic modifications contribute to perceived music quality for listeners with hearing loss and have identified compression parameters that are perceptually relevant for two genres of music. The findings can be summarized as follows:

(1) The release time of a WDRC system affects listener preferences for music. The fast release time included in this study was consistently less preferred. Linear processing was preferred overall for rock music, and the strength of this preference was greatest for listeners with a wider dynamic range. On average, the linear condition and slow WDRC were equally preferred for classical music, but further analysis showed that listeners with broader tuning curves generally preferred linear processing, while listeners with narrower curves generally

preferred slow WDRC. The fast compression that is commonly used in digital hearing aids might lead to poor music quality.

(2) The number of WDRC processing channels also plays a role in music preferences. For rock music, the 3-channel conditions were preferred to the 18-channel conditions. For classical music, the number of channels had a significant interaction with the measured tuning curves of the listeners, such that listeners with narrower tuning curves generally preferred 18 channels, and listeners with broader tuning curves generally preferred 3 channels. This finding is consistent with the idea that hearing-aid users with better frequency resolution are more resistant to spectral modifications that affect harmonic structure in classical music. Further research in this area is warranted.

(3) The contribution of music-industry CL to music preferences is secondary to the importance of WDRC processing parameters for listeners with hearing loss. While heavy CL was less preferred for classical music, the main effect of CL was not significant for rock music. The addition of CL to stimuli at the input stage of the hearing aid did not change the overall preferences for specific WDRC parameters. For listeners using real hearing aids with audio streaming devices, the current simulations suggest that hearing-aid settings might be more important than the level of compression applied to music stimuli prior to the input stage.

The outcomes presented here further support the use of psychophysical measures beyond audiometric threshold when making hearing-aid fitting and design decisions. Factors such as frequency resolution and dynamic range are not only important for speech intelligibility, but the current results suggest that suprathreshold processing capabilities might also affect music perception for listeners with hearing loss. Additional research is needed to better understand the relationships between perceptual characteristics and music processing in hearing aids.

Chapter 4

General Discussion

The perception of recorded music with hearing aids is a highly complex problem that can be described in the context of five components: the acoustic features of music vs. speech, the effects of music-industry recording, the effects of hearing-aid processing, the effects of hearing loss, and the individual characteristics of the listener. The goal of this dissertation was to improve our understanding of music perception with hearing aids by analyzing multiple factors within this framework. Two experiments examined several forms of dynamic-range compression, a type of processing that is important for both music-industry recording and hearing-aid design. The two types of compression were music-industry CL and hearing-aid WDRC. The results of the two experiments provide insight into perceptions of dynamic-range compressed music for listeners with normal hearing and with hearing loss.

4.1 Specific Aim 1

Specific Aim 1 was to characterize the acoustic effects of CL on recorded music and to determine how listeners with normal hearing perceive recorded music processed by CL.

Experiment 1 quantified the effects of CL on classical and rock music with normal-hearing listeners. Acoustic measurements revealed a loss of signal fidelity with increasing compression. When compression had the effect of increased loudness, listeners preferred a small amount of CL compared to none, on average. When the role of loudness was minimized through loudness equalization, low levels of compression were neither harmful nor beneficial to music quality, but higher amounts of compression were judged to be less pleasant and less preferred. Listeners were categorized into three patterns of preferences: the most common profile that showed a preference for low levels of compression, a second profile that showed a preference for greater amounts of compression, and a third profile that showed a preference for no compression at all. In addition to the listener profiles, Experiment 1 revealed differences between musician and non-musician listeners. Musicians tended to perceive a smaller range of loudness than non-musicians and were better able to detect differences in dynamic range caused by compression.

The findings of Experiment 1 indicate that conservative use of CL may improve the perceived quality of music for most normal-hearing listeners. Additionally, the results suggest that signals that are highly compressed within the music industry (e.g., commercial pop/rock music) are distorted prior to hearing-aid processing.

4.2 Specific Aim 2

Specific Aim 2 was to quantify the acoustic and perceptual effects of CL, WDRC, and CL+WDRC on recorded music processed through simulated hearing aids.

Experiment 2 evaluated the separate and combined effects of music-industry CL and hearing-aid WDRC for listeners with hearing loss. An acoustic analysis of the stimuli showed that both CL and WDRC caused reduced envelope contrasts and a decrease in crest factor. Additionally, WDRC caused modifications to the acoustic spectrum, and CL smoothed the peaks of the modulation spectrum. The perceptual results from Experiment 2 showed that fast WDRC generally gave poor music quality for both genres of music. However, other conditions showed differences between the two genres. For classical music, linear processing and slow WDRC were equally preferred, and the overall effect of channels was not significant. For rock music, linear processing was preferred over slow WDRC, and the 3-channel condition was preferred to 18 channels.

To observe differences in perceived quality between listeners with normal hearing and with hearing loss, the effects of CL in the two experiments may be compared. For classical music, both experimental groups judged heavy CL to be the least preferred condition. However, for rock music, the normal-hearing listeners in Experiment 1 were more sensitive to changes caused by CL than listeners with hearing loss in Experiment 2. This difference might be due to a reduced sensitivity to distortion caused by CL in listeners with hearing loss, or it might be due to the fact that the hearing-loss group had less experience listening to rock music compared to the normal-hearing listeners.

These results suggest that the fast WDRC that is commonly used in current hearing-aid design might lead to poor music quality, whereas linear processing and slow WDRC are often preferred for classical and rock music. Furthermore, the effect of WDRC is more important than CL processing applied to music stimuli prior to the hearing-aid input stage.

4.3 Specific Aim 3

Specific Aim 3 was to examine the relationship between preferred compression parameters and individual perceptual characteristics of listeners with hearing loss.

In addition to the average preference judgments pertinent to Specific Aim 2, Experiment 2 gathered psychoacoustic measurements to address Specific Aim 3. Several psychoacoustic measurements corresponded to listeners' preference ratings. For classical music, auditory filter bandwidth as estimated from psychophysical tuning curves predicted preferences for WDRC processing. Listeners with broader filters preferred 3-channel processing, while listeners with narrower filters preferred 18-channel processing. A proposed explanation for this finding is that participants with better frequency resolution were better able to make use of spectral detail provided by the 18-channel processing. The group of listeners with narrower tuning curves also showed a preference for slow WDRC over linear processing for classical music, whereas the listeners with broader tuning curves preferred the linear stimulus. For rock music, the dynamic-range measurement predicted listeners' preferences, in that the degree of preference for the linear condition was greater for listeners who had a larger dynamic range.

These results indicate that variability observed in hearing-aid users' perceptions of music quality might depend, in part, on suprathreshold processing abilities, suggesting that clinical measurements of psychoacoustic variables may help guide hearing-aid fittings for music.

4.4 Conclusions

The experiments presented in this dissertation have investigated several components of music perception with hearing aids. Within the context of dynamic-range compression, the perception of music depends on the properties of the input signal, including the genre of music and the pre-processing applied during audio recording. The use of WDRC during hearing-aid signal processing leads to reduced acoustic fidelity, and in some cases, lower music preferences. Variability in responses to hearing-aid processed music can be partially explained by individual perceptual characteristics. The present findings have implications for both music-industry recording and hearing-aid design.

References

Arehart, K. H. (1994). Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners. *Journal of the Acoustical Society of America*, 95(6):3574–3585.

Arehart, K. H., Kates, J. M., and Anderson, M. C. (2011). Effects of noise, nonlinear processing, and linear filtering on perceived music quality. *International Journal of Audiology*, 50(3):177–190.

Arehart, K. H., Kates, J. M., Anderson, M. C., and Harvey, L. O. (2007). Effects of noise and distortion on speech quality judgments in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 122(2):1150–1164.

Bacon, S. P. and Viemeister, N. F. (1985). Temporal-modulation transfer-functions in normal-hearing and hearing-impaired listeners. *Audiology*, 24(2).

Balfour, P. B. and Hawkins, D. B. (1992). A comparison of sound quality judgments for monaural and binaural hearing-aid processed stimuli. *Ear and Hearing*, 13(5):331–339.

Bhatara, A., Tirovolas, A. K., Duan, L. M., Levy, B., and Levitin, D. J. (2011). Perception of emotional expression in musical performance. *Journal of Experimental Psychology-Human Perception and Performance*, 37(3):921–934.

Bor, S., Souza, P., and Wright, R. (2008). Multichannel compression: Effects of reduced spectral contrast on vowel identification. *Journal of Speech Language and Hearing Research*, 51(5):1315–1327.

Borg, I. and Groenen, P. (2005). *Modern multidimensional scaling: Theory and applications.* Springer, New York, 2nd edition.

Bregitzer, L. (2009). *Secrets of recording: Professional tips, tools & techniques.* Elsevier, Amsterdam.

Byrne, D. and Dillon, H. (1986). The national-acoustic-laboratories (NAL) new procedure for selecting the gain and frequency-response of a hearing-aid. *Ear and Hearing*, 7(4):257–265.

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

Carney, A. E. and Nelson, D. A. (1983). An analysis of psychophysical tuning curves in normal and pathological ears. *Journal of the Acoustical Society of America*, 73(1):268–278.

Chasin, M. (2003). Music and hearing aids. *The Hearing Journal*, 56(7):36–41.

Chasin, M. (2006). Hearing aids for musicians. *Hearing Review*, 13(3):11–16.

Chasin, M. (2010). Amplification fit for music lovers. *The Hearing Journal*, 63(9):27–30.

Chasin, M. and Russo, F. (2004). Hearing aids and music. *Trends in Amplification*, 8(2):35–47.

Chisolm, T. H., Johnson, C. E., Danhauer, J. L., Portz, L. J. P., Abrams, H. B., Lesner, S., McCarthy, P. A., and Newman, C. W. (2007). A systematic review of health-related quality of life and hearing aids: Final report of the american academy of audiology task force on the health-related quality of life benefits of amplification in adults. *Journal of the American Academy of Audiology*, 18(2):151–183.

Coombs, C. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, 57:145–158.

Cox, R. M., Alexander, G. C., Taylor, I. M., and Gray, G. A. (1997). The contour test of loudness perception. *Ear and Hearing*, 18(5):388–400.

Croghan, N. B. H., Arehart, K. H., and Kates, J. M. (2012). Quality and loudness judgments for music subjected to compression limiting. *Journal of the Acoustical Society of America*, 132(2):1177–1188.

Dahlquist, M. and Leijon, A. (2012). Paired-comparison rating of sound quality using map parameter estimation for data analysis. In *ISCA Tutorial and Research Workshop on Auditory Quality of Systems*, pages 79–84.

Davies-Venn, E., Souza, P., Brennan, M., and Stecker, G. C. (2009). Effects of audibility and multichannel wide dynamic range compression on consonant recognition for listeners with severe hearing loss. *Ear and Hearing*, 30(5):494–504.

Davies-Venn, E., Souza, P., and Fabry, D. (2007). Speech and music quality ratings for linear and nonlinear hearing aid circuitry. *Journal of the American Academy of Audiology*, 18(8):688–699.

de Laat, J. (1985). The effect of competing melodies on melody recognition by hearing-impaired and normal-hearing listeners. *Journal of the Acoustical Society of America*, 78:1574–1577.

Dillon, H. (1984). A procedure for subjective quality rating of hearing aids. Technical report, NAL.

Emiroglu, S. and Kollmeier, B. (2008). Timbre discrimination in normal-hearing and hearing-impaired listeners under different noise conditions. *Brain Research*, 1220:199–207.

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). Mini-mental state - practical method for grading cognitive state of patients for clinician. *Journal of Psychiatric Research*, 12(3):189–198.

Franklin, C. A., Thelin, J. W., Nabelek, A. K., and Burchfield, S. B. (2006). The effect of speech presentation level on acceptance of background noise in listeners with normal hearing. *Journal of the American Academy of Audiology*, 17(2):141–146.

Freyaldenhoven, M. C., Smiley, D. F., Muenchen, R. A., and Konrad, T. N. (2006). Acceptable noise level: Reliability measures and comparison to preference for background sounds. *Journal of the American Academy of Audiology*, 17(9):640–648.

Gabrielsson, A. and Sjogren, H. (1979). Perceived sound quality of sound-reproducing systems. *Journal of the Acoustical Society of America*, 65(4):1019–1033.

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

Geringer, J. M. (1995). Continuous loudness judgments of dynamics in recorded music excerpts. *Journal of Research in Music Education*, 43(1):22–35.

Glasberg, B., Moore, B., and Bacon, S. (1987). Gap-detection and masking in hearing-impaired and normal-hearing subjects. *International Journal of Audiology*, 45:153–171.

Glasberg, B. R. and Moore, B. C. J. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5):331–342.

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

Harvey, L.O., J. and Gervais, M. (1981). Internal representation of visual texture as the basis for the judgment of similarity. *Journal of Experimental Psychology: Human Perception and Performance*, 7(4):741–753.

Henning, R. L. W. and Bentler, R. A. (2008). The effects of hearing aid compression parameters on the short-term dynamic range of continuous speech. *Journal of Speech Language and Hearing Research*, 51(2):471–484.

Higgins, P., Searchfield, G., and Coad, G. (2012). A comparison between the first-fit settings of two multichannel digital signal-processing strategies: music quality ratings and speech-in-noise scores. *American Journal of Audiology*, 21(June 2012):13–21.

Jenstad, L. M., Pumford, J., Seewald, R. C., and Cornelisse, L. E. (2000). Comparison of linear gain and wide dynamic range compression hearing aid circuits II: Aided loudness measures. *Ear and Hearing*, 21(1):32–44.

Jenstad, L. M. and Souza, P. E. (2005). Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *Journal of Speech Language and Hearing Research*, 48(3):651–667.

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

Kates, J. M. (2010). Understanding compression: Modeling the effects of dynamic-range compression in hearing aids. *International Journal of Audiology*, 49(6):395–409.

Kates, J. M. and Arehart, K. H. (2005). Multichannel dynamic-range compression using digital frequency warping. *Eurasip Journal on Applied Signal Processing*, 2005(18):3003–3014.

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

Katz, B. (2007). *Mastering audio: The art and the science*. Focal Press, Amsterdam.

Kefauver, A. (2001). *The audio recording handbook*. A-R Editions, Inc., Middleton.

Kellaris, J. and Rice, R. (1993). The influence of tempo, loudness, and gender of listener on responses to music. *Psychology & Marketing*, 10:15–29.

King, A. and Martin, M. (1984). Is agc beneficial in hearing aids? *British Journal of Audiology*, 18:31–38.

Kochkin, S. (2012). Marketrak VIII: The key influencing factors in hearing aid purchase intent. *Hearing Review*, March 2012:12–20.

Koelsch, S., Schroger, E., and Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, 10(6):1309–1313.

Leek, M. R., Molis, M. R., Kubli, L. R., and Tufts, J. B. (2008). Enjoyment of music by elderly hearing-impaired listeners. *Journal of the American Academy of Audiology*, 19(6):519–526.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2):467–477.

Maempel, H.-J. and Gawlik, F. (2009). The influence of sound processing on listeners program choice in radio broadcasting. In *Audio Engineering Society 126th Convention*, volume Paper No. 7785, pages 1–10.

Moore, B. (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. (2012). *An Introduction to the Psychology of Hearing.* Emerald Group Publishing Limited, United Kingdom.

Moore, B. C. J., Fullgrabe, C., and Stone, M. A. (2011). Determination of preferred parameters for multichannel compression using individually fitted simulated hearing aids and paired comparisons. *Ear and Hearing*, 32(5):556–568.

Moore, B. C. J. and Glasberg, B. R. (1997). A model of loudness perception applied to cochlear hearing loss. *Auditory Neuroscience*, 3(3):289–311.

Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (2003). Why are commercials so loud? - perception and modeling of the loudness of amplitude-compressed speech. *Journal of the Audio Engineering Society*, 51(12):1123–1132.

Moore, B. C. J. and Tan, C. T. (2003). Perceived naturalness of spectrally distorted speech and music. *Journal of the Acoustical Society of America*, 114(1):408–419.

Muchnik, C., Amir, N., Shabtai, E., and Kaplan-Neeman, R. (2012). Preferred listening levels of personal listening devices in young teen- agers: Self reports and physical measurements. *International Journal of Audiology*, 51:7.

Munro, K. and Lutman, M. (2005). Sound quality judgements of new hearing instrument users over a 24-week post-fitting period. *International Journal of Audiology*, 44:92–101.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, 13(2):87–110.

Neuman, A. C., Bakke, M. H., Hellman, S., and Levitt, H. (1995a). Preferred listening levels for linear and slow-acting compression hearing-aids. *Ear and Hearing*, 16(4):407–416.

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., and Levitt, H. (1995b). Effect of release time in compression hearing aids: Paired-comparison judgments of quality. *Journal of the Acoustical Society of America*, 98(6):3182–3187.

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., and Levitt, H. (1998). The effect of compression ratio and release time on the categorical rating of sound quality. *Journal of the Acoustical Society of America*, 103(5):2273–2281.

NIDCD (2010). Quick statistics. Downloaded September 22, 2012 from http://www.nidcd.nih.gov/health/statistics/Pages/quick.aspx.

Olive, S. E. (2003). Differences in performance and preference of trained versus untrained listeners in loudspeaker tests: A case study. *Journal of the Audio Engineering Society*, 51(9):806–825.

Olsen, K. N., Stevens, C. J., and Tardieu, J. (2010). Loudness change in response to dynamic acoustic intensity. *Journal of Experimental Psychology-Human Perception and Performance*, 36(6):1631–1644.

Olsen, W. (1998). Average speech levels and spectra in various speaking/listening conditions: A summary of the pearson, bennett, & fidell (1977) report. *American Journal of Audiology*, 7:21–25.

Oxenham, A. (2008). Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants. *Trends in Amplification*, 12:316–331.

Oxenham, A. J., Fligor, B. J., Mason, C. R., and Kidd, G. (2003). Informational masking and musical training. *Journal of the Acoustical Society of America*, 114(3):1543–1549.

Parbery-Clark, A., Skoe, E., Lam, C., and Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, 30(6):653–661.

Parbery-Clark, A., Strait, D. L., and Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, 49(12):3338–3345.

Parbery-Clark, A., Tierney, A., Strait, D. L., and Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience*, 219:111–119.

Patterson, R. D., Allerhand, M. H., and Giguere, C. (1995). Time-domain modeling of peripheral auditory processing - a modular architecture and a software platform. *Journal of the Acoustical Society of America*, 98(4):1890–1894.

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

Portnuff, C. D. F., Fligor, B. J., and Arehart, K. H. (2011). Teenage use of portable listening devices: A hazard to hearing? *Journal of the American Academy of Audiology*, 22(10):663–677.

Pras, A., Zimmerman, R., Levitin, D., and Guastavino, C. (2009). Subjective evaluation of mp3 compression for different musical genres. In *Audio Engineering Society 127th Convention*, pages 1–7.

Rabiner, L. R., Levitt, H., and Rosenberg, A. (1969). Investigation of stress patterns for speech synthesis by rule. *Journal of the Acoustical Society of America*, 45(1):92–101.

Rammsayer, T. and Altenmuller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception*, 24(1):37–47.

Sabin, A., Gallun, F., and Souza, P. (2012). Sensitivity to dynamic range compression is predictable by distances between modulation spectra. Poster presented at the International Hearing Aid Research Conference, Lake Tahoe, CA, August 8-12.

Scheirer, E. D. (1998). Tempo and beat analysis of acoustic musical signals. *Journal of the Acoustical Society of America*, 103(1):588–601.

Schmidt, J. and Rutledge, J. (1996). Multichannel dynamic range compression for music signals. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 1013–1016.

Sek, A., Alcantara, J., Moore, B. C. J., Kluk, K., and Wicher, A. (2005). Development of a fast method for determining psychophysical tuning curves. *International Journal of Audiology*, 44(7):408–420.

Sek, A. and Moore, B. C. J. (2011). Implementation of a fast method for measuring psychophysical tuning curves. *International Journal of Audiology*, 50(4):237–242.

Skoe, E. and Kraus, N. (2012). A little goes a long way: How the adult brain is shaped by musical training in childhood. *Journal of Neuroscience*, 32(34):11507–11510.

Smeds, K. (2004). Is normal or less than normal overall loudness preferred by first-time hearing aid users? *Ear and Hearing*, 25(2):159–172.

Souza, P. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends in Amplification*, 6(4):131–165.

Souza, P. and Gallun, F. (2010). Amplification and consonant modulation spectra. *Ear and Hearing*, 31(2):268–276.

Souza, P., Wright, R., and Bor, S. (2012). Consequences of broad auditory filters for identification of multichannel-compressed vowels. *Journal of Speech Language and Hearing Research*, 55(2):474–486.

Souza, P. E. and Bishop, R. D. (1999). Improving speech audibility with wide dynamic range compression in listeners with severe sensorineural loss. *Ear and Hearing*, 20(6):461–470.

Souza, P. E. and Bishop, R. D. (2000). Improving audibility with nonlinear amplification for listeners with high-frequency loss. *Journal of the American Academy of Audiology*, 11(April 2000):214–223.

Souza, P. E. and Turner, C. W. (1998). Multichannel compression, temporal cues, and audibility. *Journal of Speech Language and Hearing Research*, 41(2):315–326.

Stender, T. (2012). Setting the score for music and hearing aids. *The Hearing Review*, July 2012:12–18.

Stone, M. A., Moore, B. C. J., Fullgrabe, C., and Hinton, A. C. (2009). Multichannel fast-acting dynamic range compression hinders perform- ance by young normal-hearing listeners in a two-talker separation task. *Journal of the Audio Engineering Society*, 57:532–546.

Strait, D. and Kraus, N. (2011). Can you hear me now? musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, 2:1–10.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., and Schroger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental Brain Research*, 161(1):1–10.

Todd, N. P. M., O'Boyle, D. J., and Lee, C. S. (1999). A sensory-motor theory of rhythm, time perception and beat induction. *Journal of New Music Research*, 28(1):5–28.

Tufts, J. B., Molis, M. R., and Leek, M. R. (2005). Perception of dissonance by people with normal hearing and sensorineural hearing loss. *Journal of the Acoustical Society of America*, 118(2):955–967.

van Buuren, R. A., Festen, J. M., and Houtgast, T. (1999). Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality. *Journal of the Acoustical Society of America*, 105(5):2903–2913.

Vickers, E. (2011). The loudness war: Do louder, hypercompressed recordings sell better? *Journal of the Audio Engineering Society*, 59(5):346–351.

Appendix A

Preliminary Study for Experiment 2: Chosen Listening Levels

A.1 Introduction

The effect of WDRC in hearing aids depends on the input level of a given stimulus. Sounds that are higher in average level will have a greater proportion of segments over time that fall above the compression kneepoint, resulting in a more heavily compressed signal overall, compared to a lower-level sound. Previous evidence shows that the perceptual effects of WDRC are greater for higher-level inputs. For example, the effect of compression speed on sound quality for both speech and music is greater for an 80-dB SPL input than for sounds that are 65 dB SPL and 50 dB SPL (Moore et al., 2011). Therefore, the selection of an appropriate input level is an important decision when conducting research related to WDRC and sound quality.

Typical levels for speech range from approximately 55 to 75 dBA, depending on the level of background noise (Olsen, 1998). An average level of about 65 dB SPL for speech often serves as an appropriate input level for hearing-aid research due to its real-world relevance. However, an average level for music is much more difficult to define. The level at which music is played depends on several factors, such as the type of music being played, the instrumentation, whether the music is live, acoustic, amplified, or recorded, and the personal preferences of the listener. Chosen listening levels for music have been described for listeners with normal hearing. For example, Portnuff et al. (2011) measured average chosen listening levels of 68.3 dBA in quiet for normal-hearing adolescents, although large individual variability was reported. Based on informal listening tests conducted with normal-hearing listeners, Smeds (2004) measured chosen levels ranging from 67 dB SPL to 82 dB

SPL for several types of music, and the authors then used these levels to measure preference ratings by listeners with hearing loss. Despite such reports of chosen listening levels for normal-hearing listeners, little research is available to document the levels of music chosen by listeners with hearing loss. The lack of evidence in this area poses a difficulty when selecting an input level for music and hearing-aid research.

The purpose of this study was to determine the levels of music chosen by listeners with hearing loss when they are wearing their personal hearing aids, with the goal of identifying what level of music is appropriate to use when conducting research related to music and hearing aids.

A.2 Method

A.2.1 Subjects

Eight adult hearing-aid users participated in the study (5 males and 3 females), ranging in age from 21 to 84 years (mean: 60.4 years). All listeners had bilateral sensorineural hearing loss consistent with a cochlear site of lesion, ranging from mild to severe hearing loss. Three had profound hearing loss above 3000 Hz in at least one ear. Seven of the eight participants used binaural amplification, and one listener wore a hearing aid in the right ear only.

A.2.2 Stimuli and instrumentation

The music samples included the same six stimuli used as input to the simulated hearing aid for Experiment 2: Classical Unprocessed (with no audio compression), Classical Mild (with mild audio compression), Classical Heavy (with heavy audio compression), Rock Unprocessed, Rock Mild, and Rock Heavy.

Stimuli were loaded onto a Dell Vostro 1500 laptop. The music samples were played through Adobe Audition 2.0. and routed via an E-MU 0404 USB 2.0 Audio/MIDI Interface to a GSI-61 audiometer. Sounds were calibrated using a 1-kHz tone RMS-matched to the music samples, which was adjusted using the audiometer VU meter. Stimuli were played in the sound field through a

GSI loudspeaker positioned at 0° azimuth, 1 meter in front of the listener. All measurements were made with the listener seated in a double-walled sound-treated booth.

A.2.3 Procedure

During the listening procedure, the listener was seated in the booth while a tester was seated at the audiometer. The tester gave instructions to the listener and adjusted the level of the stimulus using the audiometer controls. The chosen listening level procedure was based on the instructions provided by Freyaldenhoven et al. (2006) and Franklin et al. (2006) for the Acceptable Noise Level (ANL) Test. Listeners were asked to use thumbs-up and thumbs-down signs to adjust the level of the music. The initial presentation level was approximately 10 dB above the listener's right-ear pure-tone average. Listeners were first asked to turn the music up to a level that was too loud, then down to a level that was too soft, and then back up to their chosen listening level. The tester used 5-dB steps for the first two adjustments and 2-dB steps for the final adjustment. Once the listener had selected a chosen listening level, the tester recorded the audiometer level in dB HL. The instructions given to listeners are written below:

> You will listen to music through a loudspeaker. After a few moments, select the loudness of the music that you like, as if listening to a radio. You will make adjustments by showing me a thumbs-up to turn the volume up and showing me a thumbs-down to turn the volume down. First, turn the loudness up until it is too loud and then down until it is too soft. Finally, select the loudness level that you like.
>
> Now, I will turn the music on. Please show me a thumbs-up every time you want me to turn the music up. Say "okay" when you have reached a level that is too loud.
>
> (Tester uses 5-dB steps)
>
> Now, please show me a thumbs-down every time you want me to turn the music down. Say "okay" when you have reached a level that is too soft.
>
> (Tester uses 5-dB steps)
>
> Now, please select the loudness level that you like. Show me a thumbs-up every time you want me to turn the music up, and show me a thumbs-down every time you want me to turn the music down. Say "okay" when you have reached the level that you like.
>
> (Tester uses 2-dB steps)

A.3 Results

The audiometer output level was converted from dB HL to dB SPL using a correction factor of 10.5 dB. The 10.5-dB correction factor was established by recording the music samples in sound field using a Verifit real-ear recording system. The probe microphone was placed 1 meter in front of the GSI loudspeaker at 0° azimuth, and a 13-second recording was made for each unprocessed stimulus using the "live speech" mode of the Verifit. The real-ear recording was stored to file. The 1/3-octave dB HL and dB SPL values were extracted from the Verifit data file and were used to calculate the A-weighted RMS level for each type of music. For both music stimuli, the difference between the dB HL RMS level and dB SPL RMS level was 10.5 dB. Therefore, a correction factor of 10.5 was used for further analysis.

The chosen listening levels of the eight subjects were averaged for each of the six music samples. The average chosen listening levels ranged from 63.25 to 66.75 dB SPL for aided sound field listening, as shown in Figure A.1 . The average chosen listening level across all listeners and all six music samples was 64.9 dB SPL. Therefore, 65 dB SPL was selected as the input level to the simulated hearing aid in Experiment 2. Although chosen listening levels varied across subjects, with a standard deviation of 8.9 dB across stimuli, the input level of 65 dB SPL was selected to represent an average chosen listening level that was likely to be within the range of comfortable listening for most subjects.

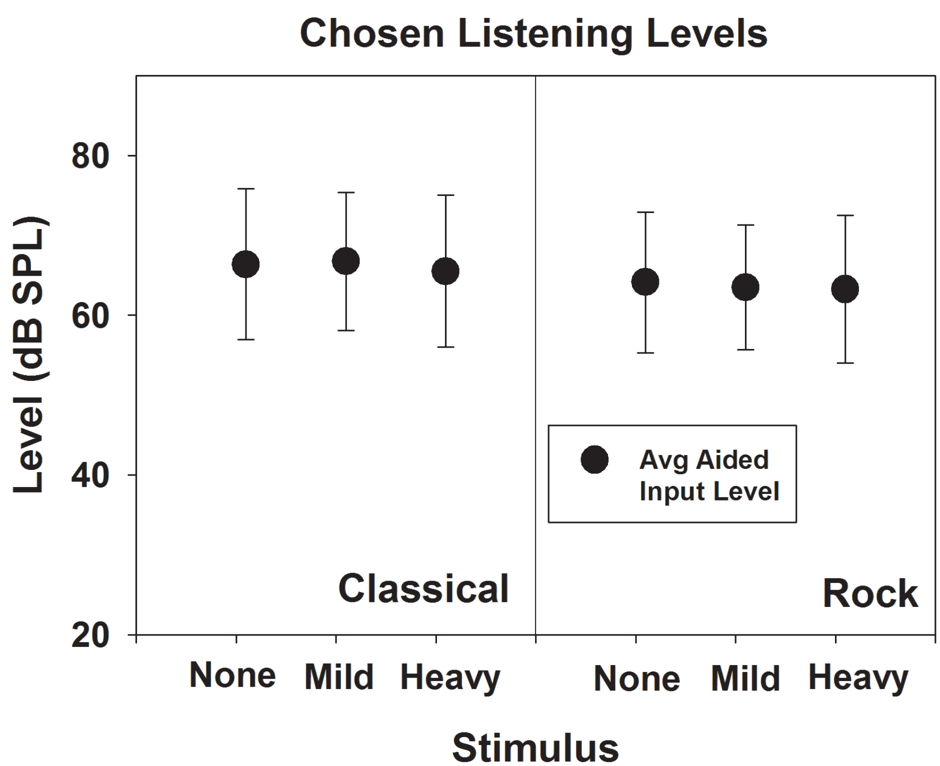

Figure A.1: Average listening levels chosen by hearing-aid users listening to the Experiment 2 music stimuli in the sound field. Error bars represent ±1 standard deviation.

Appendix B

Acoustic Measurements for All Conditions in Experiment 2

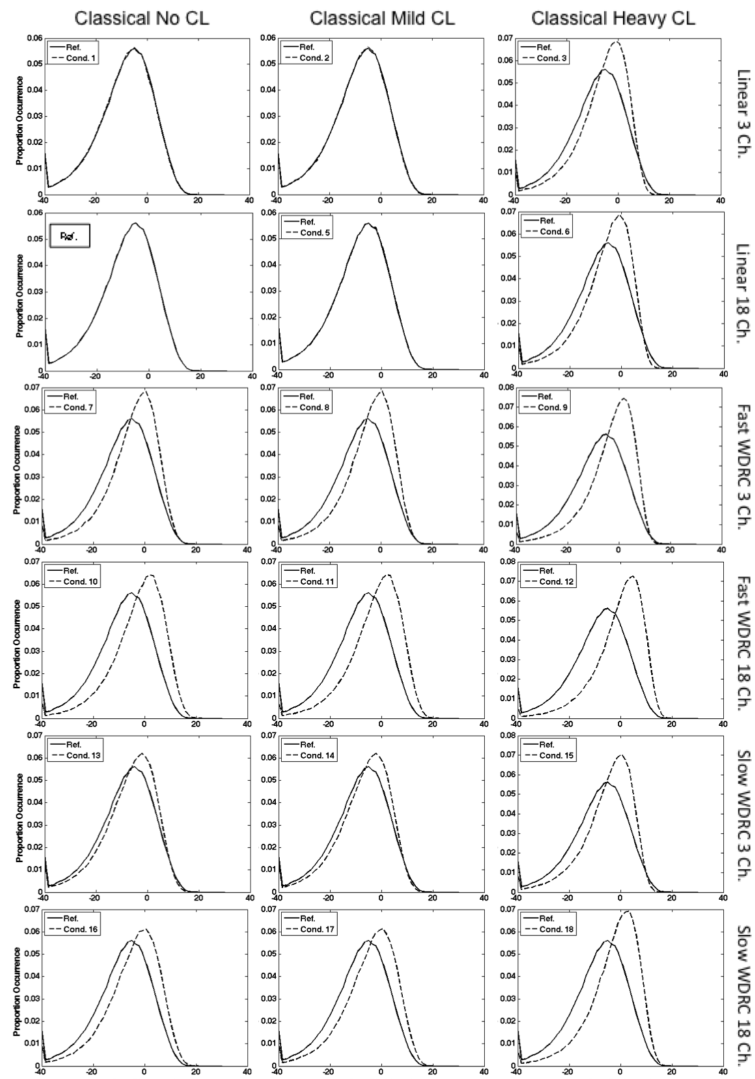

Figure B.1: Amplitude histograms comparing the reference (condition 4: linear, 18-channel processing with no CL) to the other 17 conditions with the classical stimulus. In each panel, the reference stimulus is the solid line and the comparison stimulus is the dashed line.

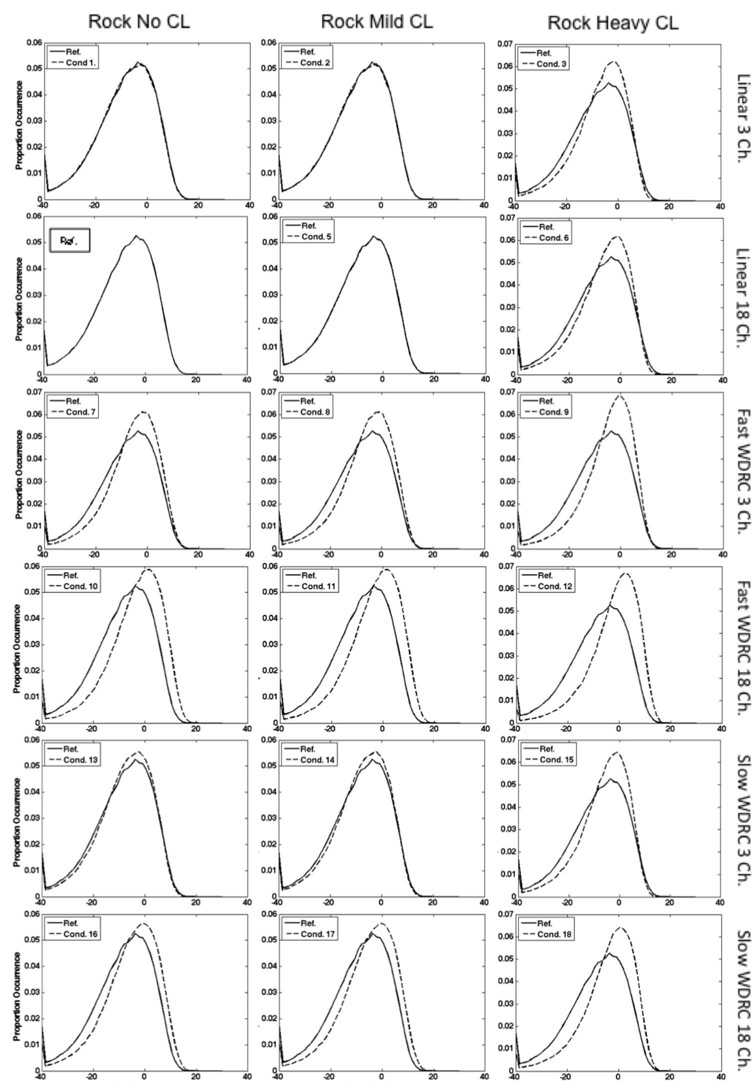

Figure B.2: Amplitude histograms comparing the reference (condition 4: linear, 18-channel processing with no CL) to the other 17 conditions with the rock stimulus. In each panel, the reference stimulus is the solid line and the comparison stimulus is the dashed line.

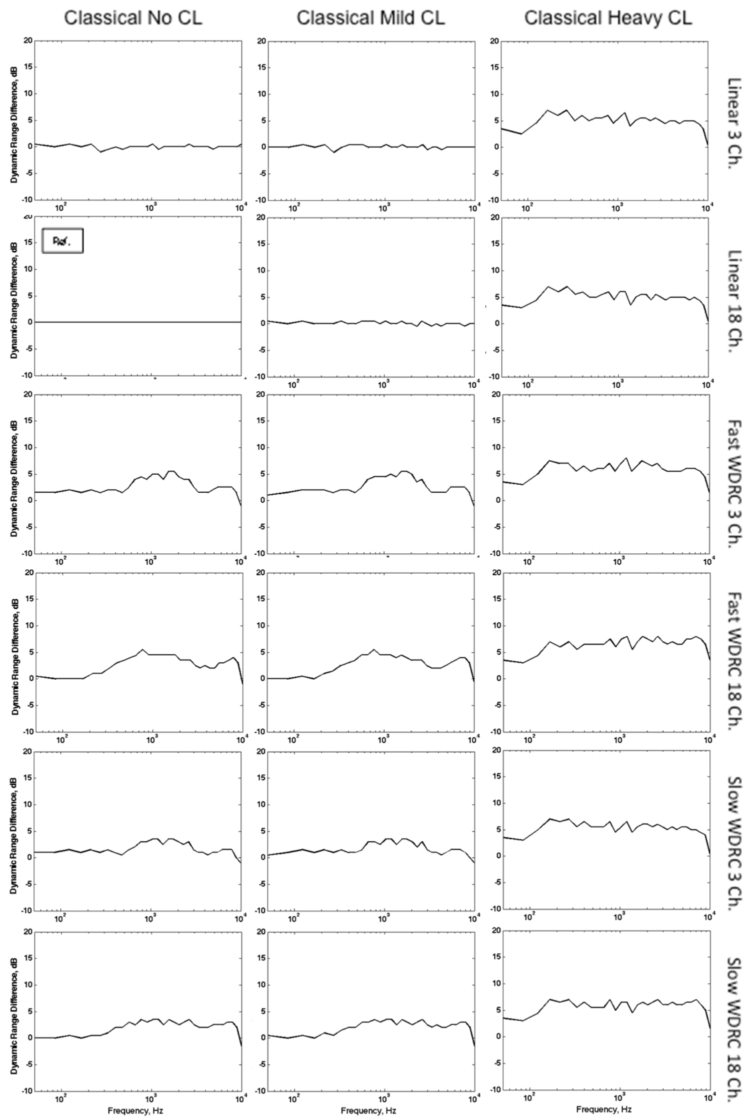

Figure B.3: Dynamic-range difference plots comparing the reference (condition 4: linear, 18-channel processing with no CL) to the other 17 conditions with the classical stimulus. Envelope distributions are analyzed within auditory filter bandwidths and represented across frequencies from 50 to 10,000 Hz. Dynamic range is defined as the difference between the 30% distribution and the 98% distribution for each stimulus. Curves are derived by subtracting the dynamic range of the comparison stimulus from the dynamic range of the condition 4 stimulus.

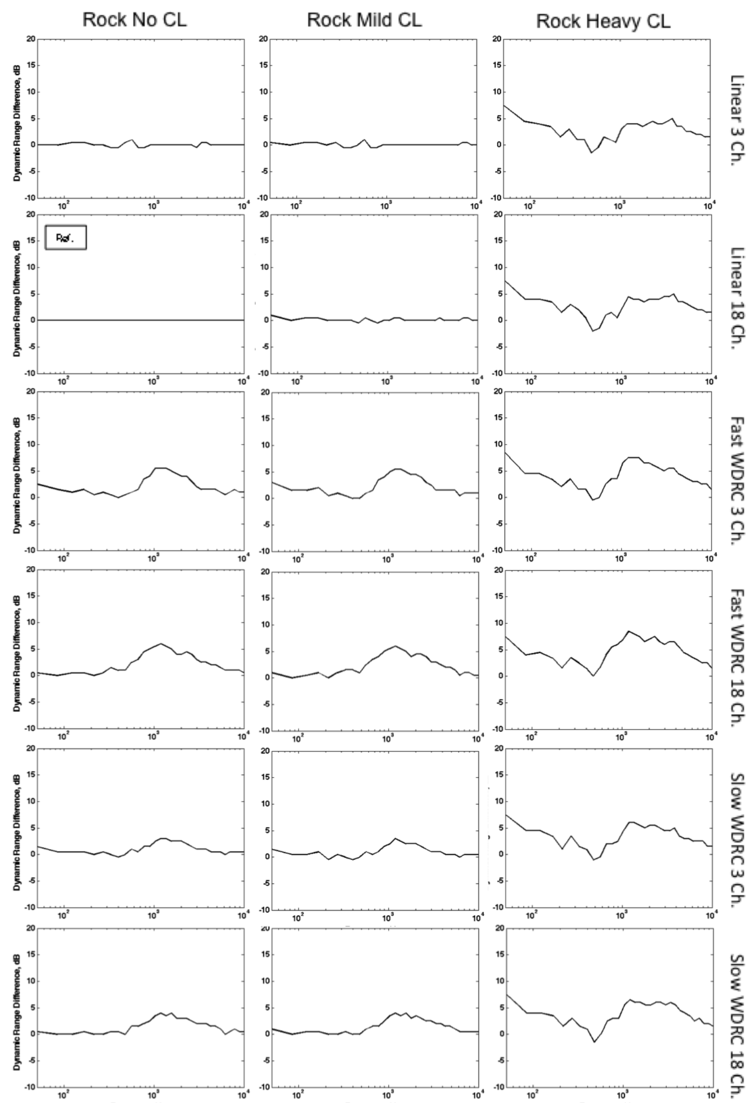

Figure B.4: Dynamic-range difference plots comparing the reference (condition 4: linear, 18-channel processing with no CL) to the other 17 conditions with the rock stimulus. Envelope distributions are analyzed within auditory filter bandwidths and represented across frequencies from 50 to 10,000 Hz. Dynamic range is defined as the difference between the 30% distribution and the 98% distribution for each stimulus. Curves are derived by subtracting the dynamic range of the comparison stimulus from the dynamic range of the condition 4 stimulus.

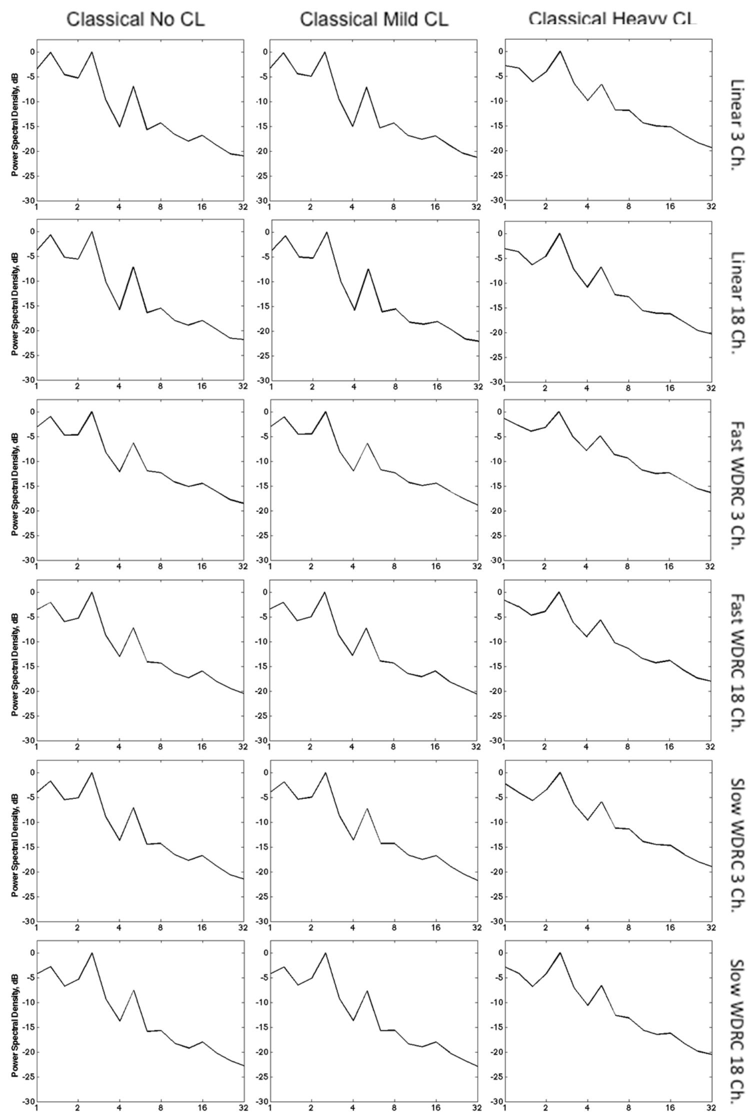

Figure B.5: Modulation spectrum analysis for the classical sample with all 18 conditions. Peaks of the modulation spectrum represent the rhythmic structure of the music.

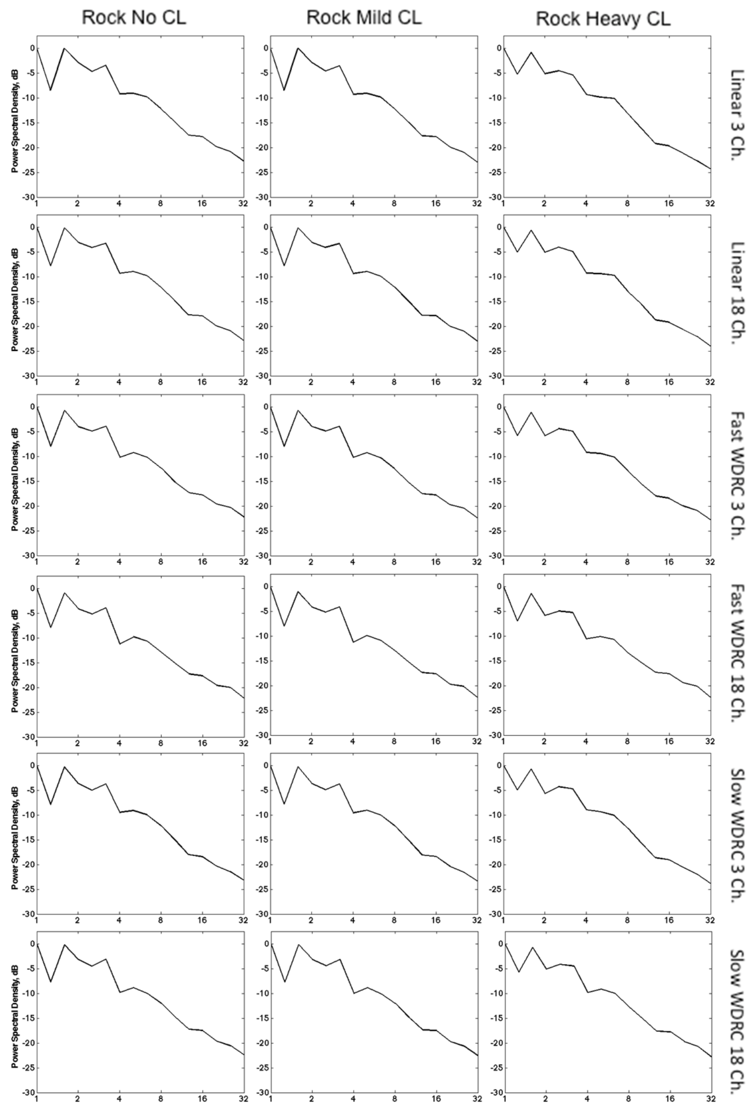

Figure B.6: Modulation spectrum analysis for the rock sample with all 18 conditions. Peaks of the modulation spectrum represent the rhythmic structure of the music.

Appendix C

Individual Subject Data for Experiment 2

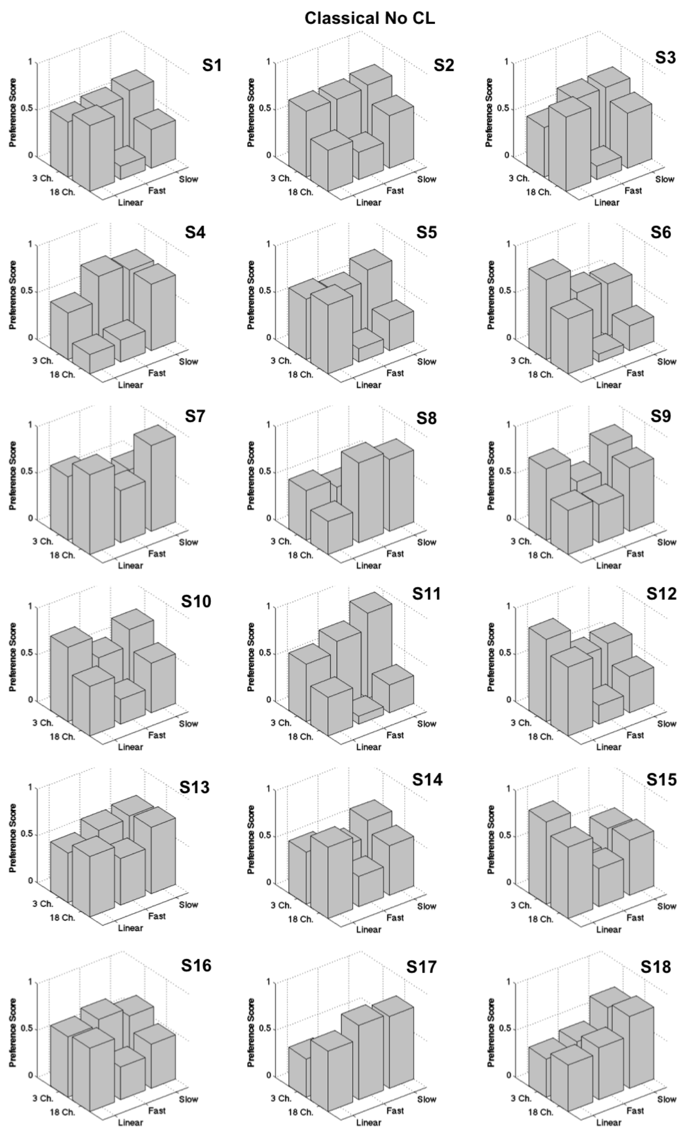

Figure C.1: Individual preference ratings for classical music with no CL. Each panel shows an individual listener's preference score between 0 and 1 plotted as a function of the number of channels (x-axis) and the WDRC processing condition (z-axis).

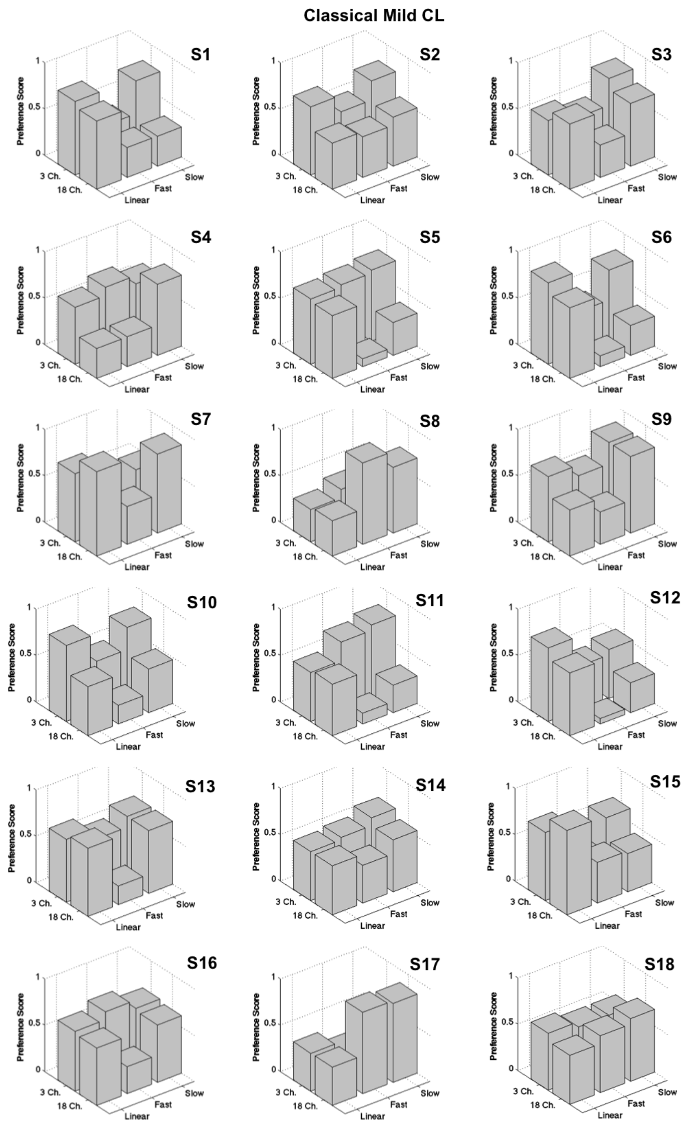

Figure C.2: Individual preference ratings for classical music with mild CL. Each panel shows an individual listener's preference score between 0 and 1 plotted as a function of the number of channels (x-axis) and the WDRC processing condition (z-axis).

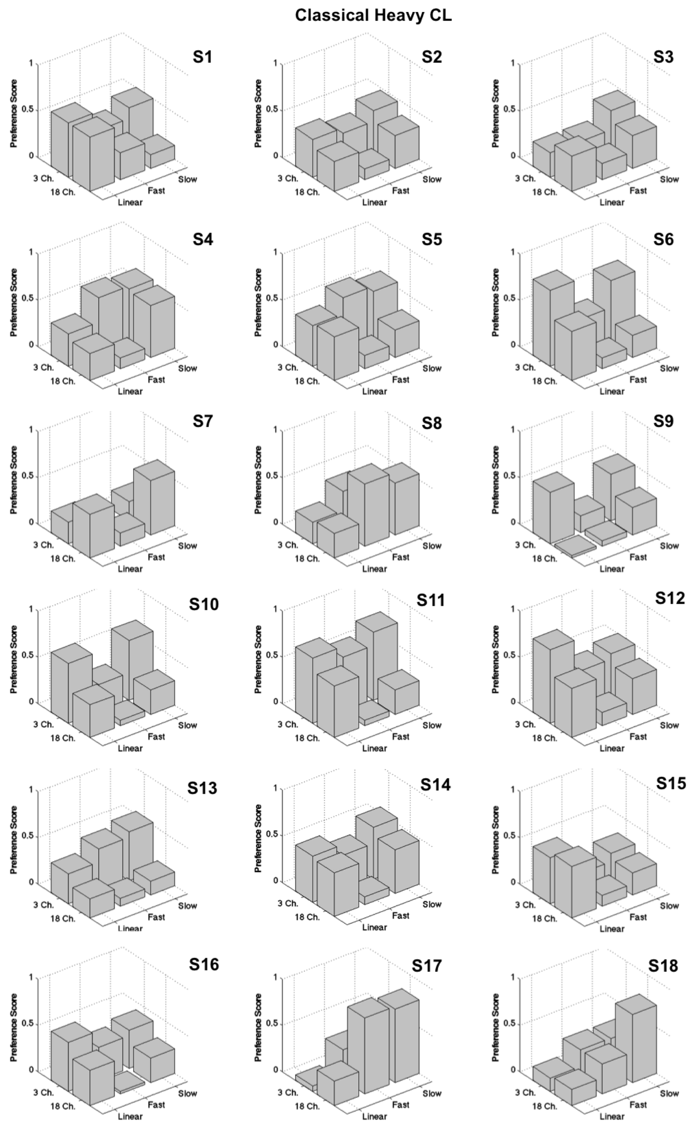

Figure C.3: Individual preference ratings for classical music with heavy CL. Each panel shows an individual listener's preference score between 0 and 1 plotted as a function of the number of channels (x-axis) and the WDRC processing condition (z-axis).

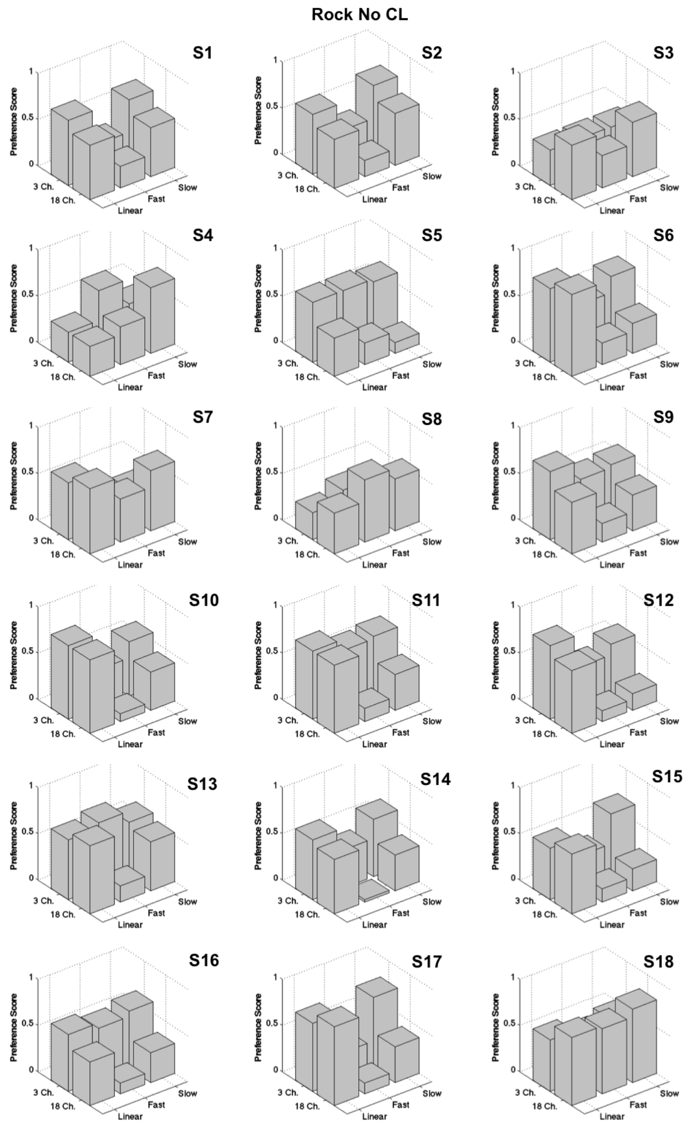

Figure C.4: Individual preference ratings for rock music with no CL. Each panel shows an individual listener's preference score between 0 and 1 plotted as a function of the number of channels (x-axis) and the WDRC processing condition (z-axis).

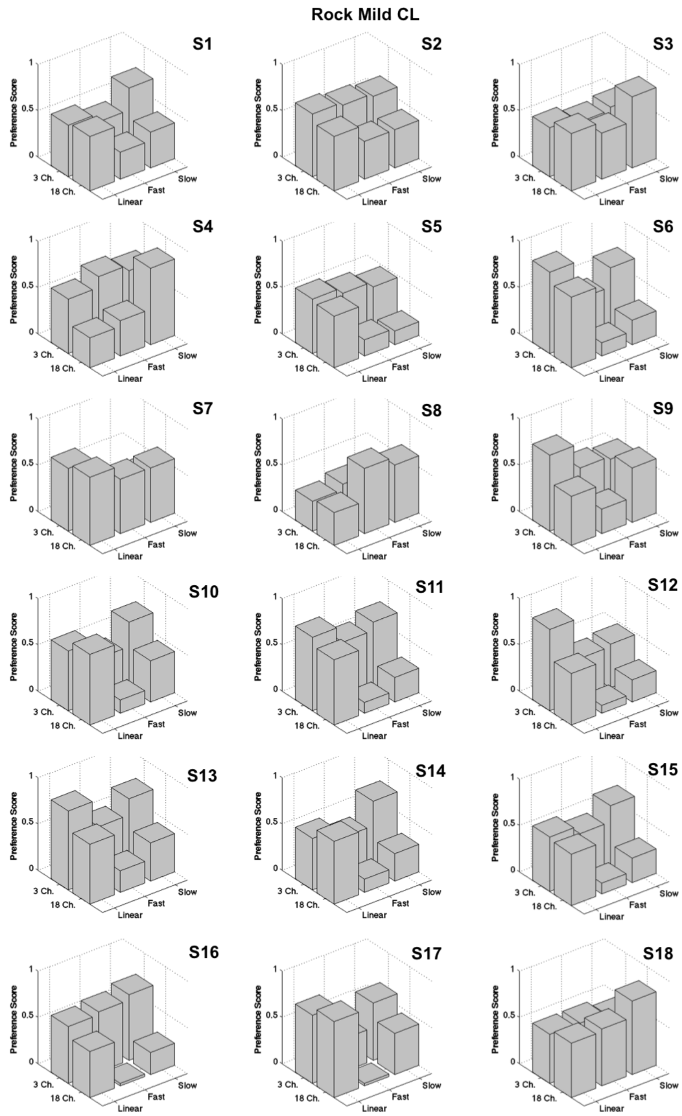

Figure C.5: Individual preference ratings for rock music with mild CL. Each panel shows an individual listener's preference score between 0 and 1 plotted as a function of the number of channels (x-axis) and the WDRC processing condition (z-axis).

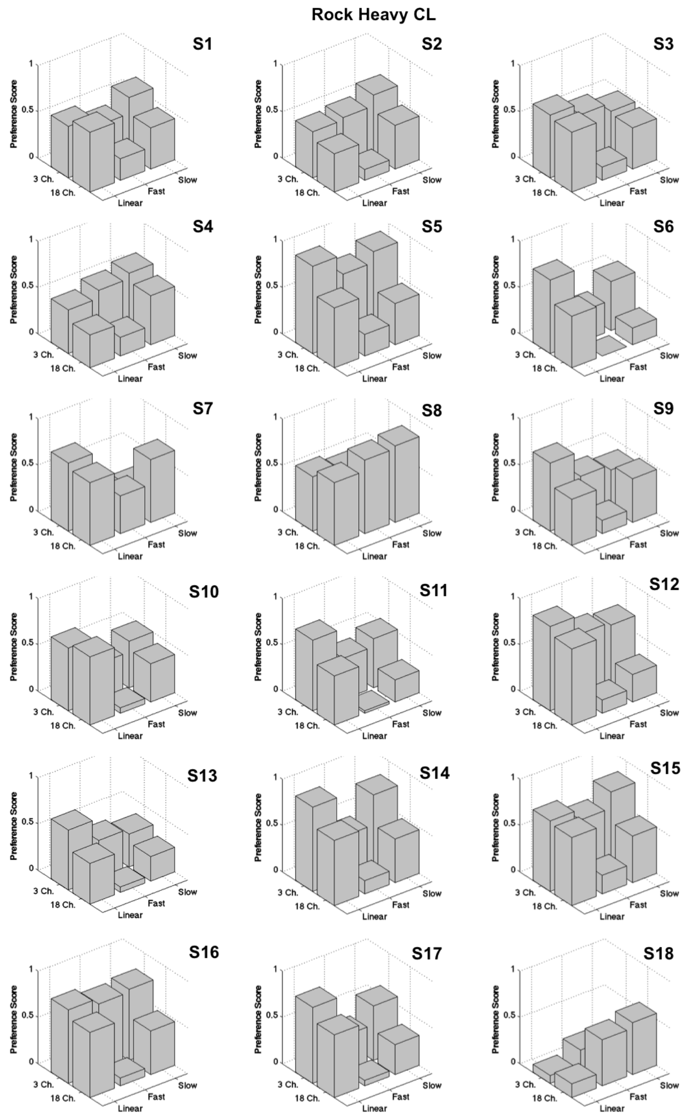

Figure C.6: Individual preference ratings for rock music with heavy CL. Each panel shows an individual listener's preference score between 0 and 1 plotted as a function of the number of channels (x-axis) and the WDRC processing condition (z-axis).

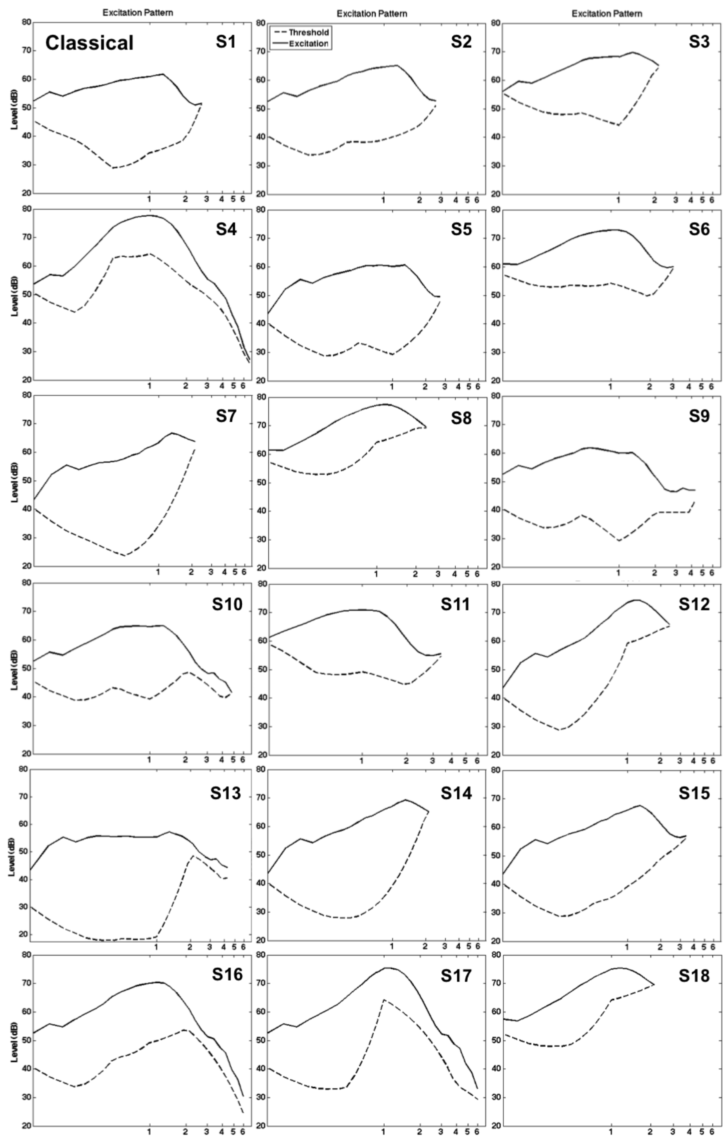

Figure C.7: Individual excitation patterns, representing the distribution of excitation along the basilar membrane, as estimated using the (Moore and Glasberg, 1997) model. Each panel demonstrates a listener's excitation threshold (dotted line) and level of excitation (solid line) across frequencies for Condition 4 (no CL, linear, 18 channels).

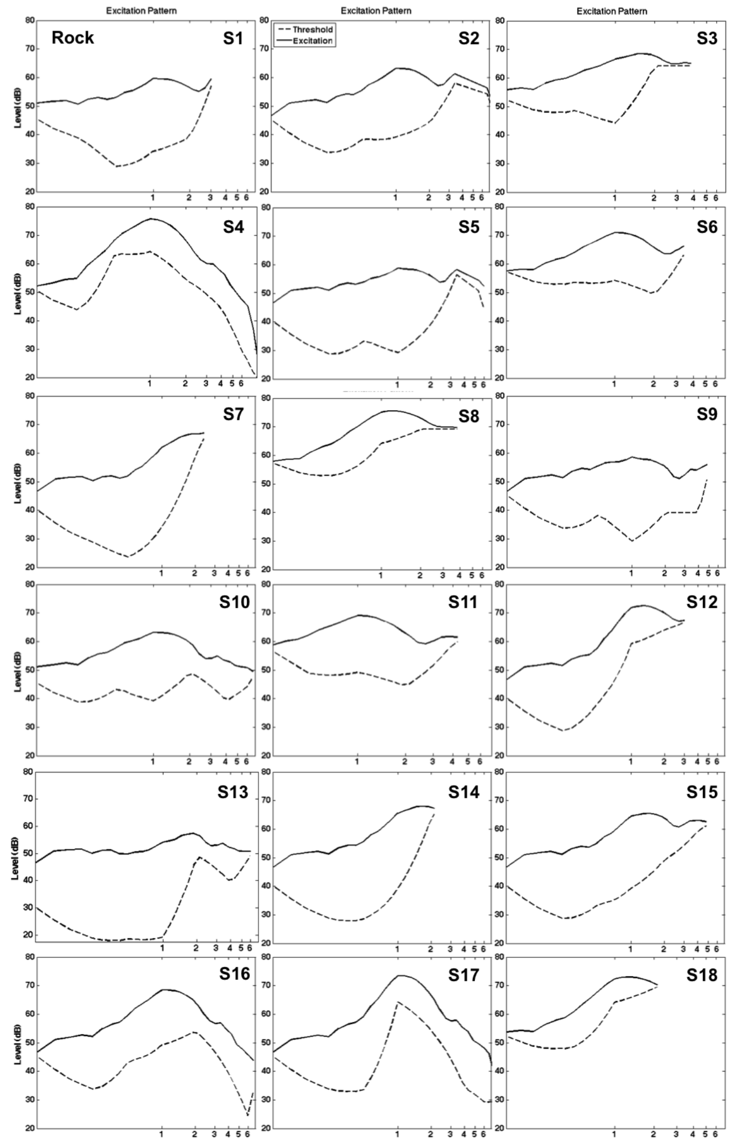

Figure C.8: Individual excitation patterns, representing the distribution of excitation along the basilar membrane, as estimated using the (Moore and Glasberg, 1997) model. Each panel demonstrates a listener's excitation threshold (dotted line) and level of excitation (solid line) across frequencies for Condition 4 (no CL, linear, 18 channels).