Hearing aid amplification at soft input levels

PhD thesis by

Helen Connor Sørensen

Technical University of Denmark

2010

Printed in Denmark.
The defense was held on the 18th of January 2010.

Abstract

Hearing loss is associated with decreased audibility for soft sounds, and impaired loudness perception. Hearing aids are used to help improve these deficits. The general aim of hearing aid fitting is to fit the dynamic range of environmental sounds into the reduced dynamic range of hearing. There is however little consensus among hearing aid researchers regarding how much gain is appropriate, in particular for sounds at low input levels.

The overall aim in this project is to determine the degree to which hearing aids should amplify soft sounds to audibility without compromising listening comfort. An important hearing aid parameter for determining gain for soft sounds is the compression threshold (CT). Lowering the CT increases the gain at low input levels. In this project, the influence of different factors on the preferred CT were investigated in a series of laboratory listening experiments and a field trial.

The influence of compression release time on the preferred CT was investigated in a laboratory listening test with 12 hearing-impaired participants. When a short release time was used, the participants predominately preferred a moderate CT, but when a long release time was used, there was equal preference for moderate and low CTs. The implication is that the release time influenced the preferred gain for soft input levels.

This finding was followed up in a field trial. Twenty hearing aid users (10 new and 10 experienced) compared two hearing aid programs (low and moderate CT) in their daily lives in two trial periods. The two CT programs were combined with either fast-acting or slow-acting compression in each trial period. Overall, the participants most often preferred the moderate CT, except in situations with quiet or distant speech when combined with slow-acting compression. In listening situations that the participants themselves nominated as important, the majority did not report a preference. Compared to the new hearing aid users, the experienced users more often preferred the low CT and thereby more gain at low input levels.

Overall, the results were not strongly in favour of the use of either a low or moderate CT. The findings provide evidence that experienced hearing aid users prefer more gain at low input levels than new hearing aid users. The compression speed also influences the preferred CT. These findings have implications for how hearing aids should be fit to new and experienced users.

Resumé

Høretab indbærer reduceret hørbarhed for svage lyd, og forandret lydstyrkeopfattelse. Høreapparater bliver anvendt til at afhjælpe disse problemer. Det generelle mål ved høreapparattilpasning er at placere alle lyde inden for høreapparat-brugerens hørbare område, sådan at taleforståeligheden og lyttekomforten bliver så god som mulig. Der er dog manglende konsensus blandt høreapparatforskere, om hvor meget forstærkning der er gavnligt ved svage lydniveauer.

Det overordnede formål i dette projekt er at bestemme, i hvor høj grad høreapparater skal forstærke svage lyde op til hørbarhed uden at gå ud over lyttekomfort. Kompressionstærsklen (KT) er en vigtig høreapparat parameter for forstærkning af svage lydniveauer. Jo lavere KT desto mere forstærkning er der ved lave lydniveauer. I dette projekt blev den foretrukne KT bestemt under forskellige omstændigheder i laboratorie lytteforsøg samt et feltforsøg.

Der blev undersøgt påvirking af kompressions-udsvingsningstid på den foretrukne KT i et laboratorie lytteforsøg med 12 testpersoner med høretab. Ved en kort udsvingsningstid, foretrak testpersonerne hovedsagligt en moderat KT, mens der ved en lang udsvingsningstid var samme præference for moderat og lav KT. Dette viser at udsvingsningstiden påvirker den foretrukne forstærkning ved svage lydniveauer.

Den foretrukne KT blev efterfølgende undersøgt i et feltforsøg. Tyve høreapparatbrugere (10 nye og 10 erfarne) sammenlignede en lav og en moderat KT indstilling i deres daglige lydomgivelser. De to KT indstillinger blev kombineret med enten en hurtigt- eller en langsomtvirkende kompression i forskellige forsøgsperioder. Testpersonerne foretrak oftest den moderate KT, bortset fra lyttesituationer med svag tale kombineret med langsomtvirkende kompression. I lyttesituationer, som testpersonerne selv angav som vigtige, havde de fleste ikke en præference for enten en moderat eller en lav KT. I forhold til de uerfarne brugere, foretrak de erfarne brugere i højere grad den lave KT.

Resultaterne viste ikke en markant forskel mellem brug af enten lav eller moderat KT. Resultaterne tyder på at erfarne høreapparatbrugere foretrækker mere forstærkning ved svage lydniveauer end nye høreapparatbrugere, og at kompressionshastigheden indvirker på den foretrukne KT. Disse resultater giver information om for hvorledes høreapparater skal tilpasses til nye og erfarne brugere

Acknowledgments

This industrial Ph.D. project was a collaboration between Widex A/S and the Department of Electrical Engineering, the Technical University of Denmark (DTU). The project was carried out between July 2006 and June 2009. The university supervisor was Torben Poulsen, an Associate Professor at the Centre for Applied Hearing Research (CAHR) at DTU. The industrial supervisor was Carl Ludvigsen, head of the Audiological Research Department at Widex. Financial support was provided by Widex A/S and the Danish Agency for Science, Technology and Innovation.

I would like to thank my supervisors, Torben Poulsen and Carl Ludvigsen. For suggesting the original idea and helping secure the funding for the project, I am thankful to Carl. For helpful discussions and for encouraging me at all times, I am grateful to Torben. Torsten Dau, head of the CAHR, was involved in an earlier stage of the project and I would like to thank him for his willingness to collaborate.

The Ph.D. was held on the 18th of January 2010 at the Technical University of Denmark. The examiners were Prof. Dr. Inga Holube from University of Applied Science, Oldenburg, Germany; Assoc. Prof. Birgitta Larsby from Linköping University, Sweden; and Assoc. Prof. Jörg Buchholz, Electrical Engineering, The Technical University of Denmark. The convenor was Assoc. Prof. Jonas Brunskog, also from Electrical Engineering, The Technical University of Denmark. Thank you to the examiners for a good dialog and constructive feedback.

The laboratory listening experiments carried out in chapters 2 and 3 were performed at CAHR, DTU. Thanks to the many fellow students and colleagues that have participated as subjects in the pilot experiments. A special thank you to Alice Lhomond and Andrew Bell for providing an earlier version of the MATLAB code used in the second pilot experiment. A thank you also goes to Professor Kathy Pichora-Fuller for helpful discussion during the development of a pilot protocol. Thanks also the hearing-impaired test participants for their time and patience.

A field trial was carried out as a part of this project, from September 2008 - Febuary 2009 at the Office of Research in Clinical Amplification—Europe (ORCA-Europe), in Stockholm, Sweden. The field trial was supervised by the principal researcher, Karolina Smeds and half of the data collection was carried out by clinical audiologist, Sara Båsjö. Karolina was an excellent sparring partner during the experimental design phase of the trial, as well as helping to finish the data collection for the last five subjects and provide feedback on earlier versions of the field trial manuscript. Sara was very thorough during the hearing aid fittings and following interviews. Additionally, I would like to thank Widex for providing 40 unmarked hearing aids for use during the study, engineer Niklas Bergman for technical assistance and Avesina Hörselrehab hearing aid clinic for assistance in contacting the test participants.

Various colleagues, both at Widex and DTU, have been helpful during the project. From Widex, Carsten Paludan-Müller made the compressor model used in chapters 2 and 3, Brent Kirkwood helped calibrate the BTE microphones used for recordings, Ole Hau helped calibrate the DAI in chapter 3, as well as teaching me how to do a level distribution analysis, and Pernille Friis helped make the boxplots easier on the eye. Erik Schmidt has provided helpful feedback and a good listening ear during all stages of the Ph.D., as well as giving feedback on manuscript drafts. At DTU, a special thanks goes to Iris Arweiler, Filip Munch Rønne and Professor Brian C.J. Moore, while here as guest researcher, for helpful feedback on earlier versions of the thesis.

A big thank you goes to my husband, Lars Korsbæk Connor Sørensen, for his patience and support. Lars' support has ranged from going out and about with BTE microphones on his head, to keeping the refrigerator stocked at home during the intense thesis writing period, to translating the abstract to Danish, as well as introducing me to Ishikawa diagrams. Finally, I would like to thank the hearing-impaired members of my family, who are my motivation to be a good audiologist and hearing researcher.

Contents

List of abbreviations

AIC	Akaike Information Criterion
AT	attack time
B&K	Brüel & Kjær
BTE	behind-the-ear hearing aid
CAHR	Centre for Applied Hearing Research
CAMEQ	Cambridge Method for Loudness Equalisation
COSI	Client Oriented Scale of Improvement
CR	compression ratio
CVR	consonant-vowel ratio
CT	compression threshold
CTLOW	low compression threshold
CTMOD	moderate compression threshold
DAI	direct audio input
dB HL	dB hearing level
dB SPL	dB sound pressure level
DSL[i/o]	Desired Sensation Level Input/Output Algorithm
DSL*m*[i/o]	Desired Sensation Level multistage Input/Output Algorithm
DTU	The Technical University of Denmark
FFT	fast fourier transform
FFtoMic	free field to microphone transfer

GLM	generalized linear model
GLMM	generalized linear mixed model
GUI	graphical user interface
HA	hearing aid
HA user	hearing aid user
ICRA	International Collegium of Rehabilitative Audiology
ISTS	International Speech Test Signal
LTASS	long-term average speech spectrum
NAL	National Acoustics Laboratory
NAL-NL1	National Acoustics Laboratory - Non-Linear 1
NAL-R	National Acoustics Laboratory - Revised
NAL-RP	National Acoustics Laboratory - Revised, Profound
PB	phonetically balanced
q	question
RAU	rationalized arcsine unit
REAG	real-ear aided gain
REAR	real-ear aided response
REIG	real-ear insertion gain
RETSPL	reference equivalent threshold sound pressure level
REUG	real-ear unaided gain
RMS	root mean square
RT	release time
SD	standard deviation
SNR	signal-to-noise ratio
SPL	sound pressure level
SPSS	Statistical Package for the Social Sciences

SSQ	Speech, Spatial and Qualities of Hearing Scale
UCL	uncomfortable loudness level
VC	vowel-consonant combination
WDRC	wide dynamic range compression
4PTA	4 Pure Tone Average

1 Introduction

1.1 Sensorineural Hearing Loss and Loudness Recruitment

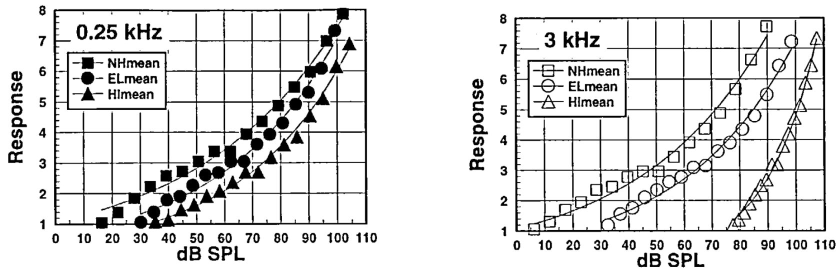
In 2001, the World Health Organization estimated that 250 million individuals globally have a disabling hearing loss, i.e., approximately 4% of the world’s population (Mathers et al., 2001). These problems are even more prevalent among the adult propulation. Davis (1989) found in Great Britain that 16% of the adult population have a “significant” hearing loss. In adults, the most common form of hearing loss is sensorineural, that is, resulting from defects in the inner ear, auditory nerve, or higher centers of the brain. Sensorineural hearing loss is associated with difficulties in understanding speech, particularly in background noise. Socially, the communication difficulties resulting from hearing loss often make it difficult to obtain, perform, and keep employment and the hearing loss may result in social stigmatization and isolation.

The central problem being investigated in this Ph.D. thesis is the lack of audibility for soft speech and environmental sounds experienced by individuals with sensorineural hearing loss. “Soft” in this project refers to sounds that are audible to normally-hearing individuals, and have a spectrum predominately below the spectrum of normal speech at 1 m distance. Soft speech can include both softly-spoken speech and distant speech. In addition to speech, environmental sounds are also an important consideration in this project. Environmental sounds give listeners an awareness of their surroundings and inform the listener about sound events in the environment e.g., the approach of another person or the beeping of a cellphone. This has an important safety element by warning the listener if there is something to which they should react. The soft components of sounds can also potentially inform the listener about the characteristics of objects involved in sound events, e.g., if a falling object is made of metal or glass. Finally, environmental soft sounds can also bring pleasure to the listener, e.g., birds singing.

Hearing loss and the resulting lack of audibility is often managed audiologically with hearing aids (HAs). With the advent of digital HAs in the late 1990s, HAs have increasingly complex sound processing and this is linked to improved HA wearer satisfaction (Kochkin, 2005). In spite of these technological improvements, some problems still remain. In the United States, 26% of HA owners do not use their HAs regularly (Kochkin, 2005). Part of the reason that some HA owners do not use their hearing aids is because they have been fit with either too much or too little amplification, potentially resulting in too little benefit from the hearing aids or experienced loudness discomfort. For instance, 22% of users were dissatisfied with their ability to hear soft sounds and 26% experienced loudness discomfort for loud sounds (Kochkin, 2005). So it seems that the amount of gain provided by hearing aids is not appropriate for all hearing-impaired individuals and the fitting rationales underlying hearing aid fitting still need to be optimised.

A part of the reason that it is difficult to fit hearing aids appropriately for a given hearing loss is because sensorineural hearing loss is also associated with impaired loudness perception (Fowler, 1936). Individuals with sensorineural hearing loss of primarily cochlear origin have elevated hearing thresholds, but uncomfortable loudness levels (UCLs) that are somewhat similar to normally- hearing individual's UCLs. That is, they have a "reduced dynamic range of hearing." As a consequence, when a sound is increased in level above the hearing threshold, the rate of growth of loudness level with increasing sound level is greater than normal. This phenomenon is known as "loudness recruitment." The origin of loudness recruitment is commonly believed to be due to a loss of the non-linear, active processes in the cochlea.

To illustrate loudness growth in individuals with hearing loss, figure 1.1 shows loudness growth curves measured by Robinson and Gatehouse (1996). They measured loudness growth for young and elderly participants with normal hearing, and participants with bilateral, moderate, sloping hearing losses. At 250 Hz, the hearing threshold for the hearing impaired participants was near-normal and loudness growth was similar for both the normally-hearing and the hearing-impaired participants. Whereas at 3000 Hz, the hearing-impaired participants had an average hearing loss of 67 dB Hearing Level (dB HL) and loudness growth was steeper for the hearing impaired participants, particularly at low intensities.



(a) 250 Hz

(b) 3000 Hz

Figure 1.1: Loudness growth functions for three listener groups, normally-hearing (NH), elderly normally-hearing (EL), and elderly hearing impaired (HI) as a function of stimulus level. Response on the y-axis refers to the listener’s rating of loudness from “softest” rated as 1, to “loudest” rated as 8. From Robinson and Gatehouse (1996).

The relationship between degree of hearing loss and uncomfortable loudness levels and frequency was investigated by Pascoe (1988) in a large study including data from 508 ears with a large range of hearing thresholds (0-120 dB HL). Figure 1.2 shows the average Most Comfortable Levels (MCL) and Uncomfortable Levels (UCLs) as a function of degree of hearing loss. It can be seen for a given increase in hearing threshold, that the average MCL and UCL do not increase to the same extent. The result is the average dynamic range of hearing reduces with increasing hearing thresholds. This effect was independent of test frequency, i.e., once the hearing threshold was accounted for, there was no test frequency that was more susceptible to loudness discomfort.

Recently, there has been some controversy about the exact shape of the loudness function for low-input levels (see Marozeau and Florentine, 2007, for review). The classical view of loudness recruitment is that hearing impaired individuals experience a “normal loudness” at threshold and a rapid growth of loudness above threshold. Buus and Florentine (2001) proposed an alternative loudness model called “‘softness

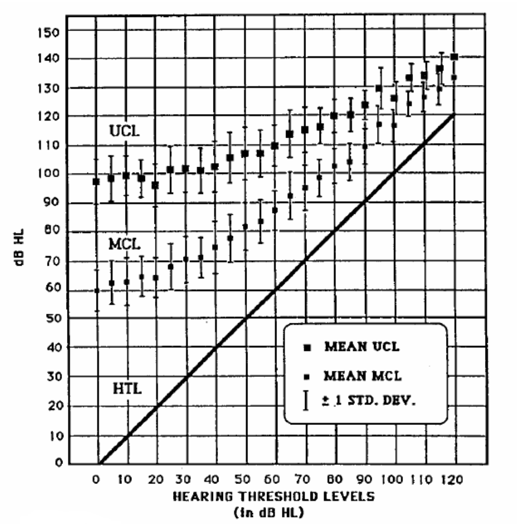


Figure 1.2: Mean most comfortable level (MCL) and uncomfortable loudness levels (UCL) as a function of hearing threshold. Since there was no significant frequency effect, once the effects of threshold were accounted for, the data for 500, 1000, 2000 and 4000 Hz frequencies were pooled together. The error bars indicate ±1 standard deviation. From Pascoe (1988).

imperception". In this model, hearing impaired individuals do not perceive sounds presented at threshold or just above threshold as being "soft", but rather at a greater loudness than experienced by normally-hearing individuals. This proposal has been controversial and Moore (2004) suggested that Buus and Florentine's data can be accounted for by a rapid loudness growth for input levels just above threshold. Recently, Marozeau and Florentine (2007) re-analysed the individual loudness growth functions from from five studies (figure 1.3). They concluded that the loudness growth function for hearing impaired individuals is quite individual: some participants exhibit a rapid growth of loudness just above threshold, while other participants exhibit softness imperception and other participants exhibit some combination of the two phenomena.

Sensorineural hearing loss also influences the temporal and spectral integration of loudness (see Moore, 1998, for review). However, these problems are beyond the scope of this thesis.

In summary, individuals with sensorineural hearing loss have a reduced dynamic

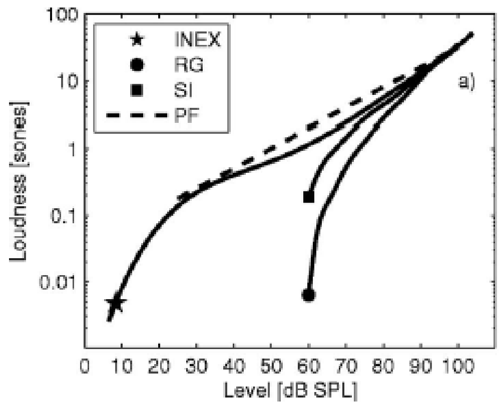


Figure 1.3: Schematised illustration of the three loudness growth functions. The line with the star represents loudness growth in individuals with normal hearing (INEX). The line with the circle represents the classical view of loudness recruitment in individuals with sensorineural hearing loss (RG). The square represents the softness imperception model for individuals with sensorineural hearing loss (SI). The dashed line represents Stevens Power Function (PF), which describes growth of loudness of stimuli at 30 dB SPL and higher levels. From Marozeau and Florentine (2007).

range of hearing and the dynamic range of hearing is negatively related to the degree of hearing loss i.e., the greater the degree of hearing loss, the smaller the dynamic range. Once the degree of hearing loss is accounted for, then the size of dynamic range is not frequency dependent, at least at the main audiometric test frequencies. The reduction of the dynamic range of hearing results in a steepening of the loudness growth curve, and there is some controversy regarding the shape of the loudness growth function at low-input levels. The next section will address how amplitude compression in hearing aids addresses the loss of audibility and loudness recruitment in hearing impaired individuals.

1.2 Overview of Compression in Hearing Aids

Most hearing aid clinicians and researchers would agree that the basic goals of hearing aid fitting for adults include (Palmer, 2002): (i) an improvement in communication ability, in a wide range of conditions, including different speaker types (man, women, children), speaker distances and environmental conditions (in the presence of background noise and/or reverberation); (ii) the hearing aid should improve audibility for environmental, non-speech sounds, to give the wearer a better awareness of their surroundings; (iii) the signal needs to be comfortable for a wide range of frequencies and input levels; (iv) the hearing aid should provide a good sound quality, free of side effects such as own voice occlusion, acoustic feedback, microphone noise and audible distortion; and (v) the hearing aid should meet the hearing aid wearer's communication needs and expectations.

Many of these goals, including improved audibility and improved loudness comfort can be met using non-linear amplification, i.e., an amplifier that adjusts its gain depending on the level of the incoming sound (amplitude compression). The general aim of compression in hearing aids is "to decrease the dynamic range of signals in the environment so that all signals of interest can fit within the restricted dynamic range of a hearing-impaired person" (Dillon, 2001, p. 160). There are many forms of compression available (e.g., compression limiting, medium level compression, etc) but this project is concerned with Wide Dynamic Range Compression (WDRC). In this form of compression, the gain is varied automatically over a wide range of input levels, such that soft sounds receive higher gain and loud sounds lower gain relative to the gain setting for medium level inputs (figure 1.4). In this way, the amplified loudness function will approach the normal loudness function, and in principle provide the listener with a comfortable and audible signal. The technical definitions of the most important compression parameters are given in appendix A.

WDRC is the most commercially widespread form of compression and the use of WDRC compression has been validated. Both Jenstad et al. (1999) and Laurence et al. (1983) found that WDRC succeeded in improving the speech intelligibility of speech presented at low-input levels. In a companion study, Jenstad et al. (2000) gave hearing aid users a "more normal" loudness growth curve. Additionally, Laurence et al. (1983) found in a field trial that the participants rated the WDRC hearing aids better in questionnaires than linear hearing aids fitted with the same gain for moderate-input levels.

There are two general approaches with regards to whether compression should be fast- or slow-acting. Fast-acting compression (often called phonemic or syllabic

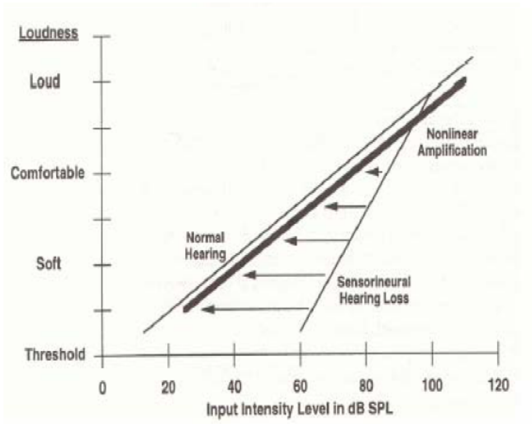


Figure 1.4: Schematised illustration of the effect of WDRC on aided loudness. The line ‘Normal Hearing’ shows the theoretical loudness growth function for an individual with normal hearing. ‘Sensorineural Hearing Loss’ show the theoretical loudness growth functions for an individual with a 50 dB hearing loss at a given frequency. ‘Non-linear Amplification’ shows the theoretical loudness growth function for hearing-impaired individual after provision with WDRC hearing aids. The horizontal arrows represents the gain as a function of input level and it shows the gain is gradually reducing with increasing input-level in an attempt to match impaired loudness growth to the normal loudness function. Modified from Stach (1998).

compression) acts quickly in order to adapt to the varying input levels of different speech segments. The aim of this approach is to improve audibility for the weak phonemes and prevent loudness discomfort for the loud phonemes by adjusting gain to reduce amplitude differences between individual phonemes or syllables (Dillon, 2001; Hickson, 1994; Maré et al., 1992). Attack times are often shorter than 5 ms, and release times may range from approximately 50 ms to approximately 200 ms. In contrast, slow-acting compressors use long release times (between 0.5-20 s), and their goal is to respond to long-term changes in overall intensity rather than to fast intensity changes that occur between speech segments (Dillon, 2001; Ludvigsen, 2001; Moore, 2008).

Previous research on the perceptual benefits of fast- versus slow-acting compression has yielded mixed results. Gatehouse et al. (2006) and Souza (2002) have made comprehensive reviews of the effect of varying time constants and found mixed results. In summary, some studies find no effects of varying time constants (Bentler and Nelson, 1997; Moore et al., 2004; Shi and Doherty, 2008), while other studies find fast-compression to be superior to slow-compression (Jenstad and Souza, 2005; Moore et al., 2004), and other studies again find slow-compression to be superior to fast-compression (Neuman et al., 1998; Hansen, 2002; Schmidt, 2006). Gatehouse et al. (2006) argued for individual differences in benefit from either fast- or slow-acting compression. They made a comprehensive field study with 50 hearing-impaired participants to investigate which participant-related factors (e.g. hearing thresholds, cognitive ability, etc.) could potentially explain the amount of individual benefit from 5 different amplification schemes (2 linear and 3 WDRC schemes.) The 3 WDRC types varied in release time (40 or 640 ms) in the low or high frequency channels. Gatehouse et al. (2006) concluded that the slow-acting compression was perceived as the most comfortable, but it did not necessarily provide the best speech intelligibility for all participants. The benefit for speech intelligibility was highly individual and some participants performed best with fast-acting compression, while other performed best with slow-acting compression. The type of compression that gave the best speech intelligibility was correlated with cognitive ability, i.e., participants with good cognitive ability benefited most from fast-acting compression, whereas participants with poor cognitive ability benefited most from slow-acting compression. Moore (2008) suggests that these individual differences in benefit for different compression speeds may be related to the individual ability to "listen in the dips" using temporal-fine structure cues.

1.3 Effects of Compression on the Signal Characteristics

Many of the perceptual effects of compression are now better understood with the development of relatively recent research interest in the effects of compression on the characteristics of the input signal.

1.3.1 Level Distribution of Speech

Henning and Bentler (2008) examined the effect of compression on the short-term dynamic range[1] of speech in quiet. Figure 1.5 illustrates that fast-acting compression reduces the levels of the speech peaks, as well as raising the level of the speech valleys. Henning and Bentler (2008) also found that the reduction in dynamic range was affected by the number of channels (also seen in figure 1.5), and by shortening the release time. In another study, Souza et al. (2006) also found that when the input signal is speech in the presence of steady-state background noise, the reduction in dynamic range due to compression is not as marked.

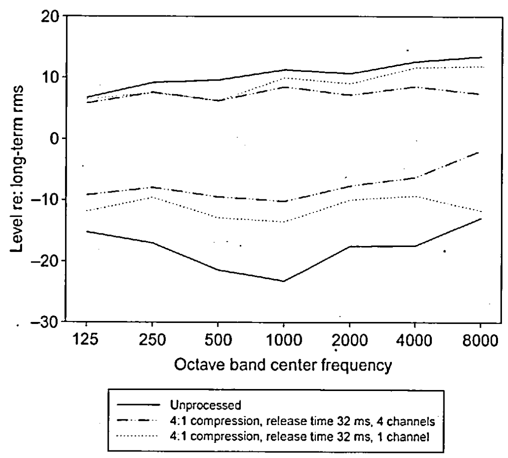


Figure 1.5: The reduction in the short-term dynamic range of speech caused by fast-acting compression. The upper curves represent the peak levels (1st percentile) and the lower curves represent the level valleys (70th percentile). The levels are analysed using a 1-octave bandwidth and are normalised to the root-mean-square (RMS) of the signal. The compressor used a fixed 4:1 compression ratio and used either 1 channel (short-dashed lines) or 4 channels (long-dashed lines). The unprocessed condition is shown for comparison (solid lines). From Henning and Bentler (2008).

[1] The short-term dynamic range was defined as the 1st and 70th percentiles for an 120-ms analysis window.

1.3.2 Vowel-Consonant Ratios and Speech Envelope

Jenstad and Souza (2005) considered the effect of compression on the consonant-vowel ratio (CVR) and the envelope[2] depth for single phonemes presented in quiet. The compressor used single-channel processing with compression ratio of 3:1. They found that compression increases the CVR i.e., compression increases the level of the consonant relative to the vowel (figure 1.6) and this effect was reduced for longer release times. They additionally found that compression also reduced envelope depth and this too was affected by release time. They then measured phoneme recognition scores for hearing-impaired participants with moderately-severe hearing losses using the same compressed material and they found that the changes to the envelope depth and CVR were mildly correlated with phoneme recognition score.

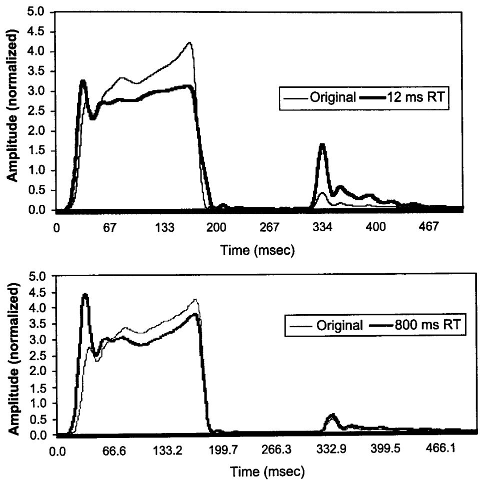


Figure 1.6: An example of the envelope for the syllable /ip/ presented at 65 dB SPL for a single-channel compressor with a 12-ms (top) and 800-ms (bottom) release time. The thin line is the normalised amplitude envelope for the unprocessed syllable, and the thick line is the normalised amplitude envelope for the processed syllable. The attack time was fixed at 4 ms and the compression ratio at 3:1. From Jenstad and Souza (2005).

[2] Qualitatively, the envelope of a signal is that boundary within which the signal is contained, when viewed in the time domain.

1.3.3 Signal-to-Noise Ratio

Until recently, results regarding the effect of fast-acting compression on the signal-to-noise ratio (SNR) at the compressor output have been contradictory. Olsen et al. (2005) found that compression *improves* output SNR relative to the input SNR, while Souza et al. (2006) using the same analysis technique found that compression *degrades* output SNR. Much of the apparent contradiction in results can be explained by considering the input SNRs and the modulation characteristics of the noise.

Naylor and Johannesson (2009) investigated the effect of input SNR and noise type on the output SNR of a commercial hearing aid. The input signal was sentence material presented at varying SNRs with three possible noise types: unmodulated speech noise, two-speaker noise and reverse single-speaker noise. Figure 1.7 shows for unmodulated speech noise (N1), that fast-acting compression degrades the output SNR from the hearing aid, particularly at positive input SNRs. This was attributed to the more instantaneous gain being applied during the pauses of speech, hence amplifying the noise up. Another pattern is seen in figure 1.7 for reverse speech (N3), which has the same modulation characteristics as the speech signal. At positive input SNRs, the output SNR was also degraded, but for negative input SNRs, fast-acting compression *improved* the output SNR. This was attributed to gain reduction being applied to the high-level components of the noise signal, and extra gain was applied to the relatively weak components of the speech signal, during the pauses in the noise. These observations seemed sufficient to explain the otherwise contradictory findings of Olsen et al. (2005) and Souza et al. (2006) because Olsen et al. (2005) using a modulated noise presented at negative SNRs and Souza et al. (2006) used an unmodulated noise presented at mostly positive SNRs.

Naylor and Johannesson (2009) also examined the effects of compression parameters on the output SNR. They found that lengthening the time constants, lowering the compression ratio and decreasing the number of channels reduced the effects of compression on the output SNR.

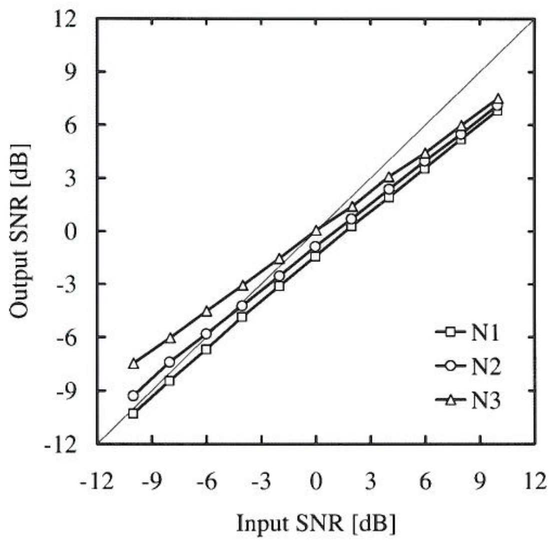


Figure 1.7: Output signal-to-noise ratio (SNR) from a measurement hearing aid for an input of speech at varying SNRs. The hearing aid was a single-channel with fast-acting compression and a fixed compression ratio of 2:1. The three noise types were: N1, unmodulated noise; N2, 2-talker modulated noise; and N3, reverse speech. From figure 3 in Naylor and Johannesson (2009).

1.3.4 Other Effects

Stone and Moore (2008) found that when the input signal to the system is a mixture of voices from different talkers, fast-acting, multi-channel compression introduces "cross-modulation" between the voices because the compressor applies the same time-varying gain to all of the voices. Hence, voices that are independently amplitude modulated acquire a common component of modulation at the output of the compressor. This gives the impression that the voices become "perceptually fused" and affects speech intelligibility of heavily compressed speech or vocoded speech.

Bor et al. (2008) found that compression leads to spectral flattening of the formant peaks of vowels (i.e., reduction in the amplitude of the peaks). They also measured vowel identification in 20 listeners with mild to moderately-severe sensorineural hearing losses and correlated vowel identification with a spectral flattening effect. Leijon and Stadler (2008) modeled the effects of fast-acting compression on speech-in-noise.

They argue that while fast-acting compression improves audibility for speech, it degrades speech information transmission due to a reduction in spectral contrasts.

1.3.5 Summary of the Effects of Compression on the Signal Characteristics

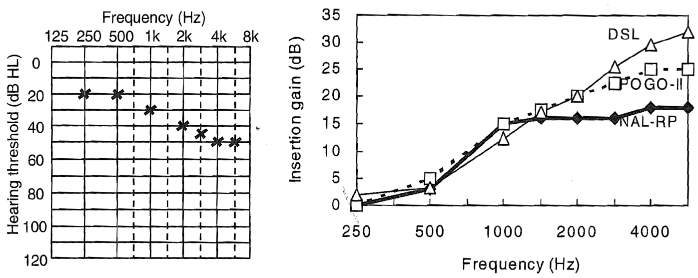
In summary, compression has a number of effects on the characteristics of speech. Some of these effects are potentially helpful, e.g., the increase in the low-level components of the signal (Henning and Bentler, 2008), including some low-level consonants (Jenstad and Souza, 2005). However, some of the effects are potentially harmful, e.g., the introduction of co-modulation to auditory objects that are otherwise uncorrelated (Stone and Moore, 2008), the spectral flattening of formant peaks (Bor et al., 2008), and potential degradation of the SNR for certain input signals (Souza et al., 2006; Naylor and Johannesson, 2009). It is important to note that the effects of compression on the signal can be ameliorated by increasing the time constants, and by reducing the compression ratio.

1.4 Hearing Aid Gain Prescription

As discussed in section 1.2, the general aim of hearing aid fitting is to improve communication ability by improving audibility, while maintaining a comfortable signal at a wide range of input levels. In order to meet these goals, the individual hearing loss should be considered, as hearing losses vary widely in their degree, configuration and type. In order that audiologists and hearing aid manufacturers have an initial estimate of how much gain is appropriate for a given individual, hearing aid prescription targets need to be available. These targets are specified by a hearing aid rationale, which are the principles and assumptions used to meet the goals of the hearing aid fitting. Hearing aid rationales have been developed for both linear and non-linear hearing aids. The gain targets prescribed by rationales can be based on either hearing thresholds, supra-threshold loudness measures, or some combination of the two. This review concentrates on threshold-based prescription rationales, since these are most widely in use.

1.4.1 Linear Hearing Aid Prescription

For mild-to-moderate hearing losses, most linear HA prescriptions recommend that the overall gain should be approximately half the amount of the hearing loss (known as the half-gain rule). The prescriptions usually differ in the prescribed frequency response due to differences in the underlying rationales. For instance, the National Acoustics Laboratory-Revised (NAL-R, Byrne and Dillon, 1986) and National Acoustic Laboratories-Revised, Profound (NAL-RP, Byrne et al., 1990) prescriptions use loudness equalisation rationales, i.e., all frequency regions of the speech spectrum should be amplified to MCL, such that they contribute equally to its loudness. In contrast, Prescription of Gain and Output II (POGO II, McCandless and Lyregaard, 1983) prescription uses a half-gain rationale with a low-frequency reduction to compensate for the upward spread of masking from ambient noise. Finally, the Desired Sensation Level (DSL, Seewald et al., 1993) rationale aims to make speech in each frequency region comfortably loud, although not necessarily equally loud. Figure 1.8 shows the gain/frequency target curves for NAL-RP, POGO II and DSL prescriptions for a given mild, sloping hearing loss. The figure illustrates that there is no universal agreement about which gain/frequency targets are appropriate for linear amplification of medium input levels.



(a) Audiogram (b) Insertion gain targets

Figure 1.8: The insertion gain targets for a mild sloping hearing loss prescribed by three different linear fitting rationales: DSL, POGO II and NAL-RP. Figure from Dillon (2001, p. 244).

1.4.2 Non-Linear Hearing Aid Prescription

Non-linear hearing aid prescriptions are more complicated than linear hearing aid prescriptions because they provide targets for inputs at different sound pressure levels. The lack of consensus about hearing aid rationales is even more marked for prescription of non-linear amplification. Three threshold-based, non-linear rationales are described in the following subsections in alphabetic order: the Cambridge method for loudness equalization (CAMEQ) rationale (Moore et al., 999a,b), the Desired Sensation Level - input/output (DSL[i/o]) rationale (Cornelisse et al., 1995) and the National Acoustic Laboratories - Non-Linear, version 1 (NAL-NL1) rationale (Byrne et al., 2001).

CAMEQ

The goal of CAMEQ is to achieve a flat loudness density pattern across frequency for speech, while the overall loudness should be equal to that of normally-hearing individuals (Moore et al., 999a,b). To implement this, loudness is calculated in critical bands for a given hearing loss using the loudness model specified by Moore and Glasberg (1998). At moderate-input levels, the specific loudness pattern evoked by speech at 65 dB SPL should give an equal loudness per critical band in the range 500-4000 Hz, and the overall loudness should be similar to that evoked in a normal individual by 65-dB speech (23 sones for binaural listening). For low-input levels, speech with an overall level of 45 dB SPL should just be audible in all frequency bands from 500-4000 Hz, provided that this does not require compression ratios of greater than 3:1. There were two reasons given as to why 45 dB SPL should be the lowest level of audible speech: (i) it corresponds roughly to the lowest level of speech that a person needs to understand in everyday life (Pearsons et al., 1977); and (ii) empirical evidence suggests that if more gain is given for low-input levels, hearing aid users complain that noises in the environment are too intrusive (Laurence et al., 1983). The constraint of a maximum 3:1 compression ratio was due to empirical evidence that higher compression ratios have an adverse effect on speech intelligibility (Moore et al., 1992; Plomp, 1988).

DSL[i/o] and DSLm[i/o]

The aim of DSL[i/o] is to give complete restoration of audibility of speech sounds (within the constraints of residual hearing ability.) The rationale fits the so-called "expanded dynamic range" of environmental sounds into the reduced dynamic range of hearing (Cornelisse et al., 1995). The extended dynamic range, at a given audiometric frequency, is equal to the range from a normal-hearing individual's threshold up to the hearing-impaired individual's UCL. Figure 1.9 shows how this is calculated at each audiometric frequency. At low input levels, the average normal-hearing threshold in the soundfield is mapped to the hearing-impaired individual's hearing threshold in the ear canal (the lower dashed line). At high input levels, the hearing-impaired individual's UCL in the soundfield is mapped to the impaired UCL in the ear canal (upper dashed line). For levels in-between, the compression ratio is prescribed by fitting either a straight line (known as the *linear* procedure) or a curvi-linear line (known as the *curvilinear* procedure, shown in figure 1.9). In theory, this should mean that the hearing-impaired individual has the same audibility for low-level inputs as a normally-hearing individual.

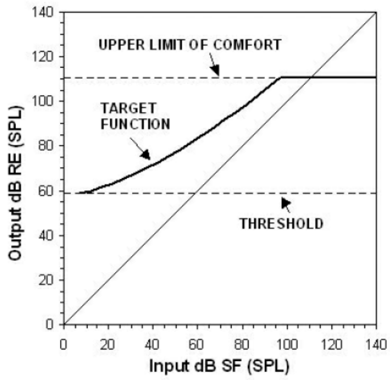


Figure 1.9: The DSL[i/o] target input/output curve (thick line) for a 50 dB HL threshold at 1000 Hz, computed using the curvi-linear procedure. The input levels are plotted in dB SPL in the soundfield (SF). The output levels are plotted in dB SPL in the real ear (RE). The dashed lines represent the impaired hearing threshold (lower dashed line) and the impaired upper limit of comfort (upper dashed line). The diagonal line represents zero gain. Figure 10 from Scollie et al. (2005).

The newly-revised, multi-stage DSL*m*[i/o] rationale (Scollie et al., 2005) now includes a linear stage for low-input levels and an output-limiting stage for high-input levels. The compression threshold is prescribed in the range between 30 dB SPL and 70 dB SPL[3] and varies with hearing threshold at the given frequency. The rationale for the compression threshold was based on theoretical considerations because "very little evidence exists that determines an appropriate compression threshold [CT] prescription, particularly as it would relate to single vs multichannel compression devices" (Scollie et al., 2005, p. 184). For mild hearing losses, the DSL*m*[i/o] prescribes a low compression threshold (approx. 30 dB SPL) in order to preserve audibility for low-level speech (assumed to be 52 dB SPL), but restricts audibility for input levels below this, which are assumed to be "low-level background noise". For profound hearing losses, the compression threshold is high to prevent loudness discomfort.

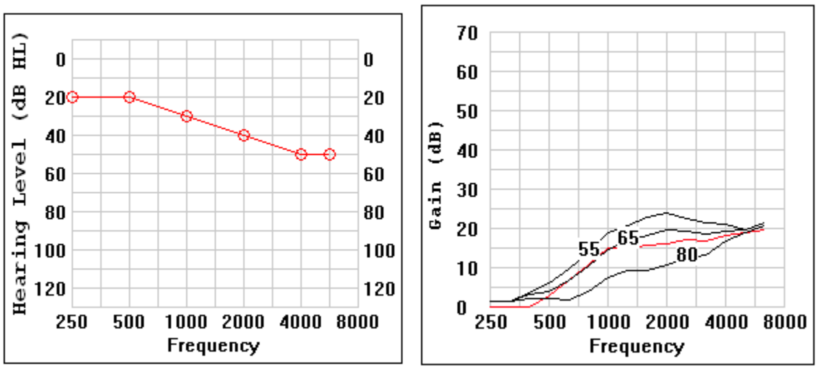
Among other changes to the DSL[i/o] rationale, the DSL*m*[i/o] rationale prescribes on average 6 dB less gain for adult fittings compared to the original DSL[i/o]. The gain reduction was based on the findings that adults have a lower preferred listening level (Scollie et al., 2005) and also clinical evidence that most hearing aid dispensers fitting to the DSL rationale, often fit to a reduced gain, particularly in the high frequencies (Scollie et al., 2005). Other modifications include corrections for conductive hearing losses and channel summation in multi-channel hearing aids.

NAL-NL1

The rationale of the NAL-NL1 is to maximize speech intelligibility, subject to the overall loudness of speech at any level being no more than that perceived by a normally-hearing individual. To derive the gain/frequency curve for each input level, two theoretical models were used. The first model is a modified version of the Speech Intelligibility Index (ANSI S3.5 (1997), with modification for the effects of hearing loss and high-input level from Ching et al. (1998)). The second model is the Moore and Glasberg (1997) loudness model, which accounts for both loudness at various input levels, as well as loudness summation across frequency (critical bands). In contrast to the NAL-RP prescription, the NAL-NL1 does not have loudness equalisation

[3] relative to the free field and specified in 1/3-octave bands

as an explicit goal. However, because they use an approach to maximise speech intelligibility input, the prescribed gain/frequency curves for a mid-level input (65 dB SPL) are similar to the NAL-RP gain/frequency curves.



(a) Audiogram (b) Insertion gain targets

Figure 1.10: The left panel shows the audiogram of a mild, sloping hearing loss. The right panel shows NAL-NL1 insertion gain targets for three speech input levels of 55, 65, and 80 dB in black and the NAL-RP target in red. From a screen capture from the NAL-NL1 software (Brewer, 2005).

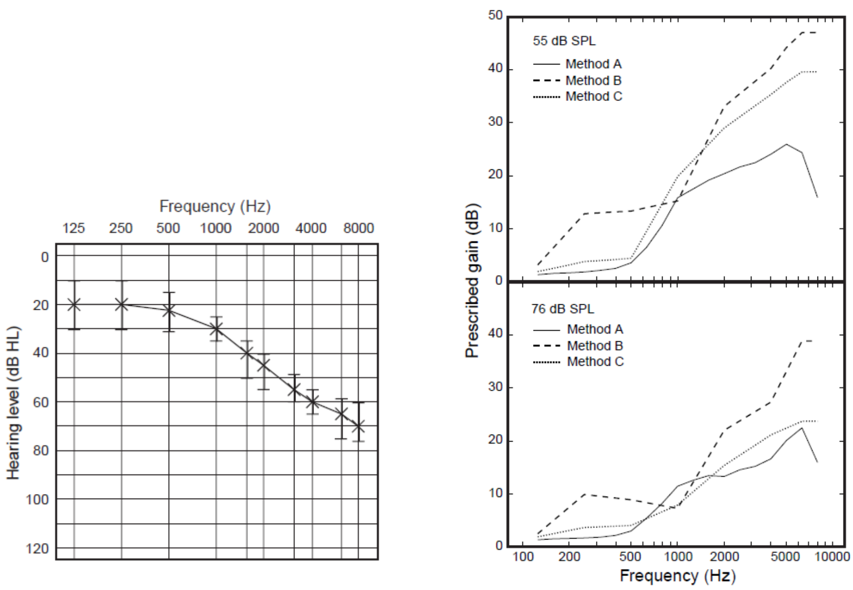
.

For low-input levels, NAL-NL1 uses a default broadband compression threshold of 52 dB SPL. This is based on empirical research by Barker and Dillon (1999), Barker et al. (2001) and Dillon et al. (1998) found that for single-channel hearing aids with fast-acting compression, the majority of hearing aid users prefer a compression threshold in excess of 60 dB SPL.

A NAL-NL version 2 (NAL-NL2) is under development. Keidser and Dillon (2006) suggest that NAL-NL2 is likely to include corrections for the acclimatisation effect, different targets for adult and child fittings and the inclusion of targets for a positive test for dead regions. Changes to the rationale are likely to include reduced gain, particularly in the low frequencies and increased compression ratios, particularly at high-input levels.

Comparison of Gain Targets

Figure 1.11 shows the prescribed gain targets for CAMEQ, DSL[i/o] and NAL-NL1 for the given audiogram. The upper panel shows prescribed gain for a 55 dB SPL input signal and the lower panel for an input signal at 76 dB SPL. The three rationales vary considerably in prescribed gain. For a 55 dB input level, NAL-NL1 prescribes substantially less gain than the other two rationales. DSL[i/o] and CAMEQ recommend similar gain in the high frequency region from 1 kHz. In the low frequency region, CAMEQ prescribes gain close to NAL-NL1, whereas DSL[i/o] gives more low frequency gain. For a 76 dB input level, NAL-NL1 and DSL[i/o] prescribe similar gain, NAL-NL1 somewhat more than CAMEQ in the 1-kHz region. DSL[i/o] prescribes more gain than the other two rationales, particularly above 2 kHz.



(a) Audiogram (b) Insertion gain targets

Figure 1.11: The left panel shows the considered audiogram and the right panel shows insertion gain targets for three hearing aid rationales (methods) for a 55 dB and 76 dB SPL input levels. Method A is based on NAL-NL1, Method B is based on DSL[i/o] and Method C is based on CAMEQ. From figure 4 in manuscript II in Smeds (2004b).

In spite of the quite different gain recommendations from the three nonlinear rationales, there is little empirical comparison between them. Marriage et al. (2004) investigated the gain requirements for hearing aid users with mild-to-moderate hearing losses. The participants were 20 experienced and 20 new hearing aid users. They were initially fitted using three different hearing aid rationales: CAMEQ, DSL[i/o] and a third rationale called CAMREST (Moore, 2000). At the fitting appointment and one-week post-fitting, the gains were adjusted when required by the minimim amount necessary to achieve acceptable fittings. On average, the adjustments were smallest for CAMEQ, wheres the largest changes were for DSL[i/o] and these were mostly to reduce the high frequency gain. Moore et al. (2001) found similar results for a smaller group of experienced hearing aid users. Marriage et al. (2004) also found that the experienced users preferred on average 3 dB more gain than the new users.

Smeds (2004a) investigated in a field trial, the preferred gain for 21 new hearing aid users with mild hearing loss. The participants compared two fitting rationales simultaneously in two different hearing aids programs. The rationales were (a) NormLoudn which was based on the CAMEQ rationale; and (b) LessLoudn which provided on average 5 dB less gain than NormLoudn in the range 1-4 kHz. Nineteen out of the twenty-one participants reported that they overall preferred LessLoudn over NormLoudn. This result may seem in contradiction to the findings of Moore et al. (2001) and Marriage et al. (2004), but there were a few key differences between these studies. Firstly. the participants in Smeds (2004a) were instructed to report the *preferred* hearing aid program, whereas the participants in Moore et al. (2001) and Marriage et al. (2004) were instructed to report when the hearing aid fitting was *acceptable*. Secondly, Smeds (2004a) used new hearing aid users with mild hearing losses, whereas Moore et al. (2001) and Marriage et al. (2004) used a combination of new users and experienced users with a higher degree of hearing losses.

Finally, Smeds et al. (2006a,b) investigated the preferred hearing aid volume (calculated loudness) relative to the NAL-NL1 fitting. Smeds et al. (2006a) included 24 participants with mild sloping to moderate hearing losses in a laboratory study and Smeds et al. (2006b) included 15 participants with similar audiograms in a field and laboratory study. In each study, the participants were a mix of new hearing aid users and experienced hearing aid users. The hearing aids were fitted to the NAL-NL1 rationale but the participants could adjust the volume control according to their preference. In both the laboratory and field studies, they found that the hearing-impaired participants preferred less gain than provided by NAL-NL1, which should in principle provide "normal loudness."

In summary, there is a lack of consensus between the major hearing aid rationales about how much gain is appropriate, particularly at low-input levels, and the lack of empirical evidence to suggest which hearing aid fitting rationale is best. Additionally, there is evidence to suggest that all three rationales provide too much gain than is otherwise preferred by hearing aid users.

1.5 Current State of Knowledge about the Preferred Compression Threshold

As discussed in section 1.4, there is little consensus about how much gain at low-input levels is appropriate. There has also been little direct investigation into preferred gain for low-input levels. This is probably due to that it is difficult to adjust one parameter (e.g., compression ratio, release time etc.) without it affecting gain for a wide range of input levels. One parameter that is of special interest in this project is *compression threshold* (defined in appendix A) because lowering the compression threshold increases gain at low-input levels, when the compression ratio is kept constant and the gain for medium levels is fixed (shown in figure 1.12). This improves audibility for low-input levels, as well as in theory normalising the loudness function. There are a wide variety of compression thresholds available among the various commercial hearing aids from different manufacturers. WDRC compression thresholds may be as low as 25-30 dB SPL or as high as 65 dB SPL (Henning and Bentler, 2008).

Barker and Dillon (1999), Barker et al. (2001) and Dillon et al. (1998) performed a series of field trials investigating the preferred compression threshold with participants with hearing losses ranging in degree from mild to severe. The participants compared a moderate compression threshold (∼65 dB SPL) with a low compression threshold (40-57 dB SPL) in two different programs of single-channel, fast-acting hearing aids with a fixed compression ratio of 2:1. The participants wore the hearing aids in their

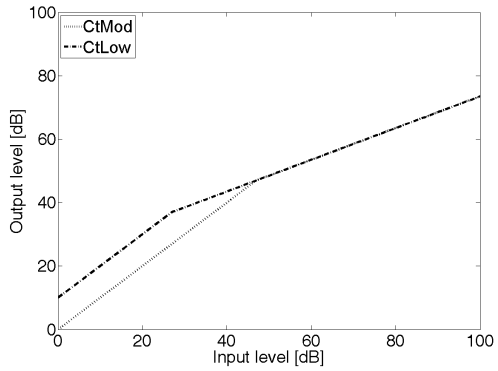


Figure 1.12: Illustration of the effect of varying compression threshold (CT), while holding gain for moderate- and high-input levels constant. The two CTs illustrated here are at a low- and moderate-input levels. The compression ratio is fixed at 2:1.

own daily listening environments for 4 weeks and then came back and reported to the experimenter their preferred hearing aid program, both overall and in a number of specific listening environments. Overall, about two-thirds of the participants preferred the moderate compression threshold over the low compression threshold. All of these studies used single-channel, fast-acting compression and the results may not be applicable to other types of compression, using multiple compression channels or other time constants. As discussed section in 1.3, varying the number of channels and the compression time constants influences the signal characteristics, including the dynamic range, envelope depth, modulation characteristics and the output SNR. Since the compression threshold influences how often the compressor is activated for a given signal, these potentially negative or positive signal effects could influence the preferred compression threshold.

1.6 Methodologies to Investigate Hearing Aid Parameters

Hearing aid settings are often evaluated using a combination of laboratory listening experiments and/or field trials. Laboratory listening experiments often include either subjective evaluation of the hearing aid settings and/or aided speech intelligibility testing. The main advantages of laboratory listening experiments are that they are often easier to implement practically, they are less time consuming and they offer the experimenter control over the experimental conditions (signal type, signal level, presentation modality, instructions, length of experimental session etc.) In contrast, the main advantage of field trials is that they offer greater face validity, but at the cost of loss of some experimental control. Since field trials are more time consuming and expensive to implement, the majority of evaluations of hearing aid settings are performed in the laboratory.

Review of field trial results and corresponding lab results indicate that the degree of correlation between lab and field data is mixed. In many previous studies investigating preferred frequency response for linear hearing aids, there has been moderately-good correlation between preferred frequency response in the lab and in the field (e.g., Byrne and Cotton, 1988; Kuk and Pape, 1992, 1993; Preminger et al., 2000). These earlier studies have identified procedure, instruction, and choice of signal as important factors in determining the relationship between lab and field studies.

For non-linear hearing aids, the correlation between lab data and field data is not always statistically significant. For instance, Smeds et al. (2006b) investigated preferred listening levels and found for the hearing impaired participants, that for low-input levels, the participants preferred more gain in the lab than in the field. Another example of poor correlation is Savage et al. (2006), who found that most of the 19 hearing impaired participants did not exhibit a clear preference for output limiting method (compression limiting or peak-clipping) in the field but had a clear preference in the lab. Conversely, Xu et al. (2008) investigated preference for compression release time and found that most participants were better at indicating a preference in the field than in the lab.

Some of the potential problems with the laboratory is that the task for the participant is unnatural and the signal presentation is static. The signal of interest is a single voice presented from a fixed sound source and the participant makes judgments of the sound quality, without having any other tasks. In real-life, there are auditory objects (or sound events) that may emerge and the listener has to detect, localise, identify and decide whether the new auditory object is worth paying attention to. This is also the case when the listener may be involved in a task where audition is not so important (e.g., reading). These attentional effects are potentially very important in the consideration of soft sounds. When extra amplification for low-input levels is provided, it will provide audibility for more auditory objects in the environment. Some of the objects may be interesting and relevant, for example, a ringing cellphone, but some of these objects may not be interesting for the listener, e.g., a computer fan. In particular, for new hearing aid users, they may hear sounds that they haven't heard for years and they may respond either positively or negatively to this.

In summary, there are some examples of both good and poor correlations between preference for hearing aid setting in the lab and corresponding field trial. Given that the degree of correlation between lab and field data is not always good, particularly when non-linear amplification is considered, it is prudent to include a laboratory and a field trial investigation in determining the optimal gain for low-input levels.

1.7 Hearing Aid Acclimatisation Effect

The term acclimatisation was first coined by Gatehouse (1989) to indicate changes in speech recognition associated with HA provision. Since then the definition has been broadened to:

> "Auditory acclimatization is a systematic change in auditory performance with time, linked to a change in the acoustic information available to the listener. It involves an improvement in performance that cannot be attributed purely to task, procedural or training effects" (Arlinger et al., 1996, p. 87S).

Keidser et al. (2008) discusses four types of HA acclimatisation that have often been investigated and discussed in the HA literature:

Improvements to speech intelligibility Some researchers have demonstrated statistically significant improvements in speech intelligibility in the 10 to 18 week period following HA provision (e.g., Cox and Alexander, 1992; Horwitz and Turner, 1997). However, these results are controversial and other researchers have found no evidence of changes in speech intelligibility (e.g., Humes et al., 2002; Humes and Wilson, 2003). There is evidence that improvements in speech intelligibility, following linear HA provision, are greater for speech at high input levels than moderate and low input levels (Connor, 1999; Gatehouse, 1989, 1992, 1993; Munro and Lutman, 2003). Yund et al. (2006) found evidence that improvement in speech intelligibility is greater for non-linear HA processing than linear HA processing, and they speculate that the greater acclimatisation effect occurs for non-linear processing because it gives the HA user access auditory input at a wider range of frequencies and input levels.

Gain adaptation Keidser et al. (2008) found a difference in preferred gain for medium input levels between new and inexperienced HA users. This difference was in the magnitude of 2 dB. New HA users gradually increased their gain preference in the 13-month period following non-linear HA provision, but this "gain adaptation" was not complete 13 months post-fitting. Similarly, Marriage et al. (2004) found that experienced HA users were prepared to accept on average 2.7 dB more gain in order to achieve an "acceptable" HA fitting compared to new HA users. These results contrast with earlier findings that preferred gain does not change up to 3 years following linear HA provision (Cox and Alexander, 1992; Horwitz and Turner, 1997; Humes and Wilson, 2003).

Changes to loudness perception There is evidence of changes in loudness perception following HA provision, particularly at medium- and high-input levels (Keidser et al., 2008; Munro and Trotter, 2006; Olsen et al., 1999; Philibert et al., 2002, 2005). Evidence from Keidser et al. (2008) and Philibert et al. (2005) suggest that this effect occurs during the 1-2 month period following HA provision.

Reported subjective benefit Improvements in subjective measures of benefit (i.e., questionnaire data) following HA provision. Again, these results are conflicting with some researchers finding a benefit in the time (e.g., Bentler et al., 1993) following HA provision and others finding no benefit (e.g., Humes et al., 2002; Humes and Wilson, 2003).

In summary, there is some controversial evidence of changes in auditory performance and preferred gain following HA provision. Given that acclimatisation is related to access to new auditory information following HA provision, it is plausible that it would take time following HA provision for the HA user to learn to make optimal use of this information, as well as learn to adapt to the extra gain for a range of everyday sounds. Multi-channel WDRC compression gives access to auditory information at a wider range of input levels and frequencies than linear amplification. So it could be that previous HA experience is an important factor in preference for gain at low input levels.

1.8 Overall Summary and Direction for the Ph.D. Project

The goal of compression in HAs is to improve audibility, particularly for speech, while maintaining a comfortable amplified signal for the HA user at a wide range of input levels. Unfortunately, there is a lack of consensus about how much gain at low-input levels is appropriate and there is a lack of empirical investigation into the matter. The few studies that have directly investigated this issue, investigated preference for compression threshold in the late 1990s using a single-channel, fast-acting compression HA. Since these experiments were performed, new research has found that compression, particularly fast-acting compression, degrades the signal and potentially the subjective signal quality. As a result, these earlier results concerning preferred compression threshold can not necessarily be applied to other HAs with multiple channels and other time constants.

The central question in this overall project is to determine under which circumstances (if any) "soft sounds" in the environment, should be amplified to audibility? This problem was tackled in different sections in this report.

Chapter 2 presents three pilot experiments concerned with optimising the laboratory listening methodologies that are typically used to investigate preferred HA settings. The first experiment was about using paired comparison data to derive scaled information about preferred HA setting. The second two experiments were concerned with improving the ecological validity of lab experiments by manipulating the listener's attention using tasks and instructions.

Chapter 3 examines in the laboratory, the relationship between preferred HA compression threshold and release time, as well as the influence of signal and instruction on this preference.

Chapter 4 examines using a combination a laboratory experiment and a field trial, the relationship between preferred compression threshold, compression time constants, listening situation and previous HA experience.

Chapter 5 summarises the main outcomes of this work and discuss the possible implications, in terms of gain prescription for low-input levels in various listening environments, different time constants and different client groups.

2

Pilot experiments to find a method for use in a laboratory listening experiment

Three pilot experiments using normal-hearing participants were performed within semesters 2 and 3 of this Ph.D. project. The general aim of these experiments was to find a methodology that was suitable for the assessment of different hearing aid settings. The experiments were:

Pilot I Deriving a response scale from paired comparison data. Previously unpublished.

Pilot II Influence of listener task on ratings of pleasantness for everyday sounds. Based on a contribution to the proceedings of International Symposium on Auditory and Audiological Research (ISAAR) in Connor and Poulsen (2007).

Pilot III Influence of the instruction and signal on the preference for compression setting in normal-hearing listeners. Previously unpublished.

2.1 Pilot I. Deriving a response scale from paired comparison data

2.1.1 Introduction

The overall aim of this Ph.D. project is to consider how sounds at low input levels should be amplified in hearing aids. In order to address this, there needs to be a method that is appropriate for assessing hearing aid wearers subjective evaluation of various hearing aid settings. There are two common methods employed in the hearing aid literature for assessment of various hearing aid settings: categorical scaling and paired comparisons (see Byrne and Cotton, 1988; Keidser, 1996; Keidser et al., 2005; Kuk and Pape, 1993, for example).

In categorical scaling, participants are asked to judge a property of a stimulus using verbal or numerical categories, for example, as shown in figure 2.1. The advantages of categorical scaling include that the task is easy for the test participants to understand, and the task gives an absolute judgment rather than a relative judgment. However, there are a number of known biases in categorical scaling (see Zielinski et al., 2008, for review). One of these biases is that participants tend to answer in the middle of the scale (central response bias tendency), which potentially results in an unequal distance between response categories (e.g., a greater difference between categories 7 and 8, than categories 5 and 6). Another disadvantage of categorical scales is that they are less sensitive to subtle differences between stimuli than a paired comparison method (Eisenberg et al., 1997).

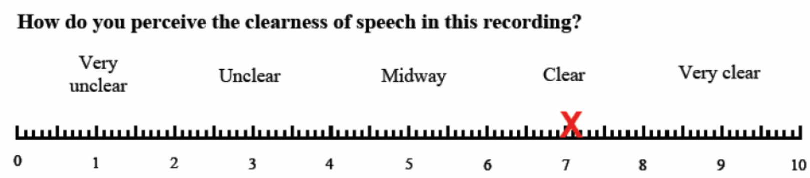


Figure 2.1: Example of a categorical scale from Schmidt (2006, p. 92)

In paired comparisons, participants compare two stimuli and report which of the two is preferred in relation to a certain parameter (e.g. loudness, clarity, sharpness). The main advantage of this method is that it is sensitive to subtle differences between stimuli (Eisenberg et al., 1997). However, the disadvantages include that the method is time consuming, and it provides only relative information.

Bradley-Terry Luce Model

In the proposed methodology, paired comparison data is used to derive a response scale, similar to a categorical scale, using the Bradley-Terry Luce (BTL) model (Bradley and Terry, 1952; Luce, 1959). The principle for deriving a scale is that the perceptual distance between two stimuli can be estimated by considering how often stimulus *a* is chosen over stimulus *b*, i.e., the preference probability (p_{ab}). The preference probability is determined in a paired comparison experiment and used as input to equation 2.1 below.

$$p_{ab} = \frac{v(a)}{v(a) + v(b),} \tag{2.1}$$

where p_{ab} is the empirically determined probability of preferring *a* over *b*, and $v(a)$ and $v(b)$ are the derived scale values of the stimuli. Note that one of the scale values, $v(a)$ or $v(b)$ can be freely selected by the experimenter.

The advantage of using the BTL model to derive scale values are that the mathematical properties underlying the derived scale are known. For example, if the scale value of stimulus *a* is 7.5 and the scale value of stimulus *b* is 1.5, then it is because stimulus *a* was preferred 5 times more often than stimulus *b*. When there are many stimuli to be compared and scaled, the scale values can be fitted to the BTL model using a MATLAB function called `OptiPt.m`, which was published by Wickelmaier and Schmid (2004). It order to fit the BTL model, there are some assumptions that need to be met and these assumptions can be tested, as described below.

Test 1: Transitivity The main assumption of the investigated model is that of unidimensionality, i.e., participants use the same listening criteria regardless of which pair of stimuli are tested. This is tested by considering the concept of transitivity, described in equation 2.2.

$$\text{If } a \geq b \text{ and } b \geq c \text{ then } a \geq c \tag{2.2}$$

In cases when uni-dimensionality does not hold, other models with more complex structures can be considered, such as the "preference tree model" or the "elimination by aspects model" (see Zimmer et al., 2004, for review).

Test 2: Likelihood Ratio Test The derived scale values fitted using the BTL model can be assessed statistically using a likelihood-ratio test. The likelihood ratio test compares the derived scale values to the direct magnitude-estimates obtained.

2.1.2 Method

The current study investigated if the BTL model can be used for comparisons of hearing aid settings to derive information about preference order and scale. The test signal, containing speech and noise, was processed offline in a compressor model with different compression ratios (CR=1, 1.5, 2 and 3). Normal-hearing participants made paired comparisons of the original test signal and the compressed stimuli. The results of the paired comparisons were examined for transitivity, and fitted to the BTL model to derive scale values. The derived scale values were then tested statistically using a likelihood ratio test.

Test Participants

Twelve participants took part in the current study. All participants were employed in the Audiological Research Laboratory in Widex A/S. Eleven of the twelve participants had normal hearing and one had a mild hearing loss. The ages of the test participants were not noted at the time of testing.

Test Stimuli

The five test stimuli were (i) the original test signal; (ii) test signal compressed with CR=1:1 [linear condition]; (iii) test signal compressed CR=1.5:1; (iv) test signal compressed CR=2:1; and (v) test signal compressed CR=3:1.

The original test signal was from the Ph.D. project of Erik Schmidt (2006). It consisted of a 7 second long recording of a male Danish speaker using a loud vocal effort at 75 dB SPL and party noise presented at 0 dB signal to noise ratio. Both the speech and noise signals originally came from Widex (1999).

The compressed stimuli also came from Schmidt (2006). The test signal was compressed in an offline 3-channel compressor model built in MATLAB SIMULINK. The attack time was 10 ms and the release time was 40 ms. There was a fixed compression ratio over the whole dynamic range of the test signal.

The stimuli were presented diotically to the test participants via Sennheiser HD515 headphones. The output level of the computer sound card was adjusted so that the level of the original test signal was approximately 75 dB SPL at the entrance to the ear canal.

Procedure

The test participants subjectively assessed the stimuli using a paired-comparison procedure based on the criteria of "clearest speech". The stimulus presentation was random and was controlled using a Graphical User Interface written in MATLAB by Wooken Song. The participants could only listen to each pair of stimuli once before making their choice. Since there were five stimuli and each stimulus was compared with each other only once, this means each participant made 10 paired comparisons.

2.1.3 Results

Preference Frequencies

Table 2.1 shows the cumulative paired comparison matrix for all the participants. This shows how often each stimulus was "preferred" over the other stimuli. For example, for the row CR1.5, it can be seen that none (0) of the participants preferred CR1.5 over either the original signal or CR1. Further along the row, it can be seen that CR1.5 was preferred over CR2 by 5 participants, and CR1.5 was preferred over CR3 by 12 participants.

Table 2.1: Cumulative paired comparison matrix for the participants. The cell entries denote the number of participants out of 12 who judged the stimuli listed in the row as clearer than the stimuli in the column.

	Original	CR1	CR1.5	CR2	CR3
Original	-	2	12	11	12
CR1	10	-	12	11	11
CR1.5	0	0	-	5	12
CR2	1	1	7	-	12
CR3	0	1	0	0	-

Test transitivity

From table 2.1, the preference order could be determined as:

$$\mathrm{CR1} > \mathrm{Original} > \mathrm{CR2} > \mathrm{CR1.5} > \mathrm{CR3} \tag{2.3}$$

Transitivity can be checked by checking the preference order for each pair of stimuli in table 2.1. For example, CR1.5 was preferred less often than CR1, Original and CR2 but more often than CR3. Since the preference order always held, regardless of which pairs of stimuli are considered, it can be said that there were no transitivity violations.

Testing The Likelihood Ratio and Deriving Scale Values

The paired comparison data were fitted using the BTL model using the `OptiPt` function. A likelihood ratio test for the fit of the BTL model indicated that the model did not fit well ($\chi^2(6) = 11$, $p < 0.05$) because a significant p-value indicates that the fitted derived scale data deviate significantly from the direct magnitude estimates. Closer examination of row CR3 in table 2.1 showed that CR3 was only selected in 1 out of a total of 48 comparisons. Since the BTL model is based on ratios of how often one stimulus is selected over the other stimuli, the model equation 2.1 cannot be solved when too many of the inputs (preference frequency, p_{ab}) are 0.

The modeling process was repeated by leaving the responses to CR3 out. This time, the likelihood ratio test indicated the model fitted adequately ($\chi^2(3) = 2.3$, $0.05 < p < 0.25$).

The derived scale values from the paired comparison data are shown in figure 2.2. The convention in the literature is to plot derived scale values on a logarithmic axis. (see Choisel and Wickelmaier, 2007, for example). The largest scale value was for CR1, followed by Original, then CR2 and the lowest scale value was for CR1.5. Hence the scale values follow the test order shown in equation 2.3. It is also possible to see that there is a large difference between CR1, and CR2, indicating a large perceptual distance between these stimuli.

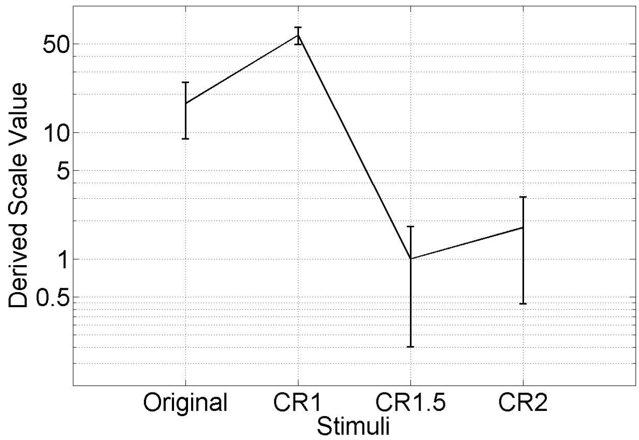


Figure 2.2: Derived scale values for the test stimuli derived from the paired comparison data using the BTL model implemented in the `OptiPt` function. The error bars represent the standard deviations.

2.1.4 Discussion

Results showed that using the paired comparison procedure, an order of preference could be established for the hearing aid processed stimuli. This order showed an absence of transitivity violations and this suggests that the participants use the same listening criteria, regardless of which pair of stimuli are compared.

An initial attempt to derive scale values was not successful because one stimulus, CR3, was so different from the other stimuli that it was only selected in a total 1 out of 48 paired comparisons. The BTL model is based on calculation of ratios of how often one stimulus is selected over another stimulus. Hence, the BTL model cannot be fitted adequately when one stimulus is seldom selected over another stimulus/stimuli. Once CR3 was dropped, then a meaningful scale could be derived using the BTL model.

In summary, the advantages of using the BTL model to derive scale values include that (i) assumptions such as transitivity and uni-dimensionality can be tested; and (ii) it is possible to say something meaningful about the perceptual difference between stimuli (e.g., meaningful intervals). The main disadvantage is that the stimuli must not be obviously different from each other or the model equation cannot be solved. Finally, the method is very time consuming because paired comparisons take more time than evaluating one stimuli at a time and many participants (or repeats) are needed to fit the model sensibly.

The BTL model shows promise as a statistical technique to analyse paired comparison data. For the rest of the Ph.D., the decision to use the model was taken on a case-by-case basis. That is, the BTL model was only used in experiments when the criteria of transitivity could be held and the data could be fit to the model adequately.

2.2 Pilot II. Influence of listener task on ratings of pleasantness for everyday sounds [1]

Abstract

The objective of the current experiment was to develop a method to investigate the influence of subject task on the evaluation of sound stimuli for use in future hearing aid experiments. Twenty listeners with normal hearing rated real-life sound stimuli under different conditions. The sound stimuli were binaurally-recorded soundscapes with low-level target sounds mixed in. The conditions were:

1. Listening only to sound stimuli without any other tasks. This condition is similar to the method used in typical hearing aid studies.
2. An 'auditory detection' paradigm, where listeners detect low-level target sounds (e.g. a microwave beep) within the sound stimuli.
3. The 'irrelevant sound' paradigm, where listeners perform cognitive tasks (e.g. simple addition of numbers), while the sound stimuli are presented.

After listening to each sound stimulus under these three conditions, listeners rated the pleasantness of the sound stimulus. The finding was that ratings of auditory pleasantness were lower under the irrelevant sound condition and under the auditory detection condition than in the listening only condition. However, there was a large degree of variability associated with the ratings, which reduces the sensitivity of the method for use of evaluating hearing aid settings.

[1] Based on Connor and Poulsen (2007)

2.2.1 Introduction

This experiment is part of a Ph.D. project that aims to investigate hearing aid wearers' preference for the audibility of soft sounds. In order to assess preference for hearing aid settings, a suitable method must first be found that can provide the results of interest. Typically in hearing aid laboratory studies, listeners evaluate hearing aid settings by passively listening to sound stimuli and then assessing the hearing aid settings based on their perception of the stimuli. However in real life, there may be many signals competing for the listeners' attention and "listeners must locate, identify, attend to and switch attention between signals" (Gatehouse and Noble, 2004, p. 86). Also in real life, listeners may find auditory information from the environment to be distracting because it directs attention away from another task (e.g. reading). This attentional aspect of hearing is not present in current typical hearing aid methodologies. It is a general aim of this experiment to create a research paradigm that combines the attentional complexity of the real world with the experimental control of the laboratory. In order to do this, the subjects are given tasks, where sound stimuli are relevant or irrelevant, in order to direct the subjects' attention either to or from the sound stimuli. In the experiment, normal-hearing listeners heard binaurally-recorded real-life sound stimuli under three conditions:

1. Listening only to the sound stimuli in a manner similar to typical hearing aid studies. This condition acts as the reference condition.

2. An auditory detection paradigm, where listeners detect low-level target sounds (e.g. a microwave beep) within the sound stimuli. This is similar to situations in real-life, in which listeners must listen to the auditory environment in anticipation of an auditory event.

3. The 'irrelevant sound effect' paradigm, in which listeners perform visual cognitive tasks (e.g. simple addition of numbers), while sound stimuli are presented. This is similar to a situation in real-life where listeners are engaged in a task, and sound is not relevant to the task. Extraneous sound has been consistently shown to impact performance on cognitive tasks, particularly short term memory tasks (see Beaman, 2005, for a recent review).

After listening to each sound stimulus under these three conditions, listeners rated the pleasantness based on their perception of the sound stimuli. The ratings in each of the three conditions are then compared in the analysis. Pleasantness was chosen as the listening criterion because it is a sound attribute that is easy for test subjects to understand and has been used as a listening criterion in a number of previous hearing aid studies. Additionally, it has been demonstrated that 'auditory unpleasantness' can be judged consistently over a wide range of stimuli (Ellermeier et al., 2004).

There are two success criteria to continue using this method in future hearing aid experiments:

1. It should be demonstrated that ratings of auditory pleasantness depend on the listening condition.

2. The variability between conditions should be sufficiently low that the method can be used to detect perceptual differences between hearing aid processed sound stimuli.

2.2.2 Method

Subjects

Twenty listeners with normal or near-normal hearing were used as subjects. Nineteen of the twenty subjects were either students or employees of the Technical University of Denmark (DTU). Five of the twenty subjects were female. The age of the subjects ranged from 21 to 44 years (mean=30 years).

Materials

Equipment: The experiment was performed in a sound proof booth at the Centre for Applied Hearing Research (CAHR), DTU. The presentation of the tasks and sound stimuli was controlled via MATLAB on a stationary computer. The computer sat outside the listening booth to minimise extraneous noise while the screen, keyboard and mouse were inside the booth. The screen was a 17 inch LCD screen. The sounds were presented via a good quality soundcard and HD580 precision circumaural headphones.

Sound Stimuli: Each sound stimulus consisted of a background soundscape and a target sound (table 2.2). The soundscapes were all from the ICRA natural sound recordings (Bjerg and Larsen, 2006) and were recorded using a Head and Torso simulator. The target sounds are taken from the Digiffects CD sound effects library (Ljudproduction AB, 2007) and were mixed in at levels determined in a previous pilot experiment to give an average 70% detection rate. The levels (dB sound pressure level[SPL]) were calibrated using an ear simulator (IEC 60318-1, 1998, Brüel & Kjær Type 4153). The loudness and fluctuation values were reported in Bjerg and Larsen (2006).

Table 2.2: Overview the sound stimuli: soundscapes and corresponding target sounds. The values are transcribed from those reported in Bjerg and Larsen (2006). Loudness refers to mean non-stationary loudness and 1 sone corresponds to a 40dB reference signal at 1 kHz. SPL is an abbreviation of sound pressure level and was analysed as an RMS level at the recording microphone. One vacil corresponds to a 100% amplitude modulated 60 dB pure tone at 1 kHz.

Soundscape	Loudness (sone)	SPL (dB SPL)	Fluctuation (vacil)	Target Sound	SNR (dB)
Dishwasher	12.9	72	0.55	Glass breaking	-27
Supermarket	13.5	61	1.19	Baby cry	-21
Kitchen	28.7	71	0.62	Microwave beep	-25
Pneumatic drill	43.5	78	1.57	Whistle	-22
Traffic, high	47.4	78	0.88	Car horn	-25

The cognitive tasks The tasks used in the 'irrelevant sound effect paradigm' were taken from the Walter Reed Performance Battery described in (Thorne et al., 1985) and coded into MATLAB using the Psychophysics Toolbox extension version 2.54 (Brainard, 1997; Pelli, 1997). The Walter Reed battery was selected because it is designed to compare intra-subject differences across test conditions and the tests are short and do not require any prior knowledge or training.

Procedure

Prior to testing, each subject listened to and rated each sound stimulus for 45 seconds to become familiar with the stimuli and the pleasantness scale. For the actual testing, subjects listened to each sound stimulus in randomised order for one minute each under the following listening conditions, ordered in a counter-balanced latin square design.

1. Listening only to the stimuli in a manner similar to typical hearing aid studies (figure 2.3).

2. An auditory detection paradigm, where listeners detect a target sound (e.g. a microwave beep) within the sound stimulus. See figure 2.4. The target sound appears five times at randomised intervals within the one minute of listening. MATLAB registered the hit and miss rates. Prior to testing in this condition, subjects were given one training round with an easy example of a dog barking in a forest.

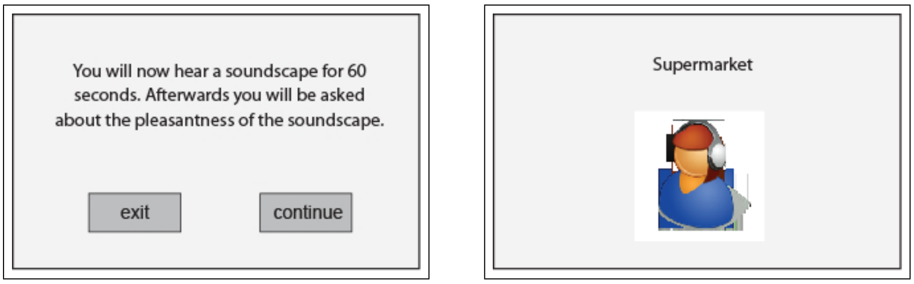
3. The 'Irrelevant Sound' paradigm, where listeners perform cognitive tasks in the presence of the sound stimuli. Subjects had one training round with the cognitive tasks prior to testing. Three cognitive tasks were performed in a counter balanced order: missing letter tasks, missing picture and two column addition. The tasks are described below:

   Missing letter Nine randomised letters appear in a row for 3.3 seconds. After a 1.7 second retention interval, eight of the nine letters are re-displayed in a different random order and the subject enters the missing letter. See figure 2.5. For each sound stimulus, each subject did five missing letter tasks.

   Missing picture Five random pictures of everyday items (e.g. animals, cars, trains) appear in a row for 0.8 seconds. After a 1.7 second retention period, one of the five pictures is re-displayed and the subject should indicate if the picture was one of the original five ('Y' or 'N'). For each sound stimulus, each subject did five missing picture tasks.

   Two column addition Two double digit numbers appear in a row. The subject task was to add them together as fast as possible. For each sound stimulus, each subject did 10 addition equations.

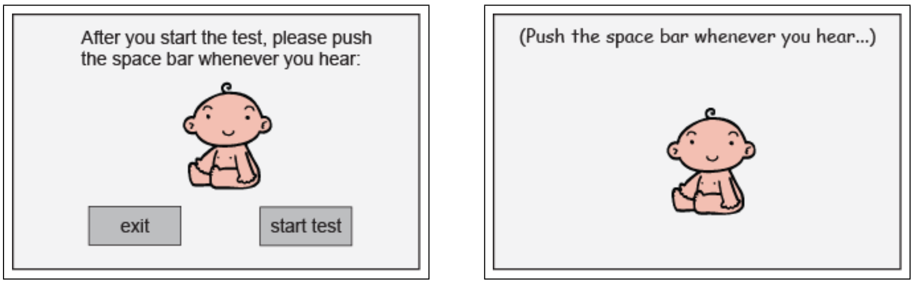
Figure 2.3: Screens shown to test subjects during the listening only condition.



(a) Instruction Screen.

(b) Screen while listening.

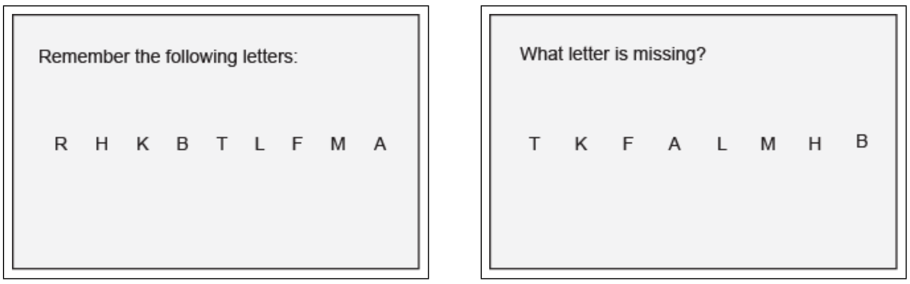
Figure 2.4: Screens shown to subjects during the 'auditory detection' test condition.



(a) Instruction Screen. The target is presented both pictorially and as a sound over the headphones.

(b) Screen while listening

Figure 2.5: Screens shown to subjects for the missing letter task during the 'irrelevant sound' condition.



(a) Nine random letters appear for 3.3 seconds.

(b) Eight of the nine letters reappear and subject should type which letter is missing.

After listening to each sound stimulus under the conditions listed above, listeners rated the pleasantness of the sound using the scale shown in figure 2.6. To avoid ceiling and floor effects, the ends of the scale are not fixed.

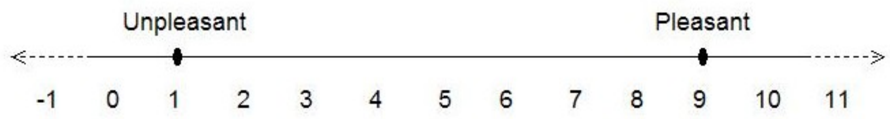


Figure 2.6: Pleasantness scale used for rating the sound stimuli.

2.2.3 Results

Effect of listening condition on ratings

Figure 2.7 shows the influence of the listening condition on average pleasantness ratings. A mixed model analysis of variance was performed using the Statistical Package for the Social Sciences (SPSS). The fixed effects were stimuli and listening condition and subjects were the repeated random effects. There was a significant sound stimuli effect ($p < 0.001$) and a significant condition effect ($p < 0.05$). The condition effect reflects that ratings of auditory pleasantness worsen, while either monitoring for a target sound or while performing a cognitive task, where sound is irrelevant. There was no significant interaction between sound stimuli and condition. Post-hoc pairwise analysis using a Bonferroni adjustment showed a significant difference between the ratings in the listening only condition and the irrelevant sound condition.

Variability between listening conditions

In order that this method can be used in future hearing aid experiments, the variability should be low enough to detect perceptual differences between hearing aid settings. This was assessed using a statistical power analysis to estimate the number of subjects required in a future hearing aid experiment. Firstly, the difference in ratings between conditions was calculated for each subject and each sound stimulus. The overall stan-

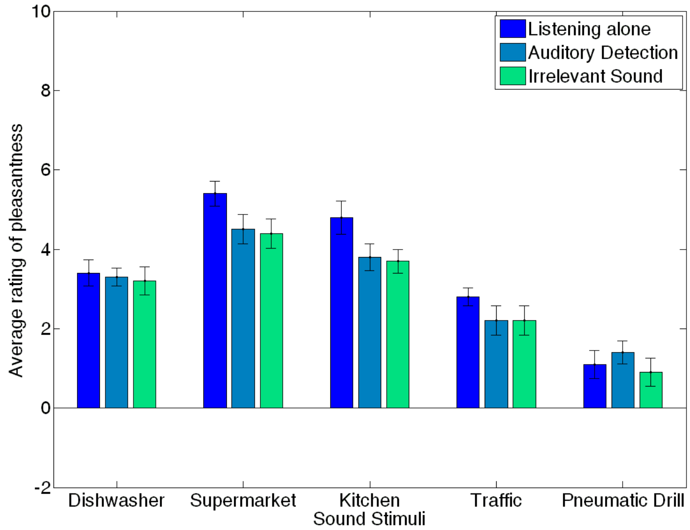


Figure 2.7: Average ratings of auditory pleasantness for the five sound stimuli under the three listening conditions. The error bars represent the 95% confidence intervals.

dard deviation for all intra-subject differences between conditions was 1.45. Secondly, a power analysis was performed in Statistical Analysis Software, SAS. It is assumed that the differences that we want to detect between hearing aid settings are as low as 0.5 on the pleasantness rating scale. The power analysis based on a paired t-test indicated that 68 test subjects would be required to detect a difference of 0.5 on the rating scale (estimated using $\alpha = 0.05$, $\beta = 0.2$ and standard deviation, $SD = 1.45$). If a less sensitive test can be accepted the number of required test subjects decrease accordingly (e.g. 19 test subjects for a difference of 1.0).

Performance in the ‘irrelevant sound’ condition

Figure 2.8 shows the effect of the sound stimuli on the missing letter task scores. The scores are percentage correct out of a total of five questions and therefore is binomially distributed and can not be treated using parametric statistics. Thus, the effect of the sound stimuli on missing letter performance was analysed using a Friedman test in Statistical Package for the Social Sciences, SPSS. The Friedman test is a non-parametric test that is similar to the parametric repeating-measure analysis of variance. The result showed a significant difference between missing letter task performance in the presence of the five different sound stimuli ($\chi^2(13, N = 19), p < 0.05$). In contrast, the sound stimuli did show any effect the performance on the missing picture task or the two column addition and they will not be further discussed.

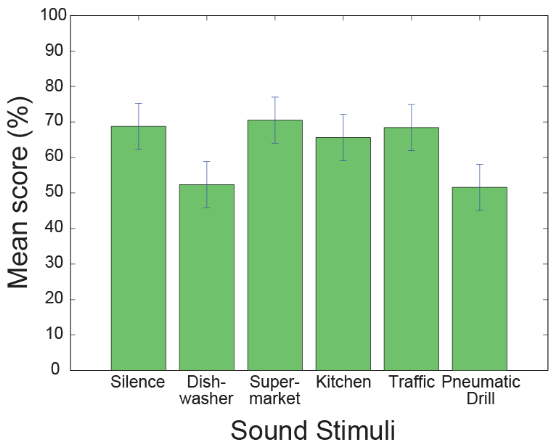


Figure 2.8: Average performance for the missing letter task for the five sound stimuli and one silent stimulus. Error bars represents the 95% confidence intervals.

2.2.4 Discussion

The objective of this experiment is to develop a method to investigate the influence of the task on the evaluation of hearing aid settings. The method is assessed using the following two criteria.

Criterion 1: Effect of listening condition on ratings of auditory pleasantness

Ratings of pleasantness on average decreased when listeners either had to detect a specific target sound or perform cognitive tasks, where the sound was irrelevant. Thus it appears that the listener task does influence perception of auditory pleasantness. This finding is consistent with audiologists' anecdotal reports that when hearing aid wearers are in attentionally-complex real-world situations their hearing aids "sound worse" than in the clinic.

Criterion 2: Inter-subject variability

As indicated in the results, 68 subjects would be required in order to detect a perceptual difference of 0.5 on the pleasantness rating scale. Perceptual differences between hearing aid settings of this small magnitude have been observed in other hearing aid studies (e.g., Neuman et al., 1998). With hearing-impaired subjects, it might be possible that the test-retest variance will be even larger and hence the required number of subjects even larger. Paired comparisons may be a more appropriate method to assess hearing aid settings rather than ratings because paired comparisons are more sensitive to small differences between stimuli (Eisenberg et al., 1997), but it would be too complicated for the test subjects to combine paired comparisons with additional tasks, like cognitive tasks.

Other remarks

It was interesting to observe that performance on the missing letter task significantly decreased for two of the sound stimuli (the industrial dishwasher and the pneumatic drill) but not for the other sound stimuli (supermarket, kitchen and traffic). There is no obvious explanation as to why the dishwasher and drill are the most disturbing because the dishwasher was a quiet stimuli with low fluctuation and the opposite is true about the drill. One possible explanation is that the industrial dishwasher and the pneumatic drill were the least familiar stimuli. However, at least for speech stimuli, the degree of disturbance is only slightly altered by whether the language is familiar or unfamiliar(Jones and Macken, 1995). Another possible explanation is that the dishwasher and the pneumatic drill recordings each had one dominant sound source, which sometimes turns on and off. Some studies have indicated that the amount of disturbance relates to the degree of change in one or more auditory streams, where one changing stream is more disturbing than three steady streams (Jones and Macken, 1995). A previous study has shown that the degree of disturbance can be reduced using sound processing, such as low pass filtering (Jones et al., 2000). It could be an interesting piece of future research to investigate how hearing aid processing influences performance on cognitive tasks.

Conclusion

The proposed method showed that ratings of pleasantness for non-processed real-life sound stimuli decreased when subjects were engaged in additional tasks. However, the ratings showed considerable inter-subject variability, which reduces the usefulness of the method to investigate perceptual differences between hearing aid settings. The performance on the missing letter task was impaired by some of the sound stimuli but not others, which poses an interesting question for future research.

Acknowledgments

A special thanks goes to Andrew Bell and Alice Lhomond who provided an earlier version the MATLAB code for the cognitive tasks used in the irrelevant sound condition. Thanks to Professor Kathy Pichora-Fuller for helping discussions during the planning stage of this experiment. Thanks also to the twenty subjects who participated.

2.3 Pilot III. Influence of the instruction and signal on the preference for compression setting in normal-hearing listeners

2.3.1 Introduction

The usual hearing aid evaluation methodologies do not include the attentional complexity of real-life listening situations. The previous experiment (Connor and Poulsen, 2007) attempted to do this by directing the participants' attention using tasks. Here normal-hearing participants rated the "auditory pleasantness" of real-life environmental sound stimuli under three different listening conditions. In condition one, the participant rated the sound stimulus without performing additional tasks. In condition two, the participant's attention was directed toward the sound stimulus by instructing them to push a button every time they detected a target sound within the background noise. In condition three, the participants' performed visual cognitive tasks in order to direct their attention away from the sound stimulus. The results showed that the task of the participant did influence the ratings of the sounds. However the variability of the ratings was large and indicated that this method would not be sensitive to detection of small perceptual differences between hearing aid processed stimuli, unless dozens of test participants were used. Anecdotally, some participants reported that it was difficult to give a rating for the sound stimuli when they were occupied with other tasks.

Another hearing aid evaluation methodology that is more sensitive and less variable than categorical ratings are paired comparisons (Eisenberg et al., 1997; Purdy and Pavlovic, 1992). However, for paired comparisons the participants' attention need to be on the stimuli, and because of this, it would be difficult to combine paired comparisons with additional tasks, such as those used in Connor and Poulsen (2007). It may instead be more feasible to direct the participants' focus of attention using instructions rather than tasks. For example, by instructing them to "listen out" for a particular target sound within the stimuli (e.g., the sound of another person's footsteps) or by instructing them to imagine that they are concentrating on something other than the sound (e.g., reading a newspaper).

Hearing aid settings are usually evaluated using test signals that are from real-life everyday environments (e.g., living room, supermarket etc). Previous experimental experience shows that test participants have difficulty hearing a difference between hearing aid settings for some test signals than for other signals (Neuman et al., 1998). It is difficult to know in advance if test participants can hear a difference between hearing aid settings for a given set of test signals before they are tried out.

The aim of the current experiment is to investigate the feasibility of using instructions to direct the participant's attention while the participants make paired comparisons of hearing aid processed stimuli. A secondary aim of the current experiment was to investigate if participant's could hear a difference between compression settings for the real-life test signals that were recorded for the current experiment.

2.3.2 Method

Test participants

The participants were ten students or members of staff from the Department of Electrical Engineering, the Technical University of Denmark (DTU). All participants had audiograms measured within 12 months prior to participating in the experiment and all had normal or near-normal hearing (i.e. hearing thresholds 30 dB HL or better). The age range of the participants was 24 to 63 years of age.

Test signals

The test signals were recorded in real-life everyday settings using bilateral behind the ear (BTE) microphones mounted in BTE hearing aid cases and placed on the ears of a volunteer who could walk freely around in the recording locations. The recording microphones were connected to a two-channel hard-disk recorder (Sound Devices 722 digital audio recorder) and the signals were recorded at a sampling frequency of 44.1 kHz in a 24-bit format. The recording equipment was calibrated using a broadband noise (0.1 - 10 kHz) at presentation levels 50-100 dB SPL in 10 dB increments presented in an Interacoustics TBS25 hearing aid test chamber.

The signals used in the experiment were recorded in three locations: in a living room, at an underground suburban train station (Nørreport S-tog station) and in a supermarket. From each recording, approximately eight second long extracts were selected. Half of these extracts had speech present and the other half had no speech. The signals all include at least one recognisable sound event, as described in table 2.3. The signal levels are also shown in table 2.3, as measured between 100-10,000 Hz with a 125-ms analysis window.

The signals were compressed offline using a 15-channel MATLAB SIMULINK model which is described in more detail in chapter 3. In the current experiment, two sets of compression thresholds (CTs) were compared:

CtMod CT in each channel was equal to the 1/3 octave level of normal speech (ANSI S3.5, 1997) in that channel; and

CtLow CT in each channel is 30 dB lower than CTMOD.

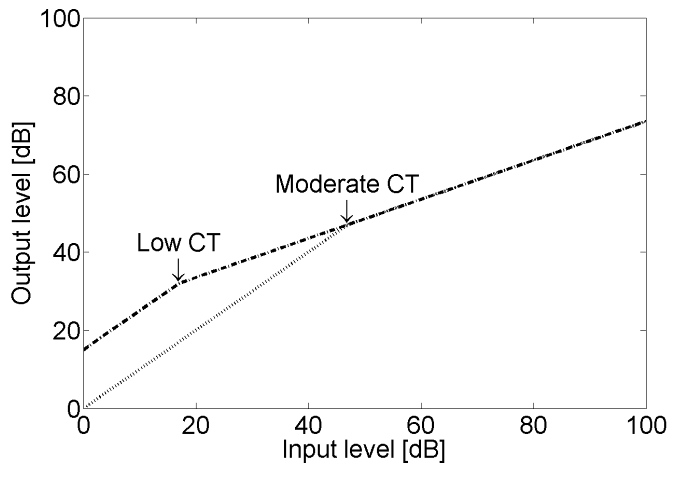


Figure 2.9: Schematic illustration of the input-output curve of the hearing aid compressor model used in the current study. CTMOD was set to the RMS levels of normal speech (ANSI S3.5, 1997) in each channel. CTLOW was set 30 dB below the CTMOD threshold for each channel. Gain for medium and high levels was the same for CTMOD and CTLOW.

The input-output curve of the two CTs settings is shown in figure 2.9. A fixed

Table 2.3: Description of the signals used in the experiment and the recording levels in dB SPL. RMS stands for root-mean-square level, and L_{10} and L_{90} for the 10th and 90th percentile levels, respectively.

Recording location	Target signal	Speech present?	Left side			Right side		
			RMS	L_{90}	L_{10}	RMS	L_{90}	L_{10}
Supermarket	Sales assistant	Yes	66	71	58	67	70	59
	Other customers	Yes	65	69	57	66	69	57
	Credit card beeps	No	63	67	56	63	67	56
	Packing groceries	No	64	67	60	65	67	60
Living room	Guest talking	Yes	71	76	56	70	75	55
	News radio	Yes	58	62	44	62	66	55
	Car outside	No	63	66	55	63	67	55
	Footsteps	No	62	66	54	57	62	44
Underground train station	Buying a train ticket	Yes	70	74	67	70	74	67
	Teenagers talking nearby	Yes	74	77	72	77	79	76
	Zipper opening	No	72	74	70	72	74	70
	Train departs	No	72	73	69	76	78	73

2:1 compression ratio was used in each channel. The attack time was fixed at 10 ms and the release time at 40 ms.

The compressed signals were presented to the participants via Sennheiser HD580 headphones. Calibration was done using an ear simulator (IEC 60318-1, 1998, B&K 4153 with flat plate).

2.3.3 Paired-Comparison Procedure

Participants subjectively assessed the signals using a paired-comparison procedure performed via a Graphical User Interface (GUI) written in MATLAB. For each of the 12 signals, the two CT settings were compared four times within a run. Within each of these runs, the order of the CT presentation was randomised. For each of the 12 signals (e.g. supermarket etc), each run was performed twice for each of the two instruction sets (a total of four runs for each test signal per participant). The two instruction sets were:

- Try to listen for a particular speaker or particular sound event (e.g. what the sales assistant at the supermarket is saying). Which setting do you prefer?
- Try to imagine that you are busy doing something other than listening out for the sound (e.g. reading). Which setting do you prefer?

The total number of comparisons was 192 (4 comparisons $\times$ 12 signals $\times$ 2 instructions $\times$ 2 repeats) and these were obtained over at least two experimental sessions. The order of the signals and instructions was counterbalanced.

The GUI allowed the participants to listen to each signal as often as they liked before deciding their preference. They could not start playing the other signal until the current signal had played until the end. Neither could they decide their preference until they had heard both signals in their entirety.

2.3.4 Results

Figure 2.10 shows the preference proportion for CTLOW i.e., how often CTLOW was chosen over CTMOD for each of the twelve test signals and the two instruction sets.

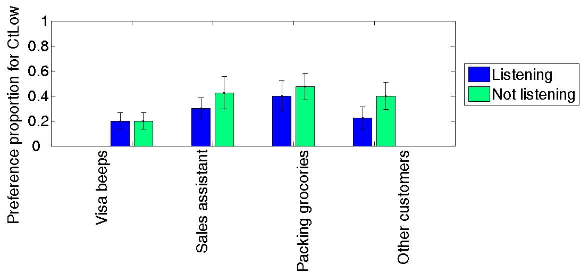
There seems to be a trend that for 8 of the 12 signals used, the proportion of times CTLOW is chosen increases when they are instructed to imagine that they concentrate on something other than the sound.

A multi-variate analysis was performed using the binomial family of the Generalized Linear Model using the R Statistical Environment (R Development Core Team, 2008). The dependent factor was the CTLOW preference proportion for each subject during each experimental run. The explanatory factors considered were Signal and Instruction and the interaction between the two. Signal was a significant factor at the 95% confidence level ($p < 0.05$). Instruction, and the interaction between Signal and Instruction were not significant factors.

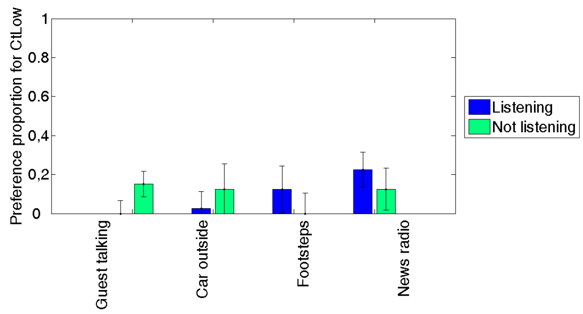
2.3.5 Discussion

The results indicated that there seems to be a slight instruction effect, although the effect was not significant. The lack of significance could potentially be due to a lack of statistical power due to too large a standard error in relation to the size of the effect. In addition, an instruction effect would be consistent with previous hearing aid studies (Neuman et al., 1998; Hansen, 2002; Keidser et al., 2005; Gatehouse et al., 2006). For instance, Keidser et al. (2005) found in a laboratory listening experiment that test participants preferred a compression setting that gave less gain when instructed to listen for listening comfort rather than for speech understanding.

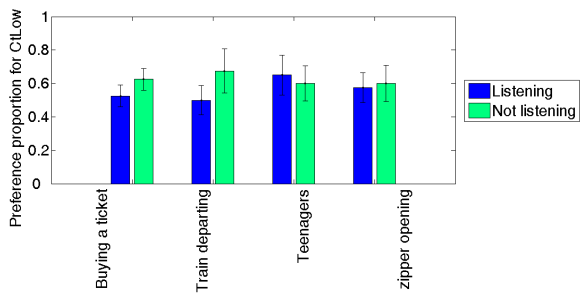
There is a significant signal effect. For the supermarket signals, the participants preferred CTLOW in about one third of the comparisons. This was similar to results to previous studies by Barker and Dillon (1999) and Dillon et al. (1998) who also found that about one-third of participants prefer a low CT, when combined with fast-acting compression. For the living room signals, the participants preferred CTLOW in about one tenth of the comparisons. Here many participants reported that they were purely making their paired comparison decision based on the level of the microphone noise, which was audibly louder in the CTLOW setting. Finally, for the underground station signals, many participants complained that they could not hear a difference between the CTLOW and CTMOD settings and this is reflected in the results because the preference for each setting is about fifty-fifty. Examination of the levels in table 2.3 shows that the input levels, also for the weak components of the signal (L_{10})



(a) Supermarket



(b) Living room



(c) Underground station

Figure 2.10: The mean proportion of times that CtLow was selected over CtMod for each of the twelve test signals and the two instruction sets. The figure legend "listen" refers to the instruction to listen for a target sound and the figure legend "not listening" refers to the instruction to imagine concentrating on something else other than the sound. The error bars indicate the standard error of the mean.

exceeded CTMOD. This means the compression was activated to the same degree for both CTLOW and CTLOW.

In summary, the primary aim of the current experiment was to investigate the feasibility of using instructions to direct the participant's attention while the participants make paired comparisons of hearing aid processed stimuli (chapter 3). Since instruction appeared to have a slight effect on preferred CT, although the effect was not significant, it was decided to continue to use the two sets of instructions. The secondary aim of the current experiment was to investigate if participant's could hear a difference between compression settings for the real-life test signals that were recorded for the current experiment. Based on the results, it was decided to (i) continue with the supermarket signals; (ii) re-record the living room signals in another apartment with a higher noise floor to avoid microphone noise; and (iii) replace the underground train station signals with signals recorded on a pedestrian mall, which has a lower root-mean-square (RMS) level and more variation in the level distribution.

3

Laboratory investigation into preferred hearing aid compression thresholds: Effects of release time and listener instructions

Abstract

There has been little direct investigation into hearing aid users' preference for gain at low input levels. One of the most important parameters for controlling gain at low input levels is the compression threshold. The current study investigated if the preferred compression threshold is influenced by (a) the compression release time, and/or (b) the instruction to the participant. Real-life everyday signals were recorded and then processed offline in a 15-channel compressor model using a combination of two compression thresholds and three Release Times (RT=40, 400 and 4000 ms). The compression thresholds were set at (i) CTMOD, the level of normal speech in each channel and (ii) CTLOW, 30 dB lower than in (i) in each channel. Both sets of compression thresholds varied across channels to follow the normal speech spectrum. Twelve experienced hearing aid users with moderate, sloping hearing losses made paired comparisons with two instructions: (a) to listen for a particular target sound within the signal, or (b) to imagine they were concentrating on something other than the sound. Preference for compression threshold was strongly influenced by the release time; as release time increased, the extent of preference for the low compression threshold also increased. Instruction did not influence the preference for compression threshold. The current findings suggest that recommendations for compression thresh-

old and, consequently, gain at low input levels should depend on the release time of the hearing aid.

3.1 Introduction

People with sensorineural hearing loss have a loss of audibility for many low level sounds. Concurrent with the loss of audibility, they usually have loudness recruitment, which is an abnormally rapid growth in loudness for a given increase in suprathreshold sound level compared with the normal rate of loudness growth (Fowler, 1936). Sensorineural hearing loss is usually managed audiologically by fitting compression hearing aids which provide more gain at low input levels to compensate for the loss of audibility and less gain at high input levels to compensate for the loudness recruitment. The gain characteristics of hearing aids are defined by gain targets specified by a hearing aid fitting rationale. Many hearing aid fitting rationales were initially developed using linear amplification, for example, National Acoustic Laboratories - Revised (NAL-R Byrne and Dillon, 1986); Desired Sensation Level version 3.1, (DSL v. 3.1 Seewald et al., 1993) and the Cambridge formula, (Moore and Glasberg, 1998). Consequently, the empirical research which supports these rationales was primarily concerned with amplification of moderate input-level signals, such as speech. Less is known about the extent to which hearing aids should amplify sounds with low input levels.

In considering how much amplification to provide at low input levels, one of the most important compression parameters is the compression threshold (CT). When the compression ratio (CR) is kept constant and the gain for medium levels is fixed, lowering the CT increases the gain at low input levels (see the schematic illustration in figure 3.1). So far, only a few studies have directly investigated the preferred CT in hearing aids. A series of field trials were carried out at the National Acoustic Laboratories (NAL) in Australia using participants with mild-moderate hearing losses (Barker and Dillon, 1999; Barker et al., 2001; Dillon et al., 1998). The studies used fast-acting, single-channel compression hearing aids and found approximately two-thirds of the participants preferred a relatively high CT ($\sim$ 65 dB sound pressure level [SPL]) over a lower CT ($\sim$ 40-57 dB SPL). These studies have been influential and have informed developers of generic hearing aid rationales, such as the National Acoustic Laboratories - Non-Linear 1 (NAL-NL1) rationale (Byrne et al., 2001) and the Desired Sensation Level - multistage Input-Output (DSL *m*I/O) rationale (Scollie et al., 2005).

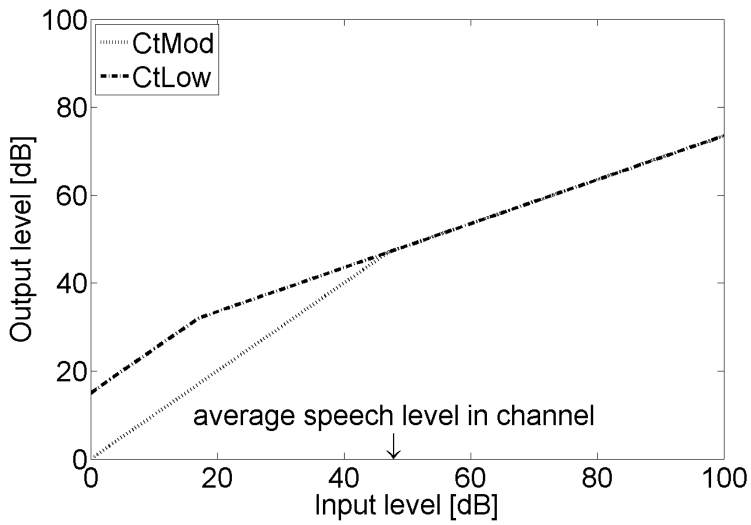


Figure 3.1: Schematic illustration of the input-output curve of the hearing aid compressor model used in the current study. The moderate CT (CtMod) were set to the root-mean square (RMS) levels of normal speech (ANSI S3.5, 1997) in each channel. The low CT (CtLow) were set 30 dB below the CtMod threshold for each channel. Gain for medium- and high-input levels was the same for CtMod and CtLow, whereas at low-input levels CtLow provides more gain than CtMod.

In spite of the broad acceptance of these findings from NAL, the results are not necessarily applicable to other types of hearing aids with either multi-channel compression and/or slower time constants. In a single-channel hearing aid, steady background noises may appear to be modulated by the dominate sounds in the listening environment, such as speech (Laurence et al., 1983). This is because when the signal level exceeds the CT and the compressor is activated, the applied gain follows the slow-modulations in the dominate signal (Stone and Moore, 1992) but these gain changes are then applied across the whole signal. In some cases, this regulation of the gain can be audible to the hearing aid user (HA user) and can be disturbing (Laurence et al., 1983). Since single-channel compression is seldom implemented commercially in single-channel hearing aids any more, it was important in the current study that compression was implemented in a multi-channel compressor to reflect the dominant current technology.

The first prediction of this study was that when slower time constants are used, HA users will be more likely to accept a low CT. The reasoning is that compression, and in particular fast compression, can introduce many negative side effects to the signal, including reduction in the signal-to-noise ratio for signals at positive SNRs (Naylor and Johannesson, 2009; Souza et al., 2006), envelope depth (Souza et al., 2006) and speech intelligibility for vocoded signals (Stone and Moore, 2008). Also increasing the release time increases the perception of pleasantness and decreased the perception of background noise (Neuman et al., 1998). Given that fast-acting compression can be detrimental to the signal, in a fast compression system the lower the CT, the more time is spent in compression and the worse the negative effects. Since many of these negative effects of compression are ameliorated when the release time is lengthened, this gives rise to the first prediction that when the release time is lengthened, the HA users will be willing to accept a lower CT.

The second prediction of this study was that when listeners evaluate hearing aid settings and they are instructed to listen for a weak target sound (e.g. the voice of a shop assistant or the beeps from a register) within the listening environment (e.g., a recording from a supermarket), they will prefer a low CT because it improves audibility for the target sound. The converse prediction is that when listeners are instructed to imagine that they are concentrating on another task (e.g. reading a newspaper), they will prefer a moderate CT because it reduces gain for low input levels and hence reduces the intrusiveness of background noises. It is intuitive to expect that when auditory information is relevant to the listener's current task, they will prefer a hearing aid setting that improves audibility, whereas when auditory information is not relevant to the current task, they will more likely prefer a setting that reduces gain. In support of this, it is well established that when listeners are instructed to understand speech, they generally prefer a hearing aid setting that improves audibility for frequency components which are important for speech (e.g. Keidser et al., 2005). In contrast, when listeners are instructed to judge listening comfort, they generally prefer a hearing aid setting with reduced gain (e.g. Keidser et al., 2005).

The current study investigated if the preferred CT is influenced by (i) the compression release time, and/or (ii) the instructions to the participant. To investigate this, real-life sounds (e.g. supermarket) were recorded bilaterally. These recorded signals were processed offline in a 15-channel compressor model with a constant 2:1 CR and 10 ms attack time. There were two possible CTs (moderate level and low level) and three possible release times (RT=40, 400 and 4000 ms). The compressed stimuli were presented through the Direct Audio Input (DAI) of bilateral Behind The Ear (BTE) hearing aids with insertion gain selected using the NAL-R rationale. Participants with moderate sloping sensorineural hearing losses made paired comparisons of the compression settings via a computer based Graphical User Interface (GUI).

3.2 Method

3.2.1 Test Participants

Twelve HA users with moderate sloping sensorineural hearing loss were tested. Their average age was 71 years (range 64 to 82 years). Participants were recruited from a database of volunteers at the Center for Applied Hearing Research at the Technical University of Denmark. All participants had at least one year of hearing aid experience and had been fitted within the last four years with multi-channel compression hearing aids. Three of the participants were fitted with open-fit hearing aids, five with in-the-ear hearing aids and the remaining four wore BTE hearing aids with earmoulds with standard tubing.

Pure tone audiograms with both air and bone conduction thresholds (figure 3.2) were measured according to ISO 8253-1 (1989). All participants had symmetrical hearing losses with less than 10 dB difference between ears for the pure tone average at 0.5, 1, 2 and 4 kHz and no more than 15 dB difference at any audiometric frequency. All participants had normal middle ear function for both ears, as indicated by both normal type A tympanogram (ASHA, 1990) and the presence of ipsilateral reflexes at

0.5, 1 and 2 kHz. This research was approved by the Copenhagen City Council ethics committee (approval No: KA04159g).

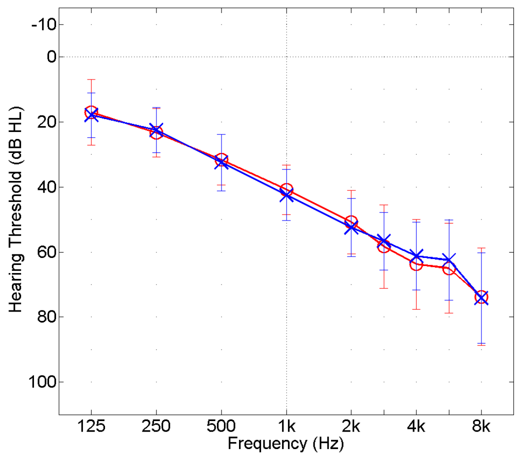


Figure 3.2: Average audiogram for the twelve participants. The circles and crosses indicate the right and left air conduction thresholds, respectively. Error bars indicate ±1 standard deviation.

3.2.2 Test Signals

The signals were recorded in real-life everyday settings using bilateral BTE microphones mounted in BTE cases placed on the ears of a volunteer who could walk freely around in the test locations. The aim was to use test signals that were 1) natural sounding, 2) contained binaural spatial information and 3) had naturally occurring signal to noise ratios. The recording microphones were connected to a two-channel hard-disk recorder (Sound Devices 722 digital audio recorder) and the signals were recorded at a sampling frequency of 44.1 kHz in a 24-bit format. The recording equipment was calibrated using a broadband noise (100-10 000 Hz) presented in a hearing aid test chamber (Interacoustics TBS25).

The signals used in the experiment were recorded in three locations: on the pedestrian mall in downtown Copenhagen (Strøget), in a living room and in a supermarket. From each recording, approximately eight-second long extracts were selected. In half of the selections, speech was present and in the other half, speech was absent. The signals included at least one recognisable target signal, as described in table 3.1. The signal levels are also shown in table 3.1, as measured between 100-10,000 Hz with a 125-ms analysis window. Figure 3.3 shows the spectra of the signals in $1/3$ octave bands. For brevity, only the spectra of the left channel are shown.

3.2.3 Hearing Aid Processing

Gain was applied to the signals in three stages, as indicated schematically in figure 3.4. In the first stage, the signals were compressed offline using a 15-channel MATLAB SIMULINK model. In the second stage, the compressed signals were routed to the hearing aids via the DAI and the gain of the DAI was adjusted to give the same gain as would be applied to an equivalent acoustic input. In the third stage, bilateral BTE hearing aids were adjusted to operate linearly and fitted with insertion gain adjusted to the NAL-R rationale.

Compressor

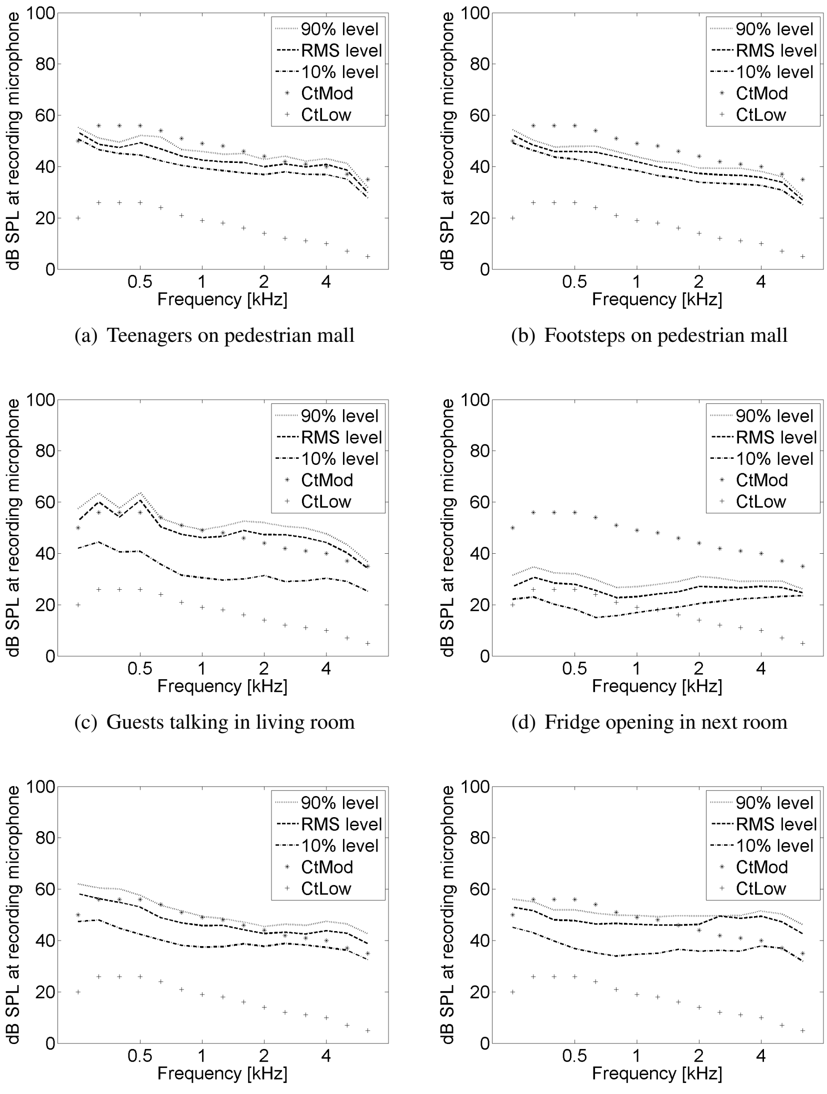
The compressor model consisted of 15 independent channels, each with an approximate $1/3$-octave bandwidth. The attack time of the compressor model was 10 ms and the three release times (RT) were 40, 400 and 4000 ms (IEC 60118-2, 1983). The two sets of CTs were: (i) CTMOD, set so that the CT in each channel was equal to the energy of normal speech within that channel (ANSI S3.5, 1997); and (ii) CTLOW, set so the CT in each channel was 30 dB lower than CTMOD. Figure 3.3 shows how CTMOD and CTLOW varied across channel/frequency.

For both CTMOD and CTLOW, the gain of the compressor was set to 0 dB for input levels equal to the RMS level of speech in each channel (shown schematically in figure 3.1). The same amount of (negative) gain was applied by the CTMOD and CTLOW settings for levels above the RMS level of normal speech. At low input levels, the CTLOW setting applied more gain than the CTMOD setting.

For each combination of RT and CT, the gain was normalised so that each combination provided the same RMS gain a linear setting when the International Speech

Table 3.1: Description of the signals used in the experiment and the recording levels in dB SPL.

Recording location	Target signal	Speech present?	Left side RMS	Left side L_{10}	Left side L_{90}	Right side RMS	Right side L_{10}	Right side L_{90}
Pedestrian mall	Teenagers talking nearby	Yes	57	55	59	60	57	62
Pedestrian mall	Footsteps	No	56	54	58	59	56	60
Living room	Guests talking	Yes	65	52	68	65	53	68
Living room	Fridge opens in next room	No	38	33	43	40	35	43
Supermarket	Two customers talking	Yes	62	55	66	62	54	66
Supermarket	Transaction beeps	No	60	51	64	58	50	63



(a) Teenagers on pedestrian mall

(b) Footsteps on pedestrian mall

(c) Guests talking in living room

(d) Fridge opening in next room

(e) Two customers talking in supermarket

(f) Transaction beeps at supermarket

Figure 3.3: The spectra of six signals used in the experiment as recorded at the left microphone. The RMS, 10th and 90th percentile levels are plotted in approximate 1/3-octave bands using a 125-ms analysis window. The levels are at the left recording microphone (BTE microphone). The levels of CTMOD and CTLOW are also shown.

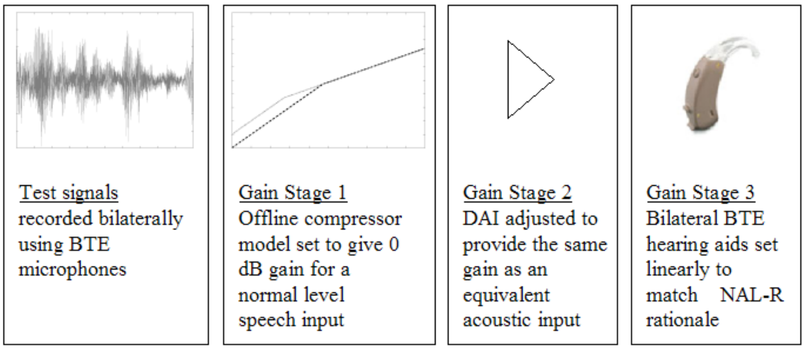


Figure 3.4: Schematic illustration of the gain stages applied to the test signal.

Test Signal (ISTS, Holube et al., 2009) presented at an RMS input level equivalent to speech at a normal vocal effort (62 dB SPL, ANSI S3.5, 1997).

Signal Presentation via the DAI

During the experiment, the left and right channels of the compressed signals were routed from the headphone output of the test computer used for signal presentation to the DAI via a 10 W amplifier with an adjustable attenuator.

Calibration of the gain of the amplifier was performed to ensure that the hearing aid gave the same output for a given electrical input as it would for an equivalent acoustical input. This was performed using a hearing aid set linearly with gain appropriate for the average participant audiogram (figure 3.2). First, the hearing aid output in a 2 cc coupler was measured for a 60 dB SPL white noise (0.125 - 10 kHz) input signal presented acoustically in an Interacoustics Equinox hearing aid test chamber. Second, the same white noise signal was presented as an electrical input via the hearing aid DAI (and amplifier) and the hearing aid coupler output was measured again. Third, the attenuator on the amplifier was adjusted manually so that the output for the electrical DAI input matched the output for the equivalent acoustic input. Figure 3.5 shows that the hearing aid output across frequency was similar for the same white noise signal presented acoustically and electrically via the DAI.

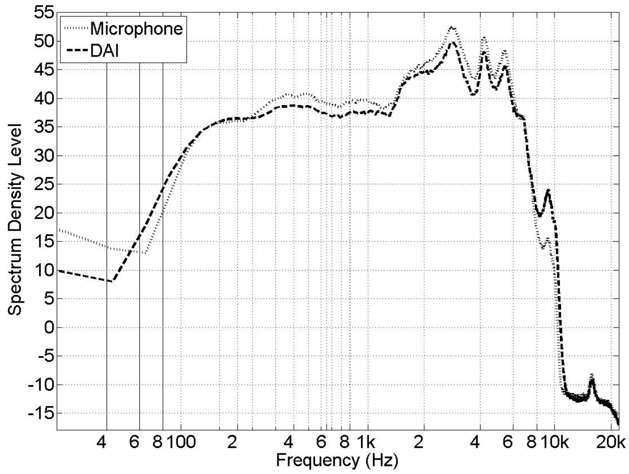


Figure 3.5: Hearing aid output measured in a 2 cc coupler for a 60 dB SPL acoustic white noise (Microphone) and it's electrical equivalent (DAI).

Hearing Aid Fitting

To ensure that the hearing-impaired participants received gain which was appropriate for their hearing losses, participants were fitted with Widex Inteo BTE (IN-9) hearing aids with standard acrylic skeleton earmoulds with 1 mm vents. The hearing aids were set to operate linearly. Additional adaptive features such as directional microphone, noise reduction and feedback canceler were disabled.

The hearing aids were fitted to the NAL-R (Byrne and Dillon, 1986) rationale. The test signals (section 3.2.2) were recorded using BTE microphones, which adds high frequency gain to the signal due to the head related transfer function, the insertion gain targets were corrected by subtracting the average Free Field to Microphone (FFtoMic) transform (Bentler and Pavlovic, 1989).

Real-ear insertion gain (REIG) was measured individually to verify the hearing aid fittings. The signal was a pseudo-random white noise presented at 65 dB SPL from 0° azimuth with a 10 second signal duration. The equipment was the Interacoustics Equinox REM440 module.

3.2.4 Illustration of the Effects of the Compression on a Test Signal

To illustrate the acoustic effects of the compression, the level distribution of a compressed wave file, "guest in living room" was calculated for $1/3$-octave bandwidths using a 125-ms time window to give the RMS, as well as the 90th and 10th percentiles for each band. The levels were then transformed to the estimated levels in the ear canal for a randomly selected test participant, KA for her left ear. Since the internal levels of the compressed wave files were known relative to the calibration of the hearing aid microphone, in order to estimate the levels to sound pressure level (dB SPL) in the ear canal, the Real Ear Aided Gain (REAG)[1] of the participant KA's left ear was added. Finally, to compare the output of the compressor to hearing threshold, the left hearing thresholds were plotted on the same graph, also expressed in dB SPL, using reference equivalent threshold sound pressure level (RETSPL) values given in ISO 389-8 (2004).

3.2.5 Paired-Comparison Procedure

The participants subjectively assessed the signals using a paired-comparison procedure. Prior to starting, the participants received some familiarisation task to give them experience with listening for small differences between stimuli. The training consisted of 17 A/B paired comparisons of real-life environmental signals which had been compressed with varying release times. For each pair of stimuli, the participants were asked to focus on an auditory feature, such as loudness, sharpness, clarity, speech intelligibility or comfort. After listening to each stimuli pair, they were asked to choose whether they preferred stimuli A or stimuli B. Responses to the training paradigm were not recorded.

After the familiarisation task, the data were collected in the following way. For each of the six signals, paired comparisons were performed so that every hearing aid setting was compared with every other hearing aid setting. Since six compressor settings were compared, each run consisted of 15 comparisons. For each of the six

[1] Aided levels in the ear canal relative to a reference microphone placed just below the ear.

signals (e.g. supermarket etc), each run was performed twice for each of the two instruction sets (a total of four runs for each test signal). The two instruction sets were:

- Try to listen for a particular speaker or particular sound event (e.g. crossing signal at lights). Which setting do you prefer?
- Try to imagine that you are busy doing something other than listening out for the sound. Which setting do you prefer?

The total number of comparisons was 360 (15 comparisons $\times$ 6 signals $\times$ 2 instructions $\times$ 2 repeats) and these were obtained over at least two experimental sessions. The participants had breaks at least every half hour. The order of the signals and instructions was counterbalanced.

The signal presentation and response collection were controlled using a Graphical User Interface (GUI) written in MATLAB. The GUI allowed participants to listen to each signal as often as they liked before deciding their preference. They could not start playing the other signal until the current signal had played until the end. Neither could they decide their preference until they had heard both signals in their entirety.

3.3 Results

3.3.1 Real Ear Insertion Gain Measurements

Figure 3.6 below shows the mean achieved minus prescribed gain. It can be seen that the mean REIG was within ± 3 dB of the insertion gain targets between 0.5-4 kHz. Above 4 kHz, the insertion gain rolls off. At 250 Hz the achieved gain was on average 5 dB more than target. This is because at 250 Hz, the targets were often negative and negative gain is very difficult to achieve in practice. Across frequencies, the standard deviation at each frequency was in the range of 4-5 dB.

3.3.2 Statistical Analysis using the Generalized Linear Model

Analysis was performed using the Generalized linear mixed model (GLMM) with a binomial link in the R statistical environment (R Development Core Team, 2008)

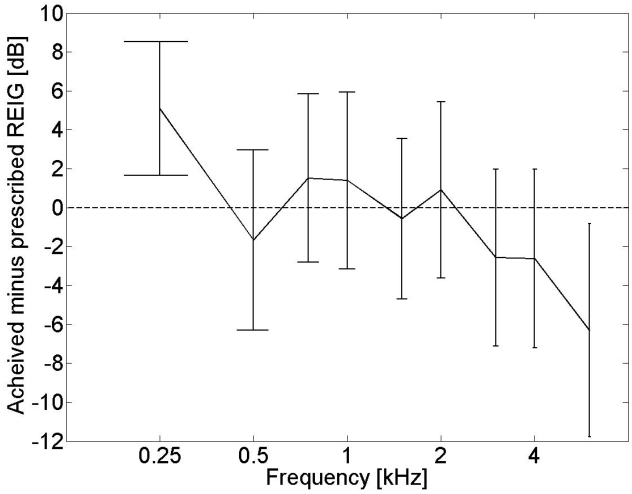


Figure 3.6: Average achieved minus prescribed REIG for a 65 dB SPL pseudo-random noise for the linearly fitted hearing aids. The error bars represent the standard deviation.

using the `lme4` package. The dependent variable were the counts of how often one setting was selected over the other settings. The fixed independent variables were CT, RT, test signal and instruction and the interaction between these variables. The fixed independent variables were analysed as repeating variables. The random independent variable was test subject.

Results of the statistical analysis showed that preference proportion was significantly influenced by CT ($p < 0.01$), RT ($p < 0.01$), test signal ($p < 0.01$) and the interaction between the CT and RT ($p < 0.01$) and the interaction between RT and test signal ($p < 0.01$). Instruction did not have an influence on preference proportion and there was no significant interaction between these instruction and the other variables.

3.3.3 Influence of RT on the Preferred CT

Figure 3.7 shows a boxplot of the preference proportion for each setting, pooled for all 6 signals and both instructions. The notches on the boxes represent the median preference proportion, that is, the proportion of times a compression setting was chosen over all the other settings. For the fastest RT (40 ms), the median preference proportion was 0.46 for CTMOD and 0.09 for CTLOW. In other words, when a fast RT was used, CTMOD was chosen five times more often than CTLOW. When the RT of 400 ms was used, the median preference proportion was 0.49 for CTMOD and 0.30 for CTLOW. When the slowest RT of 4000 ms was used, the median preference proportion was 0.63 for CTMOD and 0.58 for CTLOW, and the extent of preference for the two settings was nearly the same. In summary, the longer the RT, the more likely the participant was to select CTLOW and this observation is supported by the finding of a statistically significant interaction between the RT and CT in the GLMM analysis.

Post-hoc pairwise comparisons were performed comparing preference proportion for the two CTs for each RT. This was performed using Kruskal-Wallis tests with a Bonferroni correction to the alpha level. The analysis showed a significant difference between preference proportions for the two CTs for RTs of 40 and 400 ms ($p < 0.01$), but not for the RT of 4000 ms.

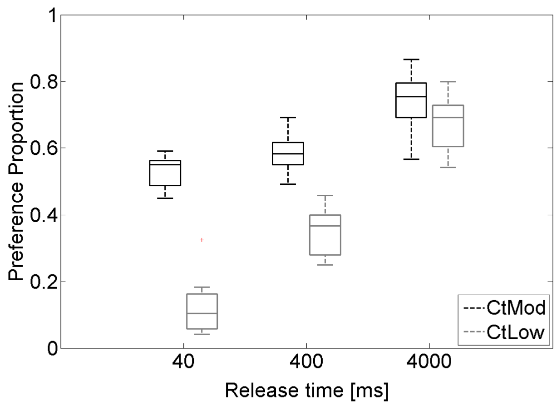
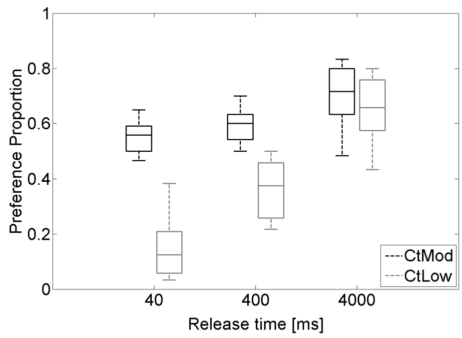


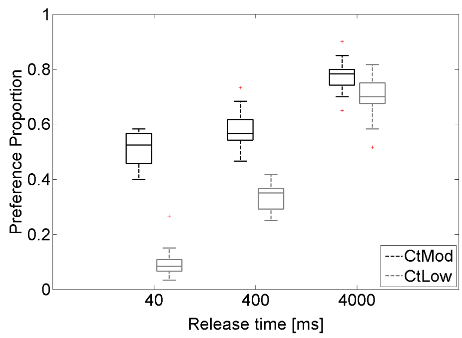
Figure 3.7: Box plot of the preference proportion for each of the 6 compression settings tested. The data are pooled from all participants for all signals, both sessions and both instruction sets. The middle horizontal line in each box shows the median values. The boundaries of the boxes show the 25th and 75th percentiles, the tails the minimum and maximum values, and the crosses show the outliers.

3.3.4 Influence of Instructions on the Preferred CT

Figure 3.8 shows a boxplot of the preference proportion for each setting and each instruction set. The instruction set did not have an obvious influence on how often CtMod was chosen over CtLow, nor did it have any obvious influence on how often any RT was selected. This is supported by the statistical analysis using the GLMM, which found no significant instruction effect, nor any significant interaction between instruction and RT.



(a) Listen for a target sound



(b) Concentrate on something other than sound

Figure 3.8: Box plot of the preference proportion for each compression setting for each instruction set. The top panel shows the preference proportions for the condition when participants were instructed to "listen for a target sound" and the bottom panel shows the choice frequencies for the condition when participants were instructed to "imagine concentrating on something other than sound." The boxplot boundaries are as described in figure 3.7.

3.3.5 Influence of Stimuli on Preferred CT

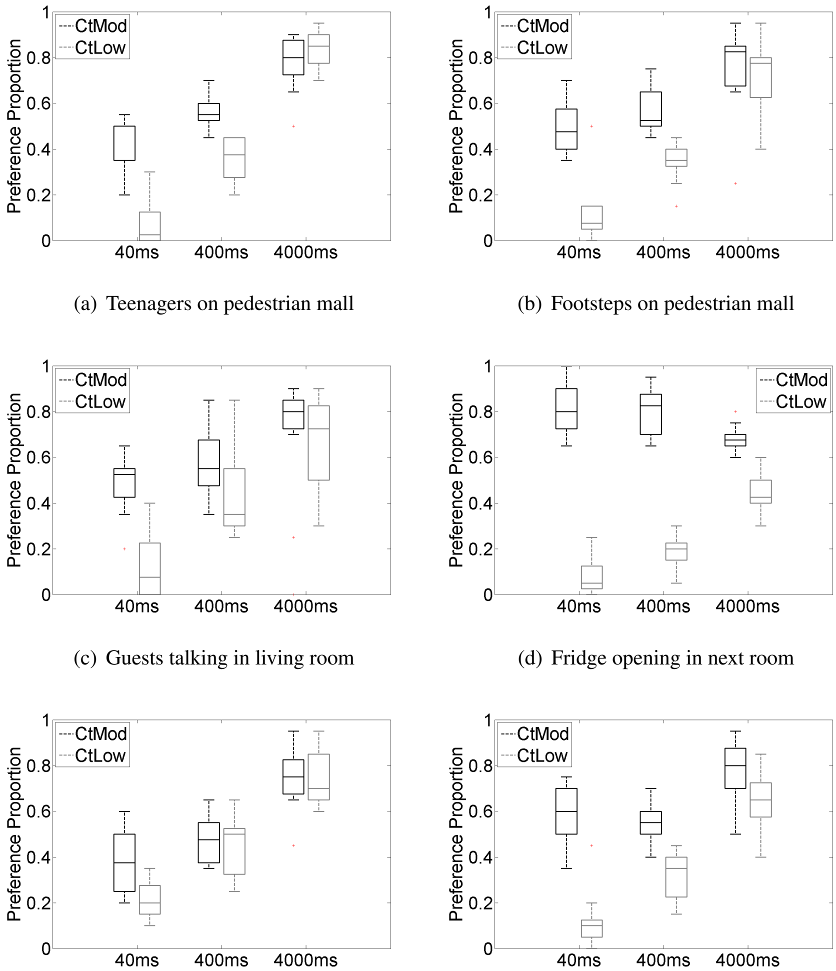
Figure 3.9 shows boxplots of the preference proportions for each setting, plotted separately for each test signal. The GLMM analysis showed that stimuli had an influence on preference proportions ($p < 0.01$). The boxplots in figure 3.9 show that the choice pattern for the 'fridge opening in the next room' test signal was quite different than that for the other five signals. When RTs are considered, the RT=40 ms/CTMOD and 400 ms/CTMOD combinations were chosen most often for the 'fridge opening in the next room' test signal and this is in contrast to the other five stimuli, where the long RT (4000 ms) was preferred most often. When CTs are considered, the CTMOD preference proportion was higher and the CTLOW preference proportion lower for the 'fridge in next room' signal than for the other signals.

3.3.6 Illustration of the Acoustic Effects of Compression on the Signals

Figure 3.10 shows the effect of the compression for the "guest in living room" signal on the estimated output levels in the right ear canal of participant KA. The main effect is that 10th percentile levels (L_{10}) are influenced by varying either the CT or the RT. When the RT is shortened or the CT lowered, the levels of the soft components of the test signal (L_{10}) increase. In contrast, the RMS level, and 50th and 90th percentiles were not altered much by varying the CT or RT.

3.4 Discussion

For the shorter RTs (40 ms and 400 ms), participants clearly preferred CTMOD over CTLOW, but for the longest RT (4000 ms), there was no significant difference between the preference proportions for CTLOW and CTMOD. Also the instructions had no influence on the preferred CT. Finally, the pattern of results was similar for five of the six signals used, in that the longest RT was chosen as 'preferred' most frequently. For one of the signals, 'fridge from room next door', the preference proportion pattern was markedly different, in that this was the only signal for which the shortest RT (40 ms) was preferred most frequently.



(a) Teenagers on pedestrian mall

(b) Footsteps on pedestrian mall

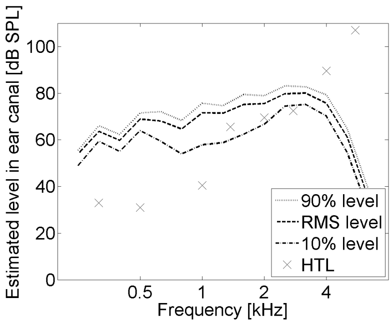
(c) Guests talking in living room

(d) Fridge opening in next room

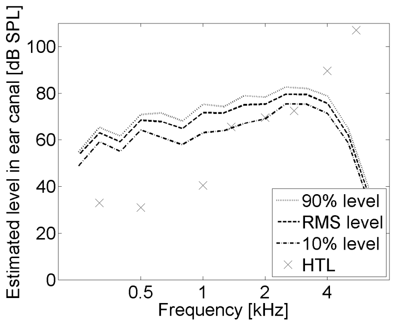
(e) Two customers talking in supermarket

(f) Transaction beeps at supermarket

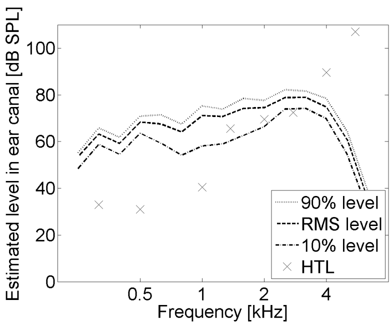
Figure 3.9: Separate boxplots of the preference proportions for each setting for each test signal. The data are pooled from all participants for both sessions and both instruction sets. The boxplot boundaries are as described in figure 3.7.



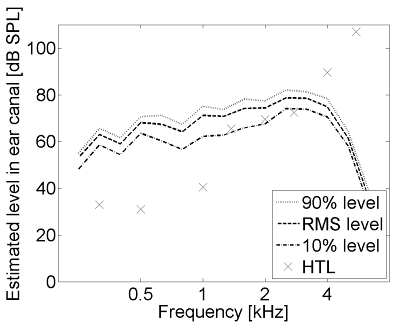
(a) CTMOD, RT = 40 ms



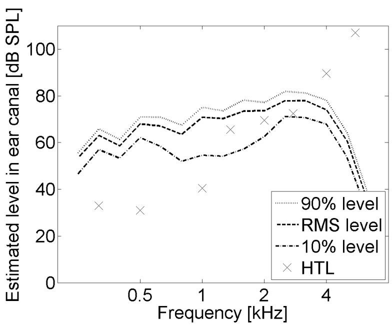
(b) CTLOW, RT = 40 ms



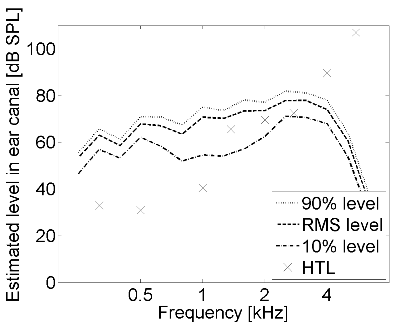
(c) CTMOD, RT = 400 ms



(d) CTLOW, RT = 400 ms



(e) CTMOD, RT = 4000 ms



(f) CTLOW, RT = 4000 ms

Figure 3.10: The effect of the six compression settings on the output level distribution of the "guest in living room" signal plotted as estimated aided levels in the ear canal for a randomly-selected participant (KA). The RMS and 10th and 90th percentiles are plotted in $1/3$-octave bands using a 125-ms analysis window. The levels were transformed to an individual estimate of aided dB SPL levels in the left ear canal as described in section 3.2.4. The black crosses show participant KA's left hearing thresholds converted to dB SPL using ISO 389-8 (2004).

The results support the first prediction of this study that the longer the hearing aid RT, the more likely it is that hearing-impaired participants will prefer a low CT. When a relatively short RT was used (40 and 400 ms), the participants preferred CTMOD more often than CTLOW and this result is consistent with studies carried out at the NAL (Barker and Dillon, 1999; Barker et al., 2001; Dillon et al., 1998). In contrast, when a long RT was used, the preference for CTs was equivocal. Figure 3.10 offers some explanation, in that the combination of a short RT (40 ms) and CTLOW results in a reduction in signal dynamic range. For the long RT, the CT did not impact the signal dynamic range very much, and it would be more difficult for the test subjects to perceive the difference in processing caused by changing the CT.

This study was carried out using a multi-channel compressor, while the earlier NAL studies were carried out using a single-channel compressor. They found that moderate CTs were preferred more often than low CTs. It was anticipated that results between the current study would differ from the earlier studies, because with single-channel compression, the gain of the whole signal is regulated by the dominate sounds in the sound environment and this is potentially annoying for the HA user. In spite of the current study using multi-channel compression, results were similar, indicating that other factors were important in the participants' preference for compression settings.

The results did not support the second prediction of this study, that when participants were instructed to listen for a weak target signal within the test sound that they would be more likely to choose a low CT. The finding that instructions had little influence on the preferred CT is surprising, especially giving the large body of evidence that instruction influences the subjective evaluation of other hearing aid settings (e.g. CR in Keidser et al., 2005). There are two potential problems with the experimental design used here that could explain the result. Firstly, while the overall levels of the recorded sound environments were known, the levels of the target sounds were not known precisely because they were spontaneously occurring acoustic events in the sound environments during the recordings. If the levels of the targets were too high, then the altering the CT would not alter the level of the target sound. Secondly, the participants were asked to "imagine themselves in the situation" and this demands a certain level of abstraction on their part to imagine how they might react to the compressed signal. It is hard to know if the participants can do this adequately.

Another general finding was that the extent of preference was greatest for the longest RT (4000 ms). This finding is consistent with results from a number of other researchers (Hansen, 2002; Neuman et al., 1998; Schmidt, 2006). The exception was for the signal 'fridge from room next door'. This was recorded at the lowest SPL of all the signals and a number of participants mentioned a 'shhh' sound on this recording, presumably due to microphone noise and this 'shhh' sound is more marked in the CtLow condition. It is interesting that, when audible microphone noise is present, hearing-impaired participants instead preferred a short RT which also increases the level of the low-level components of the signal. The author's own impression after listening to the compressed signals was that when the fastest RT was used, the microphone noise is higher in level but seemed constant in level. In contrast, when the longest RT was used, the microphone noise appeared and disappeared as the compressor gain was regulated. The disappearing and re-appearing of the microphone noise is perhaps more disturbing for the test participants than having a constant microphone noise.

One potential limitation of the current study was that while on average the hearing aids were fit well to the NAL-R rationale, there was a large spread in how well the actual insertion gain met the targets. The standard deviation at each frequency was 4-5 dB for the difference between the achieved and measured gain. It is difficult to know how this would affect results. Since some test participants received too much gain and other participants received too little gain, the effect would average out across participants.

There are other limitations in the current study regarding the implementation of the hearing aid fitting. The first was that a constant CR of 2:1 was used in all channels to allow comparison with the NAL studies (Dillon et al., 1998; Barker and Dillon, 1999). However, the implementation of a constant CR in a multi-channel is not usual in practice. Most generic fitting rationales recommend CRs that vary with frequency and degree of hearing loss. For sloping hearing losses, such as those used the study, the main generic rationales would recommend very little compression (CR< 1.5) at low frequencies and more compression (CR> 2) at high frequencies (Byrne et al., 2001;

Moore, 2000; Scollie et al., 2005). Another limitation is that a constant vent size of 1 mm was used. In clinical practice, larger vent sizes would be used and as vent sizes increase, the effective CR decreases because more direct sound enters the ear canal via the vent and mixes with the amplified sound (M. Nordahn, personal communication).

Regarding the choice of participants in the current study, the participants had fairly uniform, moderate, sloping sensorineural hearing losses. The present findings may not hold for other degrees or types of hearing loss. If the hearing losses were more mild, it would be expected that the A/B paired comparison decisions would be more often influenced by microphone noise. If the hearing losses were severe and profound, previous research indicates that there would be more inter-subject variability but more participants would prefer the hearing aid setting that operates most linearly (Barker et al., 2001; Keidser et al., 2007; Souza et al., 2005), presumably to preserve the speech envelope better. Another limitation is that all the participants were experienced HA users and the results could have been influenced by a hearing aid acclimatisation effect (Keidser et al., 2006).

The clinical implication of this study is that, for HA users with moderate hearing losses, the prescription for CT (and thereby gain at low input levels) should depend on the compression speed. Some of the most commonly used generic hearing aid rationales, such as the Cambridge Method for Loudness Equalisation (CAMEQ Moore et al., 999a,b) and NAL-NL1 (Byrne et al., 2001), recommend setting CTs to just under the level of normal speech in each channel. The current results suggest that when a long RT is used, it would be appropriate to reduce the CT for some clients, as this showed not to compromise the sound quality. Also by reducing gain at low input levels, this would improve gain and hence, audibility in low-level listening situations.

The current study showed clearly that the preferred CT depends on the compression RT. As previous research has indicated that there is not always a good correlation between preferred hearing aid settings in the laboratory and preferred hearing aid settings in the field (Smeds et al., 2006b), it was decided to extend this study to a field trial to determine if compression RT influences preferred CT when the hearing aids are worn by HA users in their own listening environments. The field trial also included both experienced and inexperienced HA users, thus addressing one of the potential limitations of this study. The field trial results are reported in chapter 4.

3.5 Conclusion

In this laboratory study, experienced HA users with moderate, sloping sensorineural hearing losses made paired-comparison judgments of real-life environmental signals processed with multi-channel compression with different combinations of CT and RT. Results showed that the preferred CT depends on the RT. For the two shorter RTs tested (40 and 400 ms), the participants clearly preferred moderate CTs more often than they preferred low CTs. This finding is consistent with findings from other researchers. The novel finding of this study was that, for the longest RT tested (4000 ms), participants preferred the moderate and low CTs almost equally. The results also showed no influence of instructions on the preferred CT. Since the preference for CT depends on the compression RT, this suggests that gain recommendations for low input levels should depend on the compression RT, such that more gain at low input levels can be allowed when combined with long RTs, in order to improve audibility in low level listening environments.

Acknowledgments

Thanks to Carsten Paladan-Müller for making the MATLAB SIMULINK compressor model. Thanks to Ole Hau for help with calibrating the input to the hearing aid DAI and Brent Kirkwood for help with making the audiogram figure. Thanks to Lars Sørensen and Kristoffer Jørgensen for wearing the recording devices during the recordings. Thanks to Professor Brian C.J. Moore and Associate Professor Torben Poulsen for helpful discussions during the development of this manuscript.

4

Field trial and laboratory investigation into the preferred hearing aid compression threshold

Experiment carried out by Helen Connor, Sara Båsjö and Karolina Smeds.

Abstract

There is a lack of empirical evidence about hearing aid (HA) gain recommendations at low-input levels. Twenty hearing-impaired participants (10 new and 10 experienced HA users) with mild-moderate hearing losses participated in a field trial supplemented with paired comparisons in the laboratory. They compared moderate and low compression thresholds (CT) implemented in two programs in a 15-channel HA, combined with either fast-acting or slow-acting compression in a double-blind, cross-over design. In the field trial, the overall CT preference was influenced by previous HA experience but not by compression speed. i.e., compared to the inexperienced HA users, the experienced HA users were more likely to prefer low CT. In specific listening situations, the participants most often preferred the moderate CT, except for situations with speech in quiet when combined with slow-acting compression. In the situations the participants nominated themselves as important, most participants did not have a preference for either a low or moderate CT. For the laboratory paired comparisons, the participants chose the moderate CT at moderate input levels and the low CT at low input levels. Overall, the results were not strongly in favour of the use of either a low or moderate CT. The findings also provide preliminary evidence that CT preference is related to the HA gain acclimatisation effect.

4.1 Introduction

People with sensorineural hearing loss have a loss of audibility for low-level sounds. Concurrent with this loss of audibility is a reduced dynamic range, which means to say that the range between hearing threshold and the level of loudness discomfort is often reduced (Fowler, 1936). This is often managed audiologically by fitting non-linear Hearing Aids (HA). One of the most commercially used form of non-linear HA amplification is Wide Dynamic Range Compression (WDRC). This type of amplification gives most gain to low level inputs and reduces gain for high level inputs, and this should improve audibility for low-level sounds while avoiding loudness discomfort. Compared to linear HA amplification, WDRC amplification improves loudness comfort over a wider range of inputs (Jenstad et al., 2000) as well as improving audibility for low level speech (Jenstad et al., 1999). In spite of the positive findings regarding WDRC not a lot is known about how much amplification is appropriate for low level inputs and there is a lack of consensus between HA rationales for gain targets for low level inputs (Byrne et al., 2001; Marriage et al., 2004).

One of the most important parameters in a HA for determining the audibility for soft sounds is the compression threshold (CT). When the gain at medium and high input levels is fixed, then lowering the CT, increases gain at low input levels. However, there is not a lot known about how CT should be fitted. In the late 1990's, Dillon et al. (1998), Barker and Dillon (1999) and Barker et al. (2001) investigated preferred CT in a series of field studies with a total of 172 participants with hearing impairment. They found that around two-thirds of HA users preferred to have a moderate level CT ($\sim$65 dB SPL) rather than a low level CT ($\sim$40-57 dB SPL). However, these studies were performed using a single-channel, fast-acting HAs. The instantaneous gain applied by a single-channel HA is usually controlled by the dominant sound source and the gain applied across the whole frequency spectrum is regulated up and down in tact with the characteristics of the dominant sound source (Laurence et al., 1983; Stone and Moore, 1992). This is potentially annoying for the HA user, particularly when combined with fast-acting compression, as this can have the effect that gain increases during the pauses of speech, thus increasing the level of any constant, low-level, background noise present. Hence, the results from Barker and Dillon (1999) and Dillon et al.

(1998) cannot necessarily be applied to HAs with multiple channels or longer time constants.

Chapter 3 investigated the preferred CT for HA users in a laboratory listening study. Twelve participants with mild-moderate hearing losses made paired comparisons of real-life environmental stimuli processed with different HA settings. The stimuli were compressed using a 15-channel offline compressor model with a fixed 2:1 Compression Ratio (CR). The compression settings were two CTs (low or moderate level) and three Release Times (RT = 40, 400 and 4000 ms). The compressed stimuli were played back to the participants via the Direct Audio Input of HAs fitted linearly to the National Acoustic Laboratory - Revised (NAL-R) rationale (Byrne and Dillon, 1986). The results showed that as RT increased, the extent of preference for the low CT also increased. In other words, the HA user's preference for CT was dependent on compression speed. This may well be because by increasing the RT, the subjective perception of pleasantness and comfort improve (Hansen, 2002; Neuman et al., 1998) and the output signal-to-noise ratio will also often improve (Naylor and Johannesson, 2009; Souza et al., 2006). Hence the test participants may have been more willing to accept that the HA spends more time in compression when a low CT is used because some of the negative side effects of compression are ameliorated by using a long RT.

The general aim of the current study is to further investigate if the findings in chapter 3 are substantiated in a field trial. When investigating benefit from various HA settings, it is generally prudent to confirm laboratory findings in a field trial because laboratory results do not always agree with field trial results. For example, Savage et al. (2006) found that test participants with hearing-impairment do not exhibit a preference for a particular type of output limiting compression type in the field but the participants had clearer preferences in the lab. In contrast, Xu et al. (2008) found that test participants with hearing-impairment do not give significantly different ratings for fast or slow release times in the lab but most participants expressed a clear preference in the field. Finally, Smeds et al. (2006b) found hearing-impaired test participants prefer a higher volume control (loudness) setting in the lab than in the field, particularly at low input levels. Since there are often discrepancies between lab and field data, it is important to assess preference for HA settings on the basis of the test participants' listening experience in their own listening environments. Or in other words, it is important to establish that a treatment is "effective" as used in normal circumstances and not just that a treatment shows "efficacy" in laboratory conditions (Haynes, 1999).

Another research focus in the current study is whether experienced and inexperienced HA users have different preferences for CT. There is anecdotal evidence that new HA users prefer less gain than experienced HA users, particularly at low input levels (see Convery et al., 2005, for review). In contrast, the research evidence suggests that if this effect is real, then the effect is small and dependent on the degree of hearing loss and possibly the type of HA processing i.e., linear or non-linear amplification, (Yund et al., 2006). For instance, Cox and Alexander (1992), Horwitz and Turner (1997) and Humes and Wilson (2003) investigated preferred gain usage up to 3 years post linear HA fitting and found little evidence of increasing gain (volume control) preference following first time HA fitting.

More recently, there has been evidence of a gain acclimatisation following non-linear HA fitting. Marriage et al. (2004) investigated differences in "acceptable" gain for a group of 20 new and 20 experienced HA users. They found new HA users accepted an average 2.6 dB lower gain than experienced HA users for an acceptable non-linear HA fitting. It should be noted that the experienced HA users on average had a greater degree of hearing loss than the inexperienced HA users. And also that the listening criterion was "acceptable" setting rather than a "preferred" setting. Finally, Keidser et al. (2008) investigated preferred gain at medium levels during the one-year period following non-linear HA fitting for 50 new HA users with varying degree of hearing loss. They found the new HA users preferred on average 2.7 dB less gain than a control group of 26 experienced HA users. They found for the new users that the gain preference increased during the 13-month period post HA fitting but the gain acclimatisation was not complete 13-months post-fitting. They also found that the gain acclimatisation effect exhibited among the participants with an average hearing loss greater than 43 dB HL, whereas participants with milder hearing losses did not exhibit a significant effect.

The research so far has concentrated on gain acclimatisation at medium-input levels, so it is not clear whether the gain acclimatisation effect would be enhanced or reduced at low input levels. We hypothesize that gain acclimatisation would also be present for low input levels because neural plasticity, which is the neural process believed to underlie acclimatisation, is related to the amount and time of exposure to a stimulus (Palmer et al., 1998). When hearing impaired individuals have a gradual onset hearing loss, it can take many years from when they first notice the hearing loss to when they get a HA. In the time in-between, they are deprived of some auditory information at low-input levels. When they finally get a HA, audibility is improved for a range of frequencies and input-levels, particularly for non-linear amplification. Since the new HA user is receiving access to information that they have not heard for years, it is plausible that it would take time to adjust to hearing a greater range of inputs. Lindley et al. (2000) made a case study of 3 new non-linear HA users and found that their subjective rating of the aversiveness of environmental noises improved during the first 3 months following HA fitting, presumably as the users adjust to their newly re-acquired access to low-input levels. On the other hand, Munro and Lutman (2005) found in a group of 16 monaurally-fit linear HA users, that there was no change in sound quality judgments in the 24-week period following fitting.

Finally, there is not a lot of information about in which listening situations HA users prefer low CT. Previous research indicates that preferred compression settings are highly dependent on the listening situation, both in terms of listening environment and listener intention (e.g. Keidser et al., 2005). Many modern HAs adjust gain adaptively depending on the listening environment. In order that the current study should provide useful information for adaptive gain algorithms, we have more knowledge about the interaction between listening environment and preferred CT.

The current study uses a combination of field trial and laboratory listening experiment to investigate the following research questions:

1. Is the preferred CT influenced by compression speed (i.e., fast- or slow-acting compression);

2. Does previous HA experience influences the preferred CT; and

3. Is the participants' preferred CT influenced by the listening environment i.e. the listening situation in the field and the choice of stimuli in the lab.

4.2 Method

4.2.1 Test Participants

The test participants were 10 inexperienced and 10 experienced HA users. All participants were recruited via a publically-funded HA clinic in Stockholm (Avesina Hörsel-rehab). The inexperienced HA users were recruited from a waiting list for HAs and the experienced HA users were recruited from a database of patients who had recently renewed their HAs.

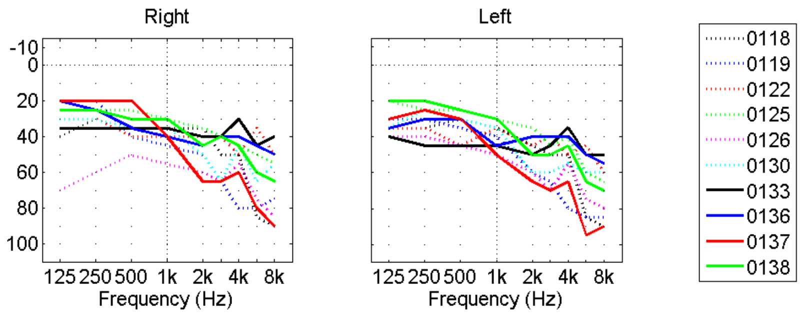
Audiometric Information

Air and bone conduction audiometry was performed using insert earphones, using procedures described in ISO 8253-1 (1989). The participants had symmetrical[1] moderate sensorineural hearing losses (see figure 4.1) with the exception of participant 0126, who had poorer low frequency thresholds on the right than on the left side. None of the participants had more than a 15 dB difference between air and bone conduction thresholds at any audiometric frequency between 500 to 4000 Hz.

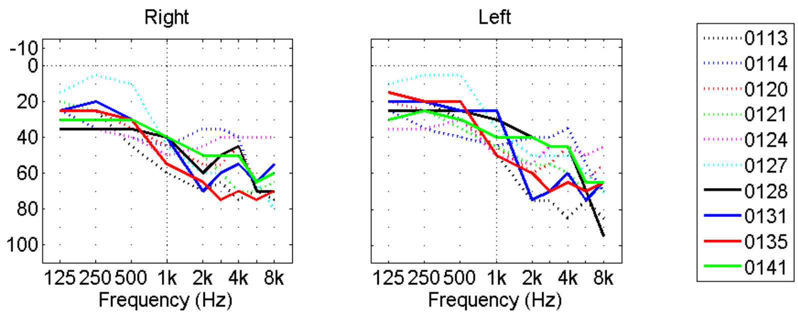
Uncomfortable Loudness Levels (UCL) were measured using pure tones at 500, 1000, 2000 and 4000 Hz on both ears using an ascending procedure with verbal instructions. All but one participant had a dynamic range (difference between UCL and hearing threshold) of at least 30 dB at 500, 1000, 2000 and 4000 Hz. Participant 0126 had an average dynamic range of 20 dB but did not show any signs of loudness discomfort when exposed to any other high-level stimuli during other procedures.

Tympanometry was performed in both ears. All participants had normal middle ear pressure[2] for both ears as indicated using tympanometry with the exception of one participant (participant 0119), who had -130 daPa middle ear pressure on his left side.

[1] Symmetrical hearing loss is defined as no more than a 10 dB difference between ears for the four frequency Pure Tone Average (4PTA); and no more than 15 dB difference at any audiometric frequency.
[2] Normal middle ear pressure is defined in the range of +50 to -100 daPa



(a) Inexperienced HA users



(b) Experienced HA users

Figure 4.1: Individual air conduction thresholds of the test participants.

Additional Participant Information

Eight of the participants were men and twelve of the participants were women. The age range of the participants was 48 to 80 years with a median of 70.5 years. Fifteen of the twenty participants spoke Swedish as a native language and the remaining five participants spoke fluent Swedish. Four of the participants were still in employment and the remaining sixteen participants were retired.

All ten of the experienced HA users had worn bilateral HAs for a total of at least five years and had renewed their HAs with modern non-linear HAs within 6 months of starting the current study. Five of them normally wore behind-the-ear (BTE) HAs with standard earmoulds, four of them wore BTE with open fit (thin tubing ear pieces) and one participant wore in-the-ear HAs. All experienced HA users reported wearing their own HAs daily.

Participants gave their written consent for participation at the start of the study. They did not receive financial compensation for their participation but they received a small gift at the end of the study to the value of approximately €10.

4.2.2 Experimental HAs

Participants were fitted binaurally with Widex Inteo IN-9 BTE HAs with standard earmolds. The HAs were not marked with either a manufacturer or a brand name. The participants could switch between two programs using a push button on the HAs but the volume control was disabled. The HAs had 15-channels and the bandwidth of each channel was approximately 1/3-octave. Aside from the compression, the other adaptive features in the HA were disabled, i.e., the directional microphones, noise reduction, and feedback system.

Participants compared two CT settings, CTMOD and CTLOW. For the CTMOD setting, the CT in each channel follows the 1/3-octave Long-Term Average Speech Spectrum (LTASS) for normal speech at 62 dB SPL (ANSI S3.5, 1997). For the CTLOW setting, the CT in each channel was 20 dB lower than CTMOD in each channel. Above the CT in each channel, a fixed compression ratio (CR) of 2:1 was applied. Figure 4.2 shows the static input-output curve of the compressor.

Two compression speeds were used: FAST and SLOW. The attack times were 14 ms and 630 ms and the release times were 44 ms and 5672 ms, respectively. The attack and release times are defined relative to the IEC 60118-2 (1983) standard. That is, the time taken for the HA output to stabilise to within 2 dB of the final output following an abrupt increase or decrease in input level between 55 and 80 dB SPL. The attack and release times were confirmed using measurements in a 2cc-coupler using the Interacoustics NOAH-based HA measurement module, HIT440.

For all compression settings tested, for a speech input signal at 62 dB SPL, the gain was adjusted individually to fit the HAs to the NAL-R rationale (Byrne and Dillon, 1986) minus a 3 dB correction for binaural loudness. The procedure to verify the

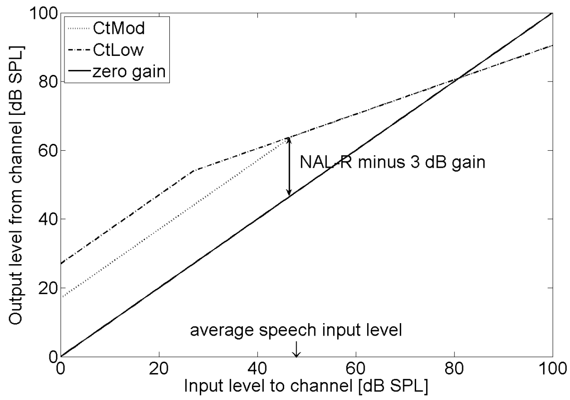


Figure 4.2: Schematic illustration of the input-output curve of the compressor in one channel. CTMOD were set to the level of normal speech in each channel (i.e., total level 62 dB SPL). CTLOW was set 20 dB below the CTMOD threshold for each channel. Gain for medium and high levels was the same for CTMOD and CTLOW. The gain applied for a normal-speech input was adjusted individually for each participant to meet the NAL-R rationale minus a 3 dB binaural correction.

insertion gain is described in the section 4.2.4. The gain was adjusted using a Graphical User Interface (GUI) written in MATLAB and programmed via the NOAHLink. The GUI also controlled the assignment of programs and compression speeds in a way that was blinded to the experimenters.

The earmolds were acrylic halfshell earmoulds with standard #13 tubing. The venting size of the earmold was selected based on recommendations in the NAL-NL1 fitting software (Brewer, 2005) and ranged from 1 to 3 mm diameter (median 2 mm).

4.2.3 Procedural Overview

The participants participated in two successive two-week trials. In each of the trials, the participants compared two different CT settings (CTLOW and CTMOD) combined with one of two possible compression speeds (FAST or SLOW). The order of the trial periods was counterbalanced, such that half of the participants started with FAST compression in the first trial and half of the participants start with SLOW compression in the first trial. Within the trial, the order of the programs (CTLOW and CTMOD) was also counter-balanced. The assignment of CT and compression speeds was double-blinded, so that neither the participant nor the experimenter knew the order of assignment. At the end of each trial, the participants reported in an interview both their overall program preference and their program preference in a number of listening situations. The interview was supplemented with two questionnaires: the Speech, Spatial and Qualities of Hearing Scale (SSQ, Gatehouse and Noble, 2004) and the Client Oriented Scale of Improvement (COSI, Dillon et al., 1997). In addition to the field study, the participants also performed in a laboratory paired comparison experiment and speech intelligibility testing.

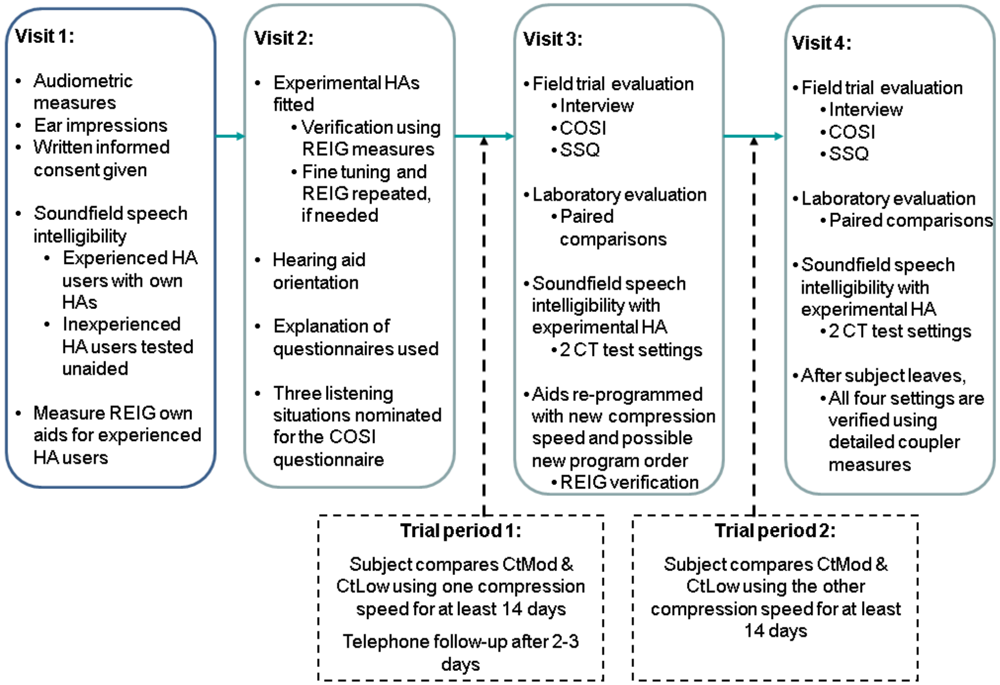
Altogether, the participants took part in four appointments and the details of the procedures undertaken at each appointment are described in table 4.1. All testing took place at the Office of Research in Clinical Amplification - Europe (ORCA Europe) in Stockholm, which is an external research laboratory for Widex A/S. The name Widex appears on the office signage. All the tests were performed in a soundproof booth.

4.2.4 HA Fitting

The HAs were fit by two clinical audiologists (Helen Connor and Sara Båsjö). The steps to fit and verify the HAs during the fitting appointment were as follows:

1. The physical fit of the earmoulds was checked.
2. The gain of the HAs was pre-adjusted based on hearing thresholds. For the safety of the participants, the coupler gain of the HAs was checked using the International Speech Test Signal (ISTS) (Holube et al., 2009) at 62 dB SPL.
3. Insertion gain was measured on each ear as described in subsection 4.2.4. The gain was adjusted to meet the NAL-R minus 3 dB target.
4. The experimenters had an informal with the participant to ascertain that the gain settings were acceptable. Fine tuning was discouraged unless the participant said they could not wear the HA as it was. If the overall gain was perceived to be unacceptably loud, the gain was reduced in all frequency bands by 3 dB. If the sound of their own voice was unacceptably loud or boomy, the gain for bands under 1 kHz was reduced by 3 dB. Finally, if there was audible feedback, the gain was reduced by 3 dB in the frequency bands above 3.2 kHz. After any

Table 4.1: Overview of the experimental design and procedures at each appointment.



necessary fine-tuning was performed, the insertion gain was re-measured for both ears.

5. After 2-3 days, there was a telephone follow-up. Only one participant returned because he thought one HA was too quiet on his right ear (participant 0135). After checking the performance of the HA and finding no fault with it, we increased the gain by only 3 dB. One of the participants also needed a gain reduction between trial one and trial two (participant 0127).

Real-Ear Insertion Gain Measurement

Real Ear Insertion Gain (REIG) was measured for the experimental HAs at the beginning of each trial period, as well as measuring the experienced HA users' own HAs. The test signal was the ISTS signal (Holube et al., 2009) presented at 0° degrees azimuth and averaged over a 45-second presentation time. When measuring the HA users' own HAs, the signal levels were 50, 62 and 70 dB SPL with a 30-second pre-conditioning time. When measuring the experimental HAs, the signal level was 62 dB SPL with a 10-second pre-conditioning time. The reference microphone sat just below the ear. The equipment was the Interacoustics REM440 Noah module.

In order to improve measurement accuracy in the high frequency region, the probe tube was placed by first finding the insertion depth which gave the minimal sound pressure level for a 6 kHz warble tone signal and then pushing the probe tube a further 8 mm deeper beyond this point (as described in Dillon, 2001, p. 94). Otoscopy was performed both before and after probe tube insertion.

Coupler Gain Measurements

Coupler gain of the four compression settings for each participant was measured using a 2cc-coupler (IEC 60813-5, 1985). The signal was the ISTS signal presented at levels between 40 to 90 dB SPL in 10 dB increments, as well as the normal speech level of 62 dB SPL. The coupler gain was averaged over one minute, after a 30 second pre-conditioning time at each level. Coupler gain was measured using the Interacoustics Equinox HIT440 HA test module with a 2cc coupler (as specified by IEC 60813-5, 1985) and the Interacoustics TBS25 testbox.

4.2.5 Field Trial Evaluation Procedure

The participants compared two settings at a time in two trial periods, each lasting at least 14 days. At the end of each trial period, the participants reported their experiences in an interview supplemented with the SSQ and COSI questionnaires.

At the fitting appointment (visit 2), the participants were given verbal and written instructions about how to operate their HAs, including how to change program. They were instructed to wear the experimental HAs as much as possible and to change programs often, particularly within the same listening environment. The experimenter explained the format of the SSQ and COSI questionnaires to the participants. For the use of the COSI, each participant nominated three important listening situations.

Interview

The interview was given in Swedish by Sara Båsjö, who was blinded to the program order and compression speed. The translations of the questions into English are given in Table 4.2 and the original questions in Swedish are given in appendix B.

SSQ Questionnaire

The SSQ (Gatehouse and Noble, 2004) is a questionnaire with 50 questions divided into three sections: Speech, Spatial and Quality. It was translated into Swedish by Professor Stig Arlinger at Linköping University. It was included in the current study because it has questions related to audibility, localisation, and speech quality in noise. Previous observation in chapter 3 suggested that “noisiness of speech” was an issue.

COSI Questionnaire

For the standardized Client Oriented Scale of Improvement (COSI) questionnaire, the participants nominates up to five listening situations in which help with hearing is required. In the current study, the participants nominated only three listening situations and at the conclusion of each field trial, the participants reported which HA program (CTMOD or CTLOW) they preferred in each situation.

Table 4.2: Interview questions used in the interview at the end of each trial period. The text in paragraphs indicated the response categories (if any). The interview was translated from Swedish. The table continues on the next page

	Question
1	How has it been? (Open response.)
2	Which program did you prefer generally? (Forced choice between program 1 and program 2.)
3a	Which program is the most comfortable? (Participant could choose between program 1 and program 2, if they could hear a difference.)
3b	How do you rate the comfort/pleasantness of each program? (Response alternatives were "Very good", "Good", "In-between", "Bad" and "Very bad".)
4a	Which program was the clearest for speech in a normal conversation in a quiet environment? (Participant could choose between program 1 and program 2, if they could hear a difference.)
4b	How would you generally describe speech in a quiet environment in each program? (Response alternatives were "Very good", "Good", "In-between", "Bad" and "Very bad".)
5a	Which program do you prefer when the environment is quiet and there are low-level sounds present? (Participant could choose between program 1 and program 2, if they could hear a difference.)
5b	Do you like being able to hear low-level sounds? (Response alternatives were "yes" and "no" with the possibility of comments.)
5c	Which low-level sounds do you want to hear? (Open response)
6a	Which program was clearest for quiet or distant speech? (Participant could choose between program 1 and program 2, if they could hear a difference.)
6b	How would you generally describe speech clarity for quiet voices or distant speech in each program? (Response alternatives were "Very good", "Good", "In-between", "Bad" and "Very bad".)
7a	Which program had the clearest speech for speech in noise? (Participant could choose between program 1 and program 2, if they could hear a difference.)

Table 4.2: Interview questions continued from the previous page. The questions were used in the interview at the end of each trial period. The text in paragraphs indicated the response categories (if any).

	Question
7b	How would you describe the clarity of speech in noisy situations in each program? (Response alternatives were "Very good", "Good", "In-between", "Bad" and "Very bad".)
8a	How is the level of background noise in noisy situations in each program? (Response alternatives were "Too quiet", "A bit too quiet", "OK", "A bit too loud" and "Too loud".)
8b	Which program is the most noisy in noisy listening situations? (Participant could choose between program 1 and program 2, if they could hear a difference.)
9	How would you describe the sound quality of the your own voice in each program? (Response alternatives were "Very good", "Good", "In-between", "Bad" and "Very bad".)
10	Can you hear the HA making a noise of it's own (microphone noise) in quiet situations in either of the programs? (Response alternatives were "yes" and "no".)
11	Have you had problems with feedback in either of the programs? (Response alternatives were "yes" and "no".)

4.2.6 Laboratory Paired Comparisons Procedures

Paired Comparison Evaluation in the Laboratory

The stimuli were pre-mixed combinations of (1) running Swedish speech recorded for this project (see the description in appendix C) and (2) background noise from the ICRA recordings of natural sound environments (Bjerg and Larsen, 2006, XY recordings). The speech was presented from 85 cm distance at 0° azimuth and the background noise presented in stereo from 1 m distance simultaneously at 45 and 315° azimuth. Table 4.3 summarizes the stimuli and presentation levels used. At the start of each stimulus, there was a fifteen-second pre-amble where the speech and noise are played at level and this pre-amble was used as a pre-conditioning period for the gain of the HA to stabilise. The stimuli levels were calibrated in the soundfield at the position of the participant's head using a class 1 sound-level meter.

Table 4.3: Description of the stimuli used in the paired comparison evaluation in the laboratory. SNR stands for Signal to Noise Ratio.

Stimuli	Background signal	Speech level (dB SPL)	Noise level (dB SPL)	SNR (dB)
1	Supermarket	60	55	+5
2	Supermarket	55	45	+10
3	Supermarket	50	45	+5
4	Supermarket	45	35	+10
5	Pedestrian Mall	60	55	+5
6	Pedestrian Mall	55	45	+10
7	Pedestrian Mall	50	45	+5
8	Pedestrian Mall	45	35	+10

The Supermarket and Pedestrian Mall signals were selected because we wanted to investigate the effect of modulation on the setting preference, as measurements by Naylor and Johannesson (2009) suggest that the degree of background modulation has an effect on the output SNR of a HA compressor. The Supermarket and Pedestrian Mall signals have similar long term average spectrum but the Supermarket signal has more fluctuation that the Pedestrian Mall signal. We also wanted to investigate the effects of signal level and SNR on CT preference, but there was a mistake made during the programming of the signal presentation interface and the result was that signals 3, 4, 7 and 8 were presented at lower levels than we intended. The actual presentation levels, as measured using a class-1 sound level meter at the position of the listener, are those shown in table 4.3.

At the start of the paired comparison procedure, the participant was instructed to judge which HA program (program 1 or 2) they "hear best with" while listening to running speech and noise. After the pre-conditioning period of each stimulus, an experimenter switched between the two HA programs approximately every three seconds using a remote control. The participant could hear the HA switch programs via indicator beep tones. The participant could listen to each stimulus as long as needed before verbally reporting their preference to the experimenter who recorded the response and started a new stimulus. For each participant at each session, a run of all eight stimuli was repeated three times, with a double-blinded re-randomisation of program order between each run. The first run was used as a practice run to give the participant the opportunity to establish their listening criteria. The order of the stimuli was counter-balanced within and across participants.

Speech Intelligibility Testing

Speech intelligibility testing was carried out at the initial appointment and at the end of each field trial. At the initial appointment, the new HA users were tested unaided and the experienced HA users were tested aided with their own HAs. In order to avoid bias from the field trial experiences, the order of the HA programs (CTMOD/CTLOW) was re-randomised before starting speech intelligibility testing. The speech material was the Swedish phonemically balanced words (Magnusson, 1995) and the background noise was the International Collegium of Rehabilitative Audiology (ICRA) unmodulated speech-shaped background noise (Dreschler et al., 2001, Track 1). Both the speech and noise were presented from 0° azimuth at a 85 cm distance from the centre of the participant's head. Two presentation levels were tested: a) speech at 71 dB SPL and noise at 65 dB SPL, and b) speech at 56 dB SPL and noise at 50 dB SPL.

In the Swedish Phonemically Balanced (PB) word lists, there are 50 monosyllable words per list and each word is proceeded by a carrier phrase and scored with whole word scoring. The word lists were counterbalanced across participants, presentation levels and HA settings. Care was taken that for each participant no word list was repeated within an experimental session.

4.2.7 Statistical Analysis

Analyses were performed using the Generalized Linear Mixed Model (GLMM), which can analyse both binomially-distributed data and normally-distributed data. Compared to many non-parametric tests, it also has the advantage of being able to include many explanatory variables, including repeating effects variables. All analyses were performed in the R statistical environment (R Development Core Team, 2008), which is a freeware package developed by and widely in use among professional biostatisticians. The specific analysis package used was `lme4`.

When performing multiple regression analyses, there are decisions to make about which explanatory variables (model terms) to include. If too many irrelevant terms are included, the standard error terms will inflate and the analysis will lose power, but if relevant terms are excluded then the analysis model will insufficiently explain the data. We built the regression analyses models for each data set using the hierarchical approach described in Cohen et al. (2002, chapter 5). First, the model for each data set took a research questions as the first consideration and an initial regression model was built with subject as the random effect variable and the fixed effect variables were HA Experience, Compression Speed and, if applicable, Listening Situation or Signal. Then secondary fixed effect variables, such as trial period, program order and any 2-way and 3-way interactions between the terms were added one at a time. These terms were included in the final model if they improved the residual deviance without appreciably raising (worsening) the Akaike Information Criteria. The residual deviance is a measure of the variance about the regression line and the Akaike Information Criterion is a measure of the goodness of fit of a model that takes the number of fitted parameters into account. Note that the otherwise commonly-reported parameter "amount of variance explained" is not available for binomially-distributed data. The terms for the final regression model for each data set are described under each relevant sub-section in the Results section.

Cohen et al. (2002, p. 161-162) warn against the stepwise-reduction approach, in which all possible terms are included and then taken out if not significant. Their argument is that the stepwise approach is easily biased by confounding and spurious relationships between variables, as well as coincidental type I errors when the sample set is small.

4.3 Results

4.3.1 Achieved Gain for the Experimental HAs

Real Ear Insertion Gain Measurements

The Real-Ear Insertion Gain (REIG) was measured bilaterally using the ISTS speech signal presented at 62 dB SPL. Prescribed and measured gain was compared for each participant (figure 4.3). For both groups of participants, the median measured REIG was within ±3 dB of the insertion gain targets up to 4.2 kHz, with the exception of 0.2 kHz which shows a greater achieved gain than target. This is because at 250 Hz, the NAL-R minus 3 dB often provides negative REIG targets and these are very difficult to achieve in practice. Above 4.2 kHz, REIG rolls off sharply. The target at 8.7 kHz was extrapolated from the 6 kHz target.

The REIG was only measured in program one for each trial. This was to avoid that the participants would start comparing the two programs based on audibility of the ISTS signal during the REIG measurements. In order to control for potential differences in gain between different programs, all settings were measured using detailed coupler measurements.

Coupler Gain Measurements

Detailed coupler gain measurements were made for all HAs employed in the study with all four HA settings (FAST CTLOW, FAST CTMOD, SLOW CTLOW and SLOW CTMOD). To give an example of the performance of the HAs over a wide dynamic range, one HA was randomly selected (participant 0114, left aid) and the coupler gain performance graphs are shown in figure 4.4. For both the FAST and SLOW settings, the CTLOW and CTMOD settings provide the same amount of gain within ±2 dB for inputs at 62 dB SPL and above. Below 62 dB SPL input, the CTLOW continues to increase in gain, whereas the CTMOD operates linearly.

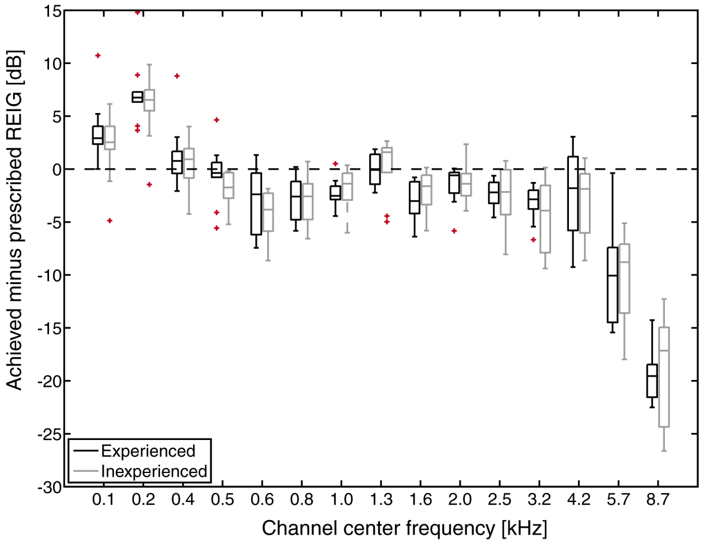


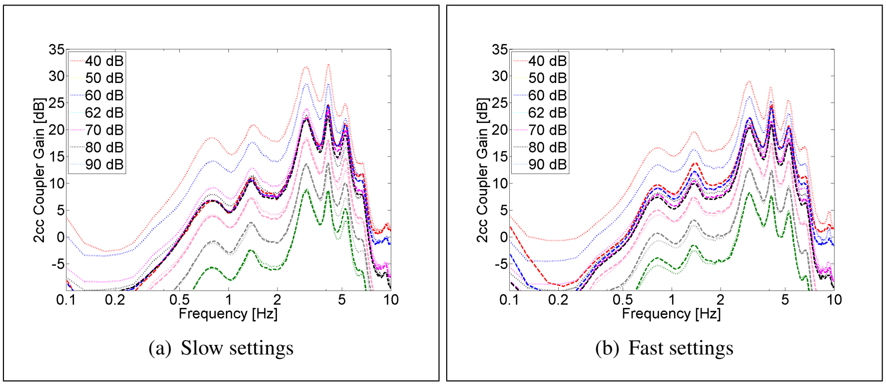
Figure 4.3: Achieved minus prescribed REIG for a 62 dB SPL speech input averaged for all four experimental settings. The analysis was performed in 1/3-octave bands centered at the center frequency of each HA channel. For each box and whisker, the center bars show median, the boxes inter-quartile values, the whiskers maximum and minimum across the participants and the crosses show the outliers.

4.3.2 Field Trial Evaluation Results

Interview Data: Overall CT Preference

Figure 4.5 shows the number of experienced and inexperienced HA users who preferred each CT setting for the SLOW and FAST trials. It can be seen that 5/10 experienced users preferred the CTLOW setting, regardless of whether it was combined with either FAST or SLOW compression. For the inexperienced HA users, 2/10 preferred the CTLOW setting, regardless of whether it was combined with either FAST or SLOW compression. Based on this data, it seems that HA experience is a more important factor in determining CT preference than compression speed. One inexperienced HA user did not make a choice during the FAST trial.

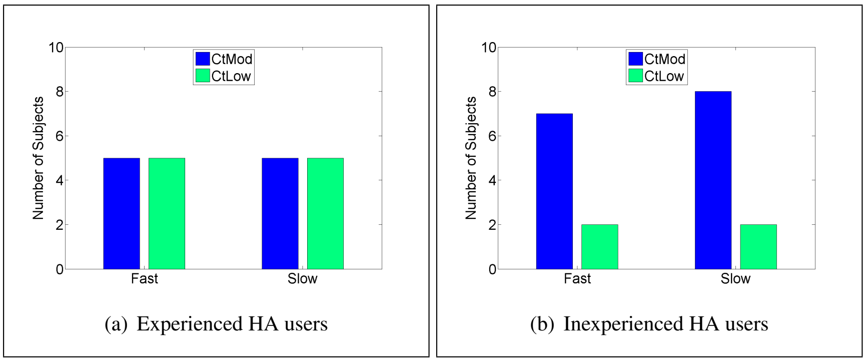
A statistical analysis of the overall CT preference was performed using the GLMM. The explanatory variables considered were HA Experience, Compression



(a) Slow settings

(b) Fast settings

Figure 4.4: Detailed coupler gain measures for a wide range of input levels for one randomly selected HA (participant 0114, left aid). On each sub-figure, the dotted lines represent the CTLOW settings and the dashed lines represent the CTMOD settings. The input signal was ISTS speech presented at input levels between 40 - 90 dB SPL with 10 dB increments, as well as 62 dB SPL. The Fast Fourier Transform (FFT) analysis bandwidth was 43 Hz.



(a) Experienced HA users

(b) Inexperienced HA users

Figure 4.5: Count of number of times either CTLOW or CTMOD were preferred overall for the FAST and SLOW compression speeds. The maximum number of times that a setting can be selected for each question is 10 because there were 10 participants in each group. One inexperienced HA user did not make a choice during the FAST trial.

Speed (repeating), Test Period (first or second trial, repeating) and Program Order (program one either CTMOD or CTLOW) and all possible 2- and 3-way interactions. The final regression model, given in table 4.4 included the terms: Experience, Speed (repeating) and Test Period (repeating). Analysis results showed that HA Experience had a significant effect on preferred CT ($p < 0.05$) but Compression Speed and Test period did not.

Table 4.4: Results of the GLMM regression analysis for overall CT preference. The p-values marked with asterisks were significant at a 5% level.

Model parameter	z value	p
(Intercept)	1.322	0.1861
Experience	-1.989	0.0467*
Compression Speed	0.0543	0.9567
Test Period	1.4117	0.1580

Residual deviance: 59.73
Akaike Information Criterion: 45.73

In order to determine if other participant factors, such as age and gender, have an influence on CT preference, a second GLMM analysis was performed with the following explanatory variables: Experience, Speed, Age, Sex and Hearing Thresholds at 0.5, 1, 2 and 4 kHz, Hearing Thresholds $\times$ Experience, as well as whether the participant had retired. There was no significant relationship between these variables and the preferred CT. Experience was not a significant factor in this analysis because when more explanatory variables are included in a regression analysis, the standard errors inflate and the statistical model loses power.

Interview Data: CT Preference in Specific Listening Situations

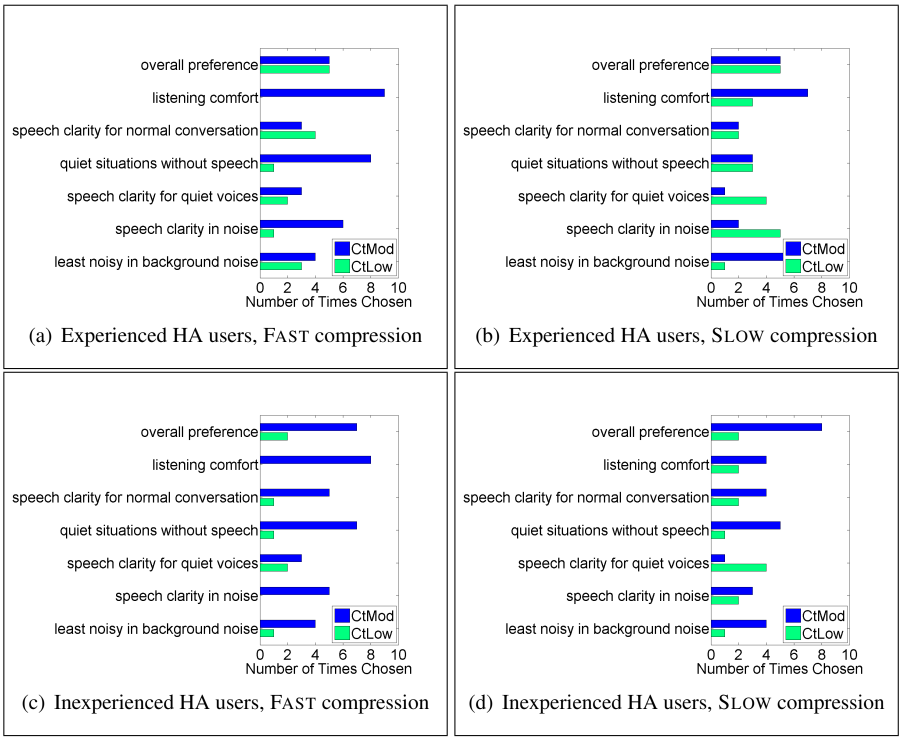
During the interviews at the end of each trial, the participants were asked about their program preferences in a number of specific listening situations, shown as questions (q.) 2-8 in table 4.2. Figure 4.6 shows the number of times that each program was selected for each of the specific questions. In general, CTMOD was preferred most often in almost all of the listening situations for both the experienced and inexperienced HA users and the extent of preference for CTMOD was more marked for the inexperienced HA users. The only situation where participants preferred CTLOW most often was speech clarity for quiet voices (q. 6a) for SLOW compression.

A statistical analysis of the CT preference in each listening situation was performed using the GLMM with subject as a random effect. The fixed effect variables considered in the model were Listening Situation (repeating), HA Experience, Compression Speed (repeating) and all possible 2- and 3-way interactions. The final regression model, given in table 4.5 included the terms: Situation (repeating), Experience, Speed and Situation $\times$ Speed. Analysis results showed that HA experience had a weakly significant effect on preferred CT in a number of listening situations ($p < 0.10$). Situation was a significant variable for quiet situations without speech (q. 5a) and for speech clarity for quiet voices (q. 6a). Compression speed did not have a significant influence on CT preference but there was an interaction between speech clarity for quiet voices (q. 6a) and compression speed ($p < 0.05$).

Table 4.5: Results of the GLMM regression analysis for CT preference in a number of specific listening situations. The p-values marked with asterisks were significant at a 5% level and with dots were weakly significant at the 10% level.

Model parameter	z value	*p*
(Intercept)	-0.1023	0.9185
q: Listening comfort	0.0815	0.9350
q: Speech clarity for normal conversation	1.8365	0.0663 .
q: Quiet situations without speech	2.3720	0.0177 *
q: Speech clarity for quiet voices	2.3954	0.0166 *
q: Speech clarity in noise	1.2783	0.2011
q: Least noisy in background noise	1.0591	0.2896
Hearing aid experience	1.7481	0.0805 .
Compression Speed	0.7382	0.4604
q: Listening comfort × Speed	-0.211	0.9573
q: Speech clarity for normal conversation × Speed	-0.1034	0.9176
q: Quiet situations without speech × Speed	-1.2024	0.2292
q: Speech clarity for quiet voices × Speed	-2.1611	0.0307 *
q: Speech clarity in noise × Speed	-0.8598	0.3899
q: Least noisy in background noise × Speed	-1.7641	0.0777 .

Residual deviance: 179
Akaike Information Criterion: 215



(a) Experienced HA users, FAST compression

(b) Experienced HA users, SLOW compression

(c) Inexperienced HA users, FAST compression

(d) Inexperienced HA users, SLOW compression

Figure 4.6: Count of number of times either CTLOW or CTMOD were preferred by the participants in each of the listening situations in the interview (table 4.2). The last question, noisiest in background noise, was phrased negatively so the responses were inverted. For example, every time CTLOW was selected as “most noisy”, then it was represented on the figure as CTMOD was least noisy.

Interview Data: Additional Questions

The interview included general questions about how satisfied the participants were with each program in each specific listening situations (q. 3b, 4b, 6b, 7b and 8a.) The median response was “good” for all settings to these questions.

Participants were also asked additional questions about the performance of their HAs. The median response to the question about own voice (q. 9) was “good” for all HA settings. For question 11 about acoustic feedback, only 3 experienced HA users (0120, 0121 and 0135) experienced feedback for the SLOW compression and it was audible for both CTLOW and CTMOD settings. Two of these participants selected CTLOW as their overall preferred setting.

For question 10 about microphone noise, 7 of the 20 participants reported hearing microphone noise: 5 participants in the CTLOW setting and 2 in the CTMOD setting. Although microphone noise was more audible for the CTLOW settings, the reporting of microphone noise does not easily explain CT preference because participants 0118 and 0131 preferred CTLOW in spite of the presence of microphone noise. A GLMM analysis did not find any significant relationship between overall preference and the reporting of microphone noise.

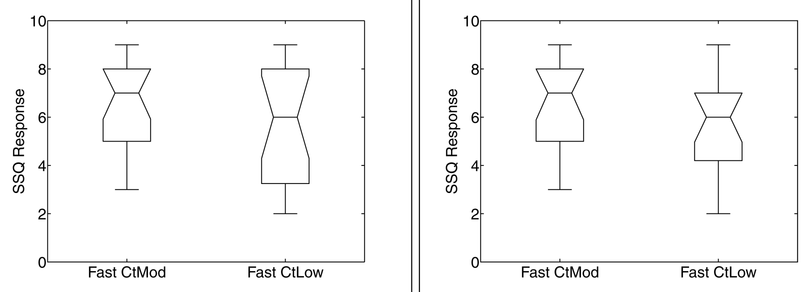
SSQ Responses

For the SLOW trial, there were no obvious differences between SSQ responses for SLOW CTMOD and SLOW CTLOW for any questions. For the FAST trial, only questions "Speech 5" and "Speech 11" showed a difference between FAST CTMOD and FAST CTLOW and the responses to these questions are shown in figure 4.7. This figure shows that median rankings and lower quartiles were lower for FAST CTLOW than FAST CTMOD. The questions when written in full are:

Speech 5 You are talking with one other person. There is continuous background noise, such as a fan or running water. Can you follow what the person says?

Speech 11 You are in conversation with one other person in a room where there are many other people talking. Can you follow what the person you are talking to is saying?

The SSQ data was analysed using the Wilcoxon matched pairs test, which is the non-parametric equivalent of the paired t-test. For FAST and SLOW compression, the responses to each question were compared for CTMOD and CTLOW. The overall significance criterion was $p \leq 0.050$ but after a correction for multiple comparisons based on Tukey's Honesty Significance Test, the criterion for each question was $p \leq 0.013$. There were no significant differences between SLOW CTLOW and SLOW CTMOD. Questions "Speech 5" and "Speech 11" showed significant differences between FAST CTLOW and FAST CTMOD at the 5% significance level.



(a) Question “Speech 5” (b) Question “Speech 11”

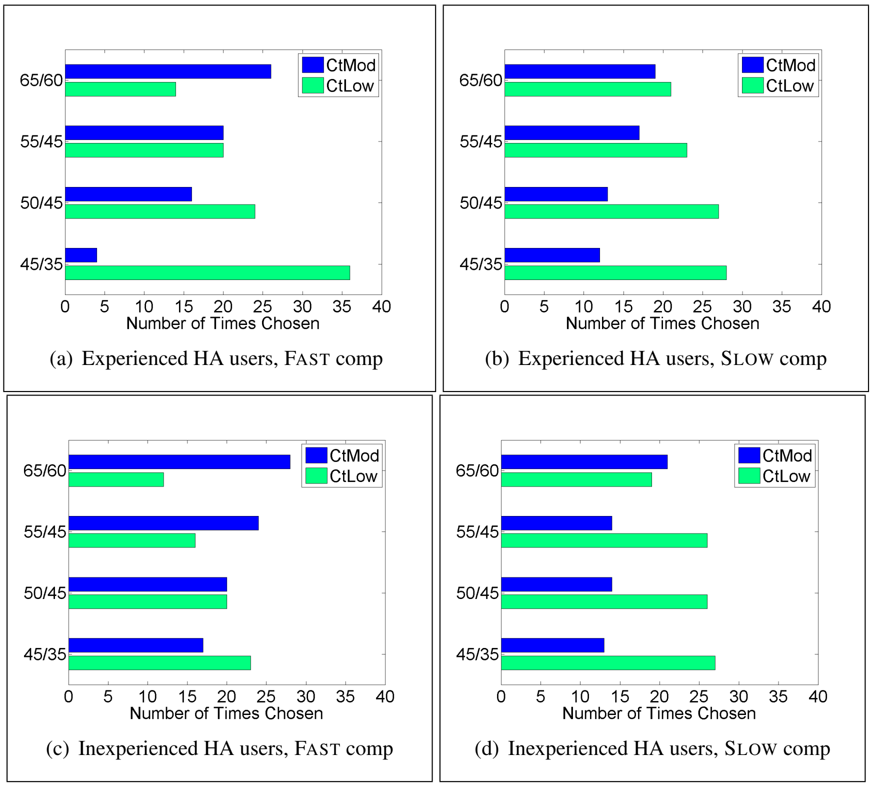
Figure 4.7: Boxplot of the responses to SSQ questions “Speech 5” and “Speech 11”. The SSQ response scale is from 0 to 10, with 0 representing no ability and 10 representing maximum ability. For each box and whisker, the center bars show median, the boxes inter-quartile values, the whiskers maximum and minimum across the participants and the crosses show the outliers.

COSI Responses

Table 4.6 shows the count for CT preference (either CtMod, CtLow or No Preference) for the experienced and inexperienced HA users in the COSI nominated situations. The nominated situations were grouped according to recommendations by Dillon et al. (1999) and only the top six nominated situations were included in table 4.6. Since there was no appreciable difference between CT preference in the Fast and Slow trials, these responses were pooled. The overall pattern to observe is most participants had no CT preference in the nominated situations (82/110 responses). When there was a preference, the participants preferred CtMod more frequently (20/110 compared to 8/110 responses), particularly for the inexperienced participants.

4.3.3 Laboratory Paired Comparisons Results

Figure 4.8 shows the total number of times that CtMod and CtLow were chosen for a given presentation level. Since noise type (Supermarket or Pedestrian Mall) did not have a significant influence on CT preference, the data for both noise types was pooled. It can be observed across all conditions, that in the laboratory the CtLow was preferred more frequently than CtMod (362 times compared to 268). It can also be observed that at the highest presentation level, CtMod was preferred most often, while at the lowest presentation level, CtLow was selected most often. This level dependence is most marked for the Fast speed and for the experienced HA users. Visually, this can be seen in figure 4.8 as the length of the bars are more even for the Slow speed (subfigures on right) and also more even for the inexperienced HA users (subfigures on bottom row).



(a) Experienced HA users, Fast comp

(b) Experienced HA users, Slow comp

(c) Inexperienced HA users, Fast comp

(d) Inexperienced HA users, Slow comp

Figure 4.8: Number of times each CT was preferred in the laboratory paired comparisons for each presentation level, pooled for both noise types (supermarket and pedestrian mall). The maximum number of times that a CT could be selected for stimulus was 40 (10 participants $\times$ 2 noise types $\times$ 2 repeats).

A GLMM analysis was performed for the preferred CT in the laboratory. The explanatory variables considered were: Presentation Level (repeating) and Noise Type (repeating), Experience, Speed (repeating) plus 2-way and 3-way interactions. The final regression model, given in table 4.7 included the terms: Level (repeating), Noise Type (repeating), Experience, Speed (repeating), Experience $\times$ Level, Speed $\times$ Level and Experience $\times$ Speed $\times$ Level. The level effects were highly significant at the 0.1% level ($p < 0.001$). The interaction between Experience and Speed, and Experience and Level, as well as the 3-way interaction between Experience, Speed and Level were significant at the 5% level ($p < 0.05$).

Correlation Between Paired Comparisons in Laboratory and Overall Preference in the Field

For each participant, the number of times they selected CTLOW and CTMOD was pooled across all laboratory signals. This number for each participant and each compression speed was correlated with their overall preference in the field using a GLMM analysis. There was no significant relationship between the number of times each CT was preferred in the lab and the overall preference in the field.

Speech Intelligibility

Figure 4.9 shows mean speech intelligibility scores in each condition. There are no obvious differences between speech intelligibility scores for the four experimental HA settings. At the 56 dB presentation level, the inexperienced HA users performed poorer when unaided than when wearing the experimental HAs.

For the statistical analysis, the speech intelligibility scores were transformed using Rationalized Arcsine Units (RAU) as described in Studebaker (1985). All RAU speech scores with the experimental HAs were analysed using a General Linear Model (not the same as the Generalized Linear Model) with presentation level (71/65 or 56/50), participant group (experienced or inexperienced HA user), compression speed (fast or slow) and CT (CtMod or CtLow) and the interaction between compression speed and CT. There were significant presentation level effects ($p < 0.001$) and participant group effects ($p < 0.001$) but no compression speed or CT effects or interactions.

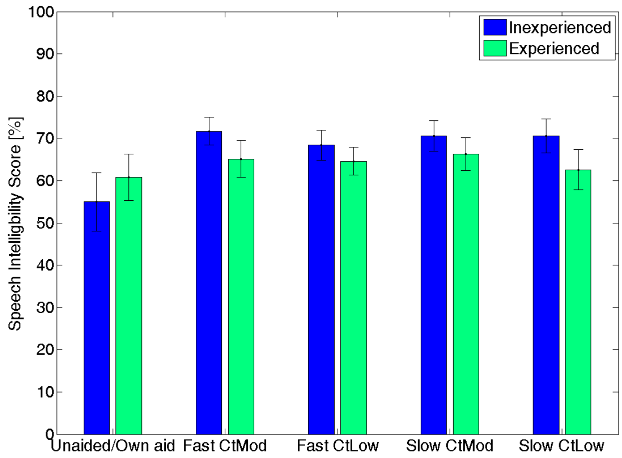
Table 4.6: The number of times each setting was preferred in the top six COSI nominated situations. The data for the FAST and SLOW settings was pooled.

Listening	Experienced			Inexperienced		
situation	CTMOD	No Pref	CTLOW	CTMOD	No Pref	CTLOW
Television/radio at normal volume	4	6	0	4	11	1
Conversation with one or two in quiet	2	7	1	3	7	2
Conversation with group in noise	1	12	1	2	6	0
Listening in church or meeting	1	5	2	2	10	0
Conversation with one or two in noise	0	4	0	0	6	0
Conversation with group in quiet	0	7	1	1	1	0
Sum	8	41	5	12	41	3

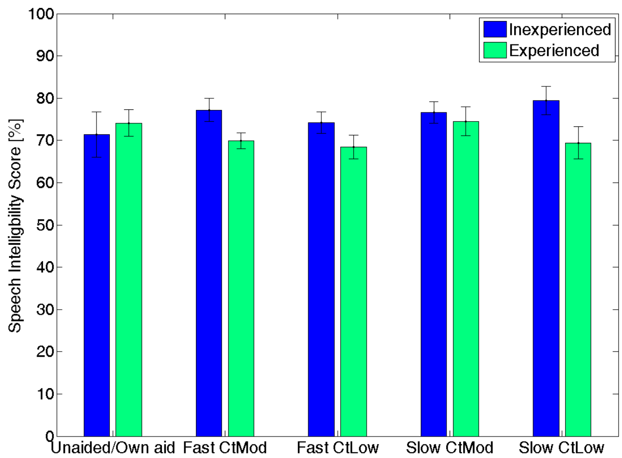
Table 4.7: Results of the GLMM regression analysis for CT preference in the laboratory. The p-values marked with asterisks were significant at a 5% level, double asterisks at the 1% level and triple asterisks at the 0.1%.

Model parameter	z value	*p*
(Intercept)	1.778	0.07545
Level: 55/45	-1.303	0.19249
Level: 50/45	-2.226	0.02603*
Level: 45/35	-4.489	0.0000***
Noise Type	-0.559	0.57601
Experience	0.446	0.65565
Speed	-1.486	0.13717
Experience × Level: 55/45	0.359	0.71926
Experience × Level: 50/45	0.301	0.76315
Experience × Level: 45/35	2.099	0.03585*
Speed × Level: 55/45	0.635	0.52571
Speed × Level: 50/45	0.635	0.52571
Speed × Level: 45/35	2.823	0.00475**
Experience × Speed × Level: 65/60	-0.062	0.95031
Experience × Speed × Level: 55/45	-1.160	0.24598
Experience × Speed × Level: 50/45	-0.500	0.61727
Experience × Speed × Level: 50/45	-2.245	0.02478*

Residual deviance: 786.3
Akaike Information Criterion: 862



(a) Speech 56 dB SPL, Noise 50 dB SPL



(b) Speech 71 dB SPL, Noise 65 dB SPL

Figure 4.9: Mean speech intelligibility scores for each HA setting used in the field trial. As a control, the experienced HA users were tested wearing their own HAs and the inexperienced HA users were tested unaided. The error bars represent the standard errors of the mean. Speech intelligibility was tested using the Swedish phonemically balanced word lists (Magnusson, 1995) with a unmodulated speech-shaped background noise (Dreschler et al, Track 1). In the top sub-figure, the speech was presented at 56 dB SPL and the noise was presented at 50 dB SPL. In the bottom sub-figure, the speech was presented at 71 dB SPL and the noise was presented at 65 dB SPL.

4.4 Discussion

Hearing aid CT preference was investigated in the field and in the lab with both experienced and inexperienced HA users. Field trial results showed that experienced HA users were significantly more likely to prefer the CTLOW setting (5/10 participants) compared with the inexperienced HA users (2/10 participants). The overall CT preference did not depend on whether compression speed was FAST or SLOW. For the specific listening situations in the field, in most situations, the subjects preferred CTMOD with the exception of "speech clarity for quiet voices" for the SLOW compression speed and There was a significant interaction between the question "speech clarity for quiet voices" and compression speed. For speech clarity in noise, the SSQ responses showed that participants rated FAST CTMOD better than FAST CTLOW. Finally, in the laboratory, CT preference was highly significantly level dependent and this level dependence was more apparent for the experienced HA users and for FAST compression.

4.4.1 Achieved Gain

A requirement for any HA investigation is that the HAs are set and perform in a manner as prescribed. The current results show that achieved insertion gain for a ISTS speech input at 62 dB SPL was close to the prescribed targets, up to and including 4.2 kHz. The participants also reported a high level of satisfaction with the performance of their HAs, supporting the claim that the participants were fitted properly. At medium- and high-input levels, the coupler gain measurements also demonstrate that the four HA settings prescribed the same gain within ± 2 dB for a realistic speech signal. Additionally, the average speech intelligibility scores showed no significant differences between the four HA settings, also supporting that the gain at medium-input levels was similar. At low-input levels, the CTLOW coupler gain was higher than the CTMOD settings, as prescribed.

4.4.2 Influence of HA Experience on CT preference

The first research question considered whether CT preference was dependent on HA experience. We found a statistically-significant difference in CT preference between the inexperienced and experienced HA users, altogether for the overall preference and preference for the laboratory paired comparisons at low input levels, as well as a weakly significant difference for the specific interview questions in the field trial. There were only 10 participants in each group, so this provides only preliminary evidence that HA experience affects CT preference. There might be (at least) three possible explanations about how HA users might acclimatise to a lower CT with HA experience.

The first explanation about how HA users might acclimatise to a lower CT, is that the new HA user might need time following non-linear HA provision to acclimatise to the extra audibility for environmental sounds at low-input levels. Following a gradually-acquired hearing loss, HA provision gives re-audibility to many new sounds that the HA user has not heard well for years, e.g., computer fan. The CTLOW setting provides greater audibility for these low-level environmental sounds than the CTMOD setting. It may take time for the new HA user to learn to identify these sounds and then "filter out" unimportant sounds, and while they are learning to do this, they may prefer a HA setting that provides less audibility for low-input levels. While this explanation sounds plausible, there is no hard evidence that either newly re-audible low-level stimuli have a particular attentional salience, or that HA users become better at filtering out irrelevant auditory stimuli with HA experience.

The second explanation about how new HA users might acclimatise to a lower CT is due to level-dependent changes in loudness perception following HA provision. However, evidence from other studies suggest that if loudness acclimatisation is level-dependent, then it occurs at medium- and high-input levels. For instance, following HA fitting, there is evidence of changes in loudness perception at medium- and high-input levels and not low-input levels (Olsen et al., 1999; Philibert et al., 2002, 2005; Munro and Trotter, 2006; Keidser et al., 2008), as well as changes in speech intelligibility scores at high-input levels (Gatehouse, 1989; Connor, 1999; Munro and Lutman, 2003). Another argument against this explanation is changes to gain preference following HA provision have a longer time course ($>$ 1 year) (Keidser et al., 2008) than level-dependent changes for loudness or speech intelligibility (2-3 months) (Keidser et al., 2008; Philibert et al., 2005).

The third explanation about how new HA users might acclimatisation to a lower CT, is that they require HA experience to get used to potential artifacts from non-linear HA processing. This explanation is not so likely because we only used a 2:1 compression ratio and previous research shows that such a low CR does not appreciably alter sound quality (Neuman et al., 1998).

Finally, we did not find a relationship between CT preference, experience and degree of hearing loss, which would otherwise suggest that HA users with a greater degree of hearing loss exhibit a greater acclimatisation effect. This is in contrast with findings from Keidser et al. (2008), who found that gain acclimatisation effect was dependent on the degree of hearing loss. That is, individuals with severe hearing losses exhibit a greater gain acclimatisation effect than individuals with mild hearing loss. However, they current study did not use participants with as great as range of hearing losses as those in Keidser et al. (2008) nor did the current study use as many participants as Keidser et al. (2008).

4.4.3 Influence of Compression Speed on CT Preference

The second main research question was whether the preferred CT is influenced by compression speed. We found in the field trial that the overall CT preference did not depend on compression speed and this was not consistent with the findings in chapter 3, in which HA users were more likely to select CTLOW when combined with long release times. However, for the interview question "speech clarity for quiet voices", there was a significant speed effect, consistent with chapter 3. In other words, the participants preferred CTMOD for quiet voices when combined with FAST compression but CTLOW when combined with SLOW compression.

Additionally, for the SSQ questionnaire, the ratings for two situations with speech in noise were poorer for CTLOW than CTMOD when combined with FAST compression. There was no difference between CTLOW and CTMOD when combined with SLOW compression. It is worth noting that chapter 3 predominantly used signals recorded in situations with moderate amounts of background noise present (supermarket and pedestrian mall). The relationship between preferred CT and compression speed in noisy situations is probably due to that when the CT is lowered, more of the signal is compressed and fast-acting compression is associated with a worsening of the SNR, at least when the input signal has a positive SNR (Naylor and Johannesson, 2009; Souza et al., 2006). When slow-acting compression is used, compression does not have such a marked effect on the SNR and this is probably why HA users are more likely to accept a low CT in noisy situations, when combined with slow-acting compression.

The laboratory paired comparison procedure in the current study also had a good agreement with earlier results in chapter 3. There was a significant interaction between preferred CT, compression speed and presentation level. At the highest signal presentation level used in the current lab study (speech 65 dB and noise 60 dB), the level was similar to the signals used in chapter 3. At this presentation level, the findings were similar to those in chapter 3. Namely, CTMOD was preferred for FAST compression but the preference for CTMOD and CTLOW was equivocal for SLOW compression. This consistency is in spite of numerous small differences in the HA specifications between the two studies, including a smaller difference between CTMOD and CTLOW in the current study, different time constants and different vent sizes.

4.4.4 Influence of Listening Environment of CT Preference

The final research question was whether CT preference depended on the listening situation. In the laboratory paired comparisons, CT preference was predominantly determined by signal presentation level; the lower the presentation level, the more likely it was that participants preferred CTLOW. The level dependence was more marked for FAST compression and experienced HA users.

In the field trial in most listening situations, CTMOD was selected most often. This result is consistent with earlier findings from the NAL laboratories (Barker and Dillon, 1999; Barker et al., 2001; Dillon et al., 1998). However, CTLOW was preferred when combined with SLOW compression for the situations "quiet or distant speech". It is interesting to note that the participants preferred CTMOD in environments without speech because it suggests that CT preference is not just dependent on input level but also on the content of the signal.

In the COSI nominated listening situations, most participants did not have a preference for CtMod and CtLow. The comments from the subjects suggested that it was difficult for them to hear a difference. Similarly, the SSQ responses showed only significant differences for 2 out of the 50 questions. Together this suggests that varying the CT settings do not give a large perceptual difference between HA settings.

As a general note, in the lab study there was a preference for CtLow but in the field study, the overall preference and the preference for the majority of listening situations was for CtMod. A lab study does not have the same issues with competing attentional demands and/or listening comfort as a field study. In the field trial, many of the participants commented that they could hear more things with the CtLow setting but found the CtMod setting more comfortable. In the lab, listeners only have one task to concentrate on and that is evaluating the signal. In the real world, individuals may not be attending to an auditory signal at all times and if they are, they may not appreciate the audibility for many other signals, particularly new HA users who are experiencing a lot of new sounds. In comparison to chapter 3 which found in the laboratory that CT preference depends on the compression speed, the current results suggest that in real-world listening situations, overall CT preference depends more on HA experience. Additionally, the participant comments suggest overall CT preference depends on how much the individual appreciates the extra audibility provided by low CT.

4.4.5 Conclusion

In the field trial, participant responses to the SSQ and COSI questionnaires did not show great differences in ratings for CT settings. When the participants had to chose between CT settings in the interview, the current results were consistent with previous findings by the NAL group (Barker and Dillon, 1999; Barker et al., 2001; Dillon et al., 1998) that the majority of HA users preferred moderate-level CT rather than low-level CT both overall and in most listening situations. But for specific listening situations, both the field trial data and laboratory paired comparisons indicated that low-level CT are preferred for speech clarity at low-input levels, when combined with slow-acting compression. The current results also found that overall CT preference was related to HA experience, and this supports including adjustment of CT in gain acclimatisation management in HA fitting.

Acknowledgments

Thanks to engineer Niklas Bergman for recording and mixing the signals for the laboratory paired comparisons and making an interface to control the signal presentation. Niklas also provided technical assistance for the real-ear and coupler measurements.

Thanks also to the 20 participants who participated in the current study and to Avesina Hörselrehab clinic for assisting us in recruiting test participants.

5

General Discussion

5.1 Summary of research findings

One of the consequences of hearing loss is that hearing-impaired individuals have a reduced audibility for sounds, and in particular soft sounds (both speech and environmental sounds). This is usually managed audiologically by fitting hearing aids. The central question in this overall project is to determine under which circumstances (if any) “soft sounds” in the environment, should be amplified to audibility? This question arose because while hearing aid (HA) compression acts to makes sounds in our environment audible and comfortable at a wide range of input levels, there is a lack of empirical information about how much gain at low-input levels HA users prefer. One of the most important parameters for determining HA gain at low input levels is the compression threshold (CT) because when CT is lowered, the gain at low input level increases, as long as gain is fixed at medium- and high-input levels. Previously, the National Acoustic Laboratory (NAL) group in Australia investigated compression threshold (CT) preference in a series of field trials using a single-channel, fast-acting compression HA (Dillon et al., 1998; Barker and Dillon, 1999; Barker et al., 2001). They found that the majority of HA users who participated in the studies preferred a moderate CT (∼65 dB SPL) over a low CT (40-57 dB SPL). This was a counterintuitive result because the low CT should give better audibility for soft sounds than the moderate CT. An implicit aim of the current project was to follow-up on the previous field studies by NAL and investigate the factors that could potentially influence the preferred CT. This was carried out in a series of pilot experiments with normal-hearing participants (chapter 2), a laboratory paired-comparison experiment with hearing-impaired individuals (chapter 3), and a field study with hearing-impaired individuals supplemented with laboratory paired comparisons (chapter 4).

The research findings from the current project are summarized in an Ishikawa diagram in figure 5.1. An Ishikawa diagram is also known as a fishbone diagram and it is commonly used within project management to illustrate cause and effect relationships. The effect that we are considering is "CT preference" and this is shown at the head of the "fish skeleton". The major factors are shown as the major bones on the fish and these were Hearing Aid, Method, Listener and Signal. The specific factors are shown as the minor bones branching from the major bones. The minor and major factors will be discussed in the rest of this section.

5.1.1 Influence of other HA settings on CT Preference

A major factor (or fishbone on figure 5.1) investigated in this project was the other hearing aid parameters (called "Hearing aid" on figure 5.1). The main hearing aid factor of interest was the influence of "Compression speed' on preferred CT. The prediction was that when slow-acting compression is used, HA users will be more likely to accept a low CT because many of the potentially negative side effects of compression are ameliorated when the release time is lengthened (e.g., lowering of the signal-to-noise ratio, envelope depth, etc). In the experiment described in chapter 3, real-life environmental stimuli (e.g. living room, supermarket) were processed offline using a 15-channel compressor model with a fixed 2:1 compression ratio. The compressed stimuli were presented to the participants via the Direct Audio Input (DAI) of bilateral hearing aids fit linearly according to the National Acoustic Laboratories - Revised (NAL-R) rationale (Byrne and Dillon, 1986). Twelve experienced HA users with moderate, sloping hearing losses made paired comparisons of the stimuli processed with a combination of two CTs and three Release Times (RT=40, 400 and 4000 ms). The two CT settings were: (i) CTMOD, the level of normal speech in each channel and (ii) CTLOW, 30 dB lower than (i) in each channel. The finding was that preference for CT was strongly influenced by the release time; as release time increased, the extent of preference for the low CT also increased. Many of the participants also mentioned informally that they made their paired comparison selections based on the level of the background noise. The findings suggest that the HA gain provided at low input levels should depend on the HA compression release time.

The finding of the influence of compression speed on preferred CT was fol-

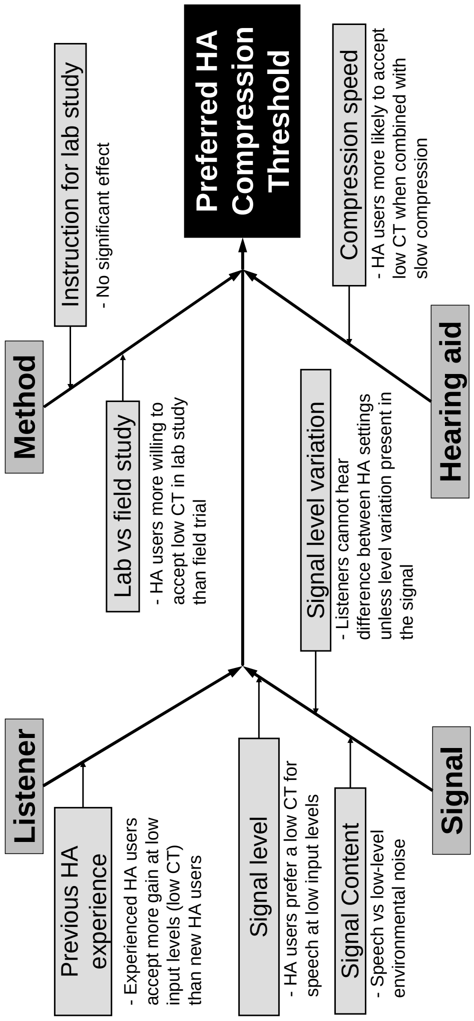


Figure 5.1: Ishikawa (fishbone) diagram of factors that influence the preferred hearing aid compression threshold (CT). The major bones represent the major categories of factors and the smaller bones represent the specific categories of factors.

lowed up with a field trial experiment described in chapter 4. The participants were 20 hearing-impaired individuals (10 new and 10 experienced HA users) with mild-moderate hearing losses. The participants wore bilateral 15-channel experimental HAs. They were fitted according to the NAL-R targets prescribed for a normal speech input (62 dB SPL ANSI S3.5, 1997). The participants compared two HA settings, CTMOD and CTLOW, in their own normal daily listening environments during two trial periods. In one trial period, the HAs were set with fast-acting compression and in the other trial period, slow-acting compression. The assignment of CT settings between programs and compression speeds between trial periods was counter-balanced and double-blind. At the end of each trial period, the participants were interviewed about their CT preference. The results showed that *overall* CT preference was not influenced by compression speed, but when the participants were asked about CT preference in specific listening situations, there was a significant interaction between speed and the listening situation "speech clarity in quiet". In other words, when listening to speech in quiet, the participants preferred CTMOD with fast-acting compression and CTLOW with slow-acting compression. Also the SSQ questionnaire revealed that in some speech in noise situations, during the fast-acting compression trials, the ratings were significantly poorer for CTLOW than CTMOD. These results also supported the conclusion that it is not desirable to combine low CTs with fast-acting compression.

Finally, the field trial described in chapter 4 was supplemented with a laboratory paired comparison task using the same experimental hearing aids. The stimuli were running speech combined with recordings of noise from a supermarket and from a pedestrian mall. It was found that at medium presentation levels (speech 65 dB SPL, noise 60 dB SPL) the compression speed influences the preferred CT. Specifically, for fast-acting compression, the participants preferred CTMOD over CTLOW and for slow-acting compression, the preference was equivocal. These result was similar to the findings in chapter 3. At low signal presentation levels (speech 45 dB, noise 35 dB), CTLOW was preferred most often, regardless of the compression speed. These paired comparisons results lend further support to the claim that low CTs are best combined with slow-acting compression.

5.1.2 Influence of Listener on CT Preference

Another major factor explored in this project was the "Listener" (figure 5.1). In the field study described in chapter 4, there were two groups of test participants: 10 new and 10 experienced HA users. The finding was that the experienced HA users were more likely to prefer CTLOW than the new HA users. This was found for overall preference, as well as a data trend for preference in specific listening situations and also responses to the COSI questionnaire. This provides evidence for the hypothesis that HA users might acclimatise to a lower CT.

The main studies in chapters 3 and 4 used test participants with mild-to-moderate, sloping hearing losses, which are the most common degree and configuration of hearing loss (Wilson et al., 1999). It is important to note that the research findings from this project may not apply to HA users with other configurations of hearing loss (e.g., reverse slope) or a more severe degree of hearing loss.

5.1.3 Influence of Signal on CT Preference

A major factor explored in this project was "Signal". The effect of signal level was investigated in the paired comparison experiment described in chapter 4. The stimuli used were running speech presented at various levels from 45-60 dB SPL with noises from a supermarket and a pedestrian mall presented at either +10 or +5 dB signal-to-noise ratio. The result for these paired comparisons was that CT preference was predominantly determined by signal presentation level; the lower the presentation level, the more likely it was that participants preferred CTLOW. The level dependence was more marked for fast-acting compression and experienced HA users. However, data from the paired comparisons in chapter 3 indicated that if the signal level is too low (i.e., signals recorded in living room), then microphone noise becomes audible and this influences the preference such that HA users prefer a moderate CT.

The findings from the field trial suggested that the "Content" of the signal as well as the signal level might influence the preference. Most participants preferred CTLOW with slow-acting compression for "quiet or distant speech", but they preferred CTMOD (also combined with slow-acting compression) for "quiet situations without speech". The limitation with this result is that the data was from the field trial, where it was not possible to control the signal levels.

Finally, it was found in the third pilot study in chapter 2, that the test participants could not hear a difference between HA settings, when the signal level was either too high or lacked level variation.

5.1.4 Influence of Research Method on CT Preference

A factor in this project is the research methodology employed. The third pilot experiment in Chapter 2 investigated the feasibility of using instructions to direct the participant's attention while the participants make paired comparisons of hearing aid processed stimuli. Ten normally-hearing participants made paired comparisons of everyday stimuli processed with fast-acting compression. The participants compared CTMOD and CTLOW. There were two instructions: (a) listen for a particular target sound within the signal, or (b) imagine that you are concentrating on something other than the sound. There was a slight trend in the data, that the participants were more likely to select CTLOW when instructed to "not listen" than when instructed to "listen". As there was a (non-significant) trend observed in the pilot study that instruction influences CT preference, the two instructions was included in the next paired comparison experiment with hearing-impaired participants described in chapter 3. In this experiment, instruction did not influence the preference for CT or RT. So it seems that instruction in this laboratory study did not have the effect that was intended.

As a general note from the experiment in chapter 4 that combined a field study with a lab study, the lab study showed a preference for CTLOW but in the field study, the overall preference and the preference for the majority of listening situations was for CTMOD. A lab study does not have the same issues with competing attentional demands and/or listening comfort as a field study. In the field trial, many of the participants commented that they could hear more things with the CTLOW setting but found the CTMOD setting more comfortable and comfort may be more important in the field than in the lab.

5.2 Clinical implications

Overall, the results were not strongly in favour of either a low or moderate CT. In the field trial (chapter 4), the preference for CTLow and CTMod was equivocal for the experienced HA users. The SSQ questionnaire used in the field trial did not reveal large differences between CT settings. In the COSI questionnaire, which probes into listening situations that the HA user themselves find important, the vast majority of responses were "No preference between settings". Finally, speech intelligibility testing in chapter 4 did not find any differences between CT settings.
The major clinical recommendations are listed below.

1. The combination of fast-acting compression with low CTs are not often preferred by HA users, particularly in listening situations with speech in background noise.

2. There may be some advantage to providing a low CT with slow-acting compression in situations with low level speech, but not situations with low-level noise. Since most modern HA have speech detection algorithms, it would be possible to apply low CTs when speech is detected and increase CT when noise is detected.

3. New HA users do not appreciate low CTs as much as experienced HA users, so this would suggest that new HA users should be fit with moderate CTs and implicitly, reduced gain for low input levels. However there is very little information about how much experience new HA users require with HA amplification before they are ready to have gain increases. It is also unclear whether it would be beneficial that the HA should automatically increase gain with time, or if the HA dispenser or HA user should increase the gain of the HA with time.

Before making clinical recommendations for gain at low input levels, it is important to recognise that the investigations performed in this project were carried out with participants with mild-moderate, sloping sensorineural hearing losses. It is not clear if the findings would apply to other HA users with other hearing loss types and configurations or more severe hearing loss. Additionally, the experiments used HAs with all other adaptive features disabled. It is not clear if results would apply to HAs with either noise reduction or directional microphones, which have the purpose to increase the signal-to-noise ratio. Finally, if more open earpieces were used instead of the standard acrylic earmoulds, the results would probably be even more equivocal, as the influence of more unprocessed direct sound entering the ear canal become greater.

5.3 Suggestions for Future Research

One of the most interesting findings in this study was that CT preference is related to HA experience, i.e., experienced HA users preferred low CTs to a greater degree than new HA users. Although, this finding is only preliminary because there were only 10 new and 10 experienced HA users in the field trial. The effect is nonetheless consistent with the "gain acclimatisation effect" described by Keidser et al. (2008). So far, there has been little investigation into the "gain acclimatisation" effect and the time course, magnitude and degree of individual variability are unclear. It is also unclear if the effect is level-dependent and if the gain-acclimatisation effect is related to changes in loudness perception or speech perception following HA provision (see Convery et al., 2005; Keidser et al., 2008, for review). It would also be interesting to know if the gain-acclimatisation effect can be augmented with auditory training, for example, with programs such as the Listening and Communication Enhancement (LACE) program (Sweetow and Sabes, 2006). Until more is known about the gain acclimatisation effect, it is difficult to know how to manage the effect audiologically.

References

ANSI S3.5 (1997). American national standard methods for the calculation of the Speech Intelligibility Index. Americal National Standards Institute Inc, New York.

Arlinger, S., Gatehouse, S., Bentler, R. A., Byrne, D., Cox, R. M., Dirks, D. D., Humes, L., Neuman, A., Ponton, C., Robinson, K., Silman, S., Summerfield, A. Q., Turner, C. W., Tyler, R. S., and Willott, J. F. (1996). Report of the Eriksholm Workshop on auditory deprivation and acclimatization. *Ear Hear*, *17*, 87S–98S.

ASHA (1990). Guidelines for screening for hearing impairment and middle-ear disorders. American Speech—Language—Hearing Association (ASHA).

Barker, C. and Dillon, H. (1999). Client preferences for compression threshold in single-channel wide dynamic range compression hearing aids. *Ear Hear*, *20*, 127–139.

Barker, C., Dillon, H., and Newall, P. (2001). Fitting low ratio compression to people with severe and profound hearing losses. *Ear Hear*, *22*, 130–141.

Beaman, C. P. (2005). Auditory distraction from low-intensity noise: A review of the consequences for learning and workplace environments. *Appl Cognitive Psych*, *19*(8), 1041–1064.

Bentler, R. and Nelson, J. (1997). Assessing the effect of varying release time in a 2-channel agc hearing aid. *Am J Audiol*, *6*, 43–51.

Bentler, R. A., Niebuhr, D. P., Getta, J. P., and Anderson, C. V. (1993). Longitudinal study of hearing aid effectiveness. II: Subjective measures. *J Speech Hear Res*, *36*(4), 820–831.

Bentler, R. A. and Pavlovic, C. V. (1989). Transfer functions and correction factors used in hearing aid evaluation and research. *Ear Hear*, *10*(1), 58–63.

Bjerg, A. P. and Larsen, J. N. (2006). Recording of natural sounds for hearing aid measurements and fitting. Technical report, Acoustic Technology, Ørsted*DTU, Technical University of Denmark.

Bor, S., Souza, P., and Wright, R. (2008). Multichannel compression: effects of reduced spectral contrast on vowel identification. *J Speech Lang Hear Res*, *51*(5), 1315–1327.

Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, *39*, 324–345.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 443–446.

Brewer, S. (2005). NAL-NL1 software. version 1.4 (dll ver 1.5) experimenter edition. Software.

Buus, S. and Florentine, M. (2001). Growth of loudness in listeners with cochlear hearing losses: recruitment reconsidered. *J Assoc Res Otolaryngol*, *3*, 120–139.

Byrne, D. and Cotton, S. (1988). Evaluation of the National Acoustic Laboratories' new hearing aid selection procedure. *J Speech Hear Res*, *31*, 178–186.

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 Hear*, *7*(4), 257–265.

Byrne, D., Dillon, H., Ching, T., Katsch, R., and Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisions with other procedures. *J Am Acad Audiol*, *12*, 37–51.

Byrne, D., Parkinson, A., and Newall, P. (1990). Hearing aid gain and frequency response requirements for the severely/profoundly hearing impaired. *Ear Hear*, *11*(1), 40–49.

Ching, T. Y., Dillon, H., and Byrne, D. (1998). Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am*, *103*(2), 1128–1140.

Choisel, S. and Wickelmaier, F. (2007). Evaluation of multichannel reproduced sound: scaling auditory attributes underlying listener preference. *J Acoust Soc Am*, *121*, 388–400.

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioural sciences* (Third Edition ed.). Laurence Erlbaum Associates.

Connor, H. (1999). The acclimatisation effect: Loudness perception and speech recognition. Master's thesis, The University of Auckland, New Zealand.

Connor, H. and Poulsen, T. (2007). Influence of listener task on ratings of pleasantness for everyday sounds. In Dau, T., Büchholz, J., Harte, J., and Christiansen, T. (Eds.), *Auditory Signal Processing in Hearing Impaired Listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR).*

Convery, E., Keidser, G., and Dillon, H. (2005). A review and analysis: Does amplification experience have an effect on preferred gain over time? *Austr NZ J Audiol*, *27*, 18–32.

Cornelisse, L. E., Seewald, R. C., and Jamieson, D. G. (1995). The input/output formula: A theoretical approach to the fitting of personal amplification devices. *J Acoust Soc Am*, *97*(3), 1854–1864.

Cox, R. M. and Alexander, G. C. (1992). Maturation of hearing aid benefit: objective and subjective measurements. *Ear Hear*, *13*(3), 131–141.

Davis, A. C. (1989). The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *Int J Epidemiol*, *18*(4), 911–917.

Dillon, H. (2001). *Hearing Aids*. Boomerang Press, Sydney.

Dillon, H., Birtles, G., and Lovegrove, R. (1999). Measuring the outcomes of a national rehabilitation program: Normative data for the Client Oriented Scale of Improvement (COSI) and the Hearing Aid User's Questionnaire (HAUQ). *J Am Acad Audiol*, *10*, 67–89.

Dillon, H., James, A., and Ginis, J. (1997). Client Oriented Scale of Improvement (COSI) and its relationship to several other measures of benefit and satisfaction provided by hearing aids. *J Am Acad Audiol*, *8*(1), 27–43.

Dillon, H., Storey, L., Grant, F., Phillips, A., Skelt, L., Mavrias, G., Woytowych, W., and Walsh, M. (1998). Preferred compression threshold with 2:1 wide dynamic range compression in everyday environments. *Aus J Audiol*, *20*, 33–34.

Dreschler, W., Verschuure, H., Ludvigsen, C., and Westermann, S. (2001). Icra noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. international collegium for rehabilitative audiology. *Audiology*, *40*, 148–157.

Eisenberg, L. S., Dirks, D. D., and Gornbein, J. A. (1997). Subjective judgments of speech clarity measured by paired comparisons and category rating. *Ear Hear*, *18*(4), 294–306.

Ellermeier, W., Mader, M., and Daniel, P. (2004). Scaling the unpleasantness of sounds according to the btl model: Ratio-scale representation and psychoacoustical analysis. *Acta Acust United Ac*, *90*, 101–107.

Fowler, E. (1936). A method for the early detection of otosclerosis. *Arch Otolaryngol*, *24*, 731–741.

Gatehouse, S. (1989). Apparent auditory deprivation effects of late onset: The role of presentation level. *J Acoust Soc Am*, *86*(6), 2103–2106.

Gatehouse, S. (1992). The time course and magnitude of perceptual acclimatization to frequency responses: evidence from monaural fitting of hearing aids. *J Acoust Soc Am*, *92*(3), 1258–1268.

Gatehouse, S. (1993). Role of perceptual acclimatization in the selection of frequency responses for hearing aids. *J Am Acad Audiol*, *4*(5), 296–306.

Gatehouse, S., Naylor, G., and Elberling, C. (2006). Linear and nonlinear hearing aid fittings–1. Patterns of benefit. *Int J Audiol*, *45*(3), 130–152.

Gatehouse, S. and Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *Int J Audiol*, *43*(2), 85–99.

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

Haynes, B. (1999). Can it work? does it work? is it worth it? the testing of healthcare interventions is evolving. *British Medical Journal*, *319*, 652–653. AA2171.

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. *J Speech Lang Hear Res*, *51*(2), 471–484.

Hickson, L. (1994). Compression amplification in hearing aids. *Am J Audiol*, *3*, 51–65.

Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (2009). Development and analysis of an International Speech Test Signal (ISTS). *Int J Audiol*, *submitted*.

Horwitz, A. R. and Turner, C. W. (1997). The time course of hearing aid benefit. *Ear Hear*, *18*(1), 1–11.

Humes, L. E. and Wilson, D. L. (2003). An examination of changes in hearing-aid performance and benefit in the elderly over a 3-year period of hearing-aid use. *J Speech Lang Hear Res*, *46*(1), 137–145.

Humes, L. E., Wilson, D. L., Barlow, N. N., and Garner, C. (2002). Changes in hearing-aid benefit following 1 or 2 years of hearing-aid use by older adults. *J Speech Lang Hear Res*, *45*(4), 772–782.

IEC 60118-2 (1983). Hearing Aids. Part 2: Hearing aids with automatic gain control circuits. International Electrotechnical Commission, (Geneva).

IEC 60318-1 (1998). Electroacoustics — Simulators of human head and ear —- Part 1: Ear simulator for the calibration of supra-aural earphones. International Electrotechnical Commission (Geneva).

IEC 60813-5 (1985). 2cm3 coupler for the measurement of hearing aids and earphones coupled to the ear by means of ear inserts. International Electrotechnical Commission (Geneva).

ISO 389-8 (2004). Acoustics - reference zero for the calibration of audiometric equipment - Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones. International Organization for Standardization, (Geneva).

ISO 8253-1 (1989). Acoustics - Acoustic test methods - Part 1: Basic pure-tone air and bone conduction threshold audiometry. International Organization for Standardization, (Geneva).

Jenstad, L., Seewald, R., Cornelisse, L., and Shantz, J. (1999). Comparison of linear gain and wide dynamic range compression hearing aid circuits: Aided speech perception measures. *Ear Hear*, *20*, 117–126.

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

Jenstad, L. M. and Souza, P. E. (2005). Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *J Speech Lang Hear Res*, *48*(3), 651–667.

Jones, D. and Macken, W. (1995). Auditory babble and cognitive efficiency: The role of number of voices and their locations. *Journal of Experimental Psychology: Applied*, *1*, 216–226.

Jones, D. M., Alford, D., Macken, W. J., Banbury, S. P., and Tremblay, S. (2000). Interference from degraded auditory stimuli: linear effects of changing-state in the irrelevant sequence. *J Acoust Soc Am*, *108*(3 Pt 1), 1082–1088.

Keidser, G. (1996). Selecting different amplfication for different listening conditions. *J Am Acad Audiol*, *7*, 92–104.

Keidser, G., Brew, C., Brewer, S., Dillon, H.and Grant, F., and Storey, L. (2005). The preferred response slopes and two-channel compression ratios in twenty listening conditions by hearing-impaired and normal-hearing listeners and their relationship to the acoustic input. *Int J Audiol*, *44*, 656–670.

Keidser, G. and Dillon, H. (2006). What's new in prescriptive fittings down under? In *Phonak Conference on Hearing Care for Adults*, (pp. 133–142).

Keidser, G., Dillon, H., and Convery, E. (2008). The effect of the base line response on self-adjustments of hearing aid gain. *J Acoust Soc Am*, *124*(3), 1668–1681.

Keidser, G., Dillon, H., Dyrlund, O., Carter, L., and Hartley, D. (2007). Preferred low- and high-frequency compression ratios among hearing aid users with moderately severe to profound hearing loss. *J Am Acad Audiol*, *18*(1), 17–33.

Keidser, G., O'Brien, A., Carter, L., McLelland, M., and Yeend, I. (2008). Variation in preferred gain with experience for hearing-aid users. *Int J Audiol*, *47*(10), 621–635.

Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., and Convery, E. (2006). The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers. *Int J Audiol*, *45*(10), 563–579.

Kochkin, S. (2005). MarkeTrak VII: Customer satisfaction with hearing instruments in the digital age. *Hear J*, *58 (9)*, 30, 32–34, 38–40, 42–43.

Kuk, F. K. and Pape, N. M. (1992). The reliability of a modified simplex procedure in hearing aid frequency-response selection. *J Speech Hear Res*, *35*, 418–429.

Kuk, F. K. and Pape, N. M. (1993). Relative satisfaction for frequency responses selected with a simplex procedure in different listening conditions. *J Speech Hear Res*, *36*, 168–177.

Laurence, R. F., Moore, B. C., and Glasberg, B. R. (1983). A comparison of behind-the-ear high-fidelity linear hearing aids and two-channel compression aids, in the laboratory and in everyday life. *Br J Audiol*, *17*(1), 31–48.

Leijon, A. and Stadler, S. (2008). Fast amplitude compression in hearing aids improves audibility but degrades speech information transmission. In *Proc of the European Signal Processing Conf (EUSIPCO)*.

Lindley, G. A., Palmer, C. V., Durrant, J., and Pratt, S. (2000). Adaptation to loudness and environmental stimuli in three newly fitted hearing aid users. *J Am Acad Audiol*, *11*(6), 316–322.

Ljudproduction AB (2007). The digiffects sound effects library. Audio Compact Disc collection. URL: http://www.ljudproduktion.se.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* Wiley, New York.

Ludvigsen, C. (2001). Audiological background and design rationale of Senso Diva. Widexpress 15, Widex A/S.

Magnusson, L. (1995). Reliable clinical determination of speech recognition scores using swedish pb words in speech-weighted noise. *Scandinavian Audiology*, *24*, 217–223.

Maré, M. J., Dreschler, W. A., and Verschuure, H. (1992). The effects of input-output configuration in syllabic compression on speech perception. *J Speech Hear Res*, *35*(3), 675–685.

Marozeau, J. and Florentine, M. (2007). Loudness growth in individual listeners with hearing losses: A review. *J Acoust Soc Am*, *122*(3), EL81–EL87.

Marriage, J., Moore, B. C. J., and Alcántara, J. I. (2004). Comparison of three procedures for initial fitting of compression hearing aids. III. Inexperienced versus experienced users. *Int J Audiol*, *43*(4), 198–210.

Mathers, C., Smith, A., and Concha, M. (2001). Global burden of hearing loss in the year 2000. Technical report, World Health Organization, Geneva. Also available on the worldwide web at www.who.int/whr.

McCandless, G. and Lyregaard, P. (1983). Prescription Of Gain/Output (POGO) for hearing aids. *Hear Instrum*, *34*(1), 16–21.

Moore, B. (2000). Use of a loudness model for hearing aid fitting. IV. Fitting hearing aids with multi-channel compression so as to restore 'normal' loudness for speech at different levels. *Br J Audiol*, *34*(3), 165–177.

Moore, B., Alcántara, J., and Marriage, J. (2001). Comparison of three procedures for initial fitting of compression hearing aids. I. experienced users, fitted bilaterally. *Br J Audiol*, *35*, 339–353.

Moore, B., Alcántara, J., Stone, M., and Glasberg, B. (1999a). Use of a loudness model for hearing aid fitting: II. Hearing aids with multi-channel compression. *Br J Audiol*, *33*, 157–170.

Moore, B. and Glasberg, B. R. (1998). Use of a loudness model for hearing-aid fitting. I. Linear hearing aids. *Br J Audiol*, *32*(5), 317–335.

Moore, B., Glasberg, B. R., and Stone, M. A. (1999b). Use of a loudness model for hearing aid fitting: III. A general method for deriving initial fittings for hearing aids with multi-channel compression. *Br J Audiol*, *33*(4), 241–258.

Moore, B. C. (1998). *Cochlear Hearing Loss*. Whurr, London.

Moore, B. C. (2004). Testing the concept of softness imperception: Loudness near threshold for hearing-impaired ears. *J Acoust Soc Am*, *115*, 3103–3111.

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

Moore, B. C., Lynch, C., and Stone, M. A. (1992). Effects of the fitting parameters of a two-channel compression system on the intelligibility of speech in quiet and in noise. *Br J Audiol*, *26*(6), 369–379.

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

Moore, B. C. J., Stainsby, T. H., Alcántara, J. I., and Kühnel, V. (2004). The effect on speech intelligibility of varying compression time constants in a digital hearing aid. *Int J Audiol*, *43*(7), 399–409.

Munro, K. J. and Lutman, M. E. (2003). The effect of speech presentation level on measurement of auditory acclimatization to amplified speech. *J Acoust Soc Am*, *114*(1), 484–495.

Munro, K. J. and Lutman, M. E. (2005). Sound quality judgements of new hearing instrument users over a 24-week post-fitting period. *Int J Audiol*, *44*(2), 92–101.

Munro, K. J. and Trotter, J. H. (2006). Preliminary evidence of asymmetry in uncomfortable loudness levels after unilateral hearing aid experience: evidence of functional plasticity in the adult auditory system. *Int J Audiol*, *45*(12), 684–688.

Naylor, G. and Johannesson, R. B. (2009). Long term signal-to-noise ratio at the input and output of amplitude-compression systems. *J Am Acad Audiol*, *20*, 161–171.

Neuman, A. C., Bakke, M. H., Mackersie, C., and Hellman, S. (1998). The effect of compression ratio and release time on the categorical rating of sound quality. *J Acoust Soc Am*, *103*, 2273–2281.

Olsen, H. L., Olofsson, A., and Hagerman, B. (2005). The effect of audibility, signal-to-noise ratio, and temporal speech cues on the benefit from fast-acting compression in modulated noise. *Int J Audiol*, *44*(7), 421–433.

Olsen, S. O., Rasmussen, A. N., Nielsen, L. H., and Borgkvist, B. V. (1999). Loudness perception is influenced by long-term hearing aid use. *Audiology*, *38*(4), 202–205.

Palmer, C. V. (2002). *Strategies for selecting and verifying hearing aid fittings* (Second ed.)., chapter Fitting strategies for patients with symmetrical hearing loss, (pp. 202–220). Thieme, New York.

Palmer, C. V., Nelson, C. T., and Lindley, G. A. (1998). The functionally and physiologically plastic adult auditory system. *J Acoust Soc Am*, *103*(4), 1705–1721.

Pascoe, D. (1988). Clinical meaurements of the auditory dynamic range and their relations to formulas for hearing aid gain. In Hartvig Jenson, J. (Ed.), *Proceedings of the 13th Danavox Symposium*, (pp. 129–151).

Pearsons, K., Bennett, R., and Fidell, S. (1977). Speech levels in various noise environments. project report on contract 68 01-2466. Technical report, Washington, DC: U.S. Environmental Protection Agency.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Philibert, B., Collet, L., Vesson, J.-F., and Veuillet, E. (2002). Intensity-related performances are modified by long-term hearing aid use: a functional plasticity? *Hear Res*, *165*(1-2), 142–151.

Philibert, B., Collet, L., Vesson, J.-F., and Veuillet, E. (2005). The auditory acclimatization effect in sensorineural hearing-impaired listeners: evidence for functional plasticity. *Hear Res*, *205*(1-2), 131–142.

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

Preminger, J., Neuman, A., Bakke, M., Walters, D., and Levitt, H. (2000). An examination of the practicality of the simplex procedure. *Ear Hear*, *21*, 177–193.

Purdy, S. C. and Pavlovic, C. V. (1992). Reliability, sensitivity and validity of magnitude estimation, category scaling and paired-comparison judgements of speech intelligibility by older listeners. *Audiology*, *31*(5), 254–271.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robinson, K. and Gatehouse, S. (1996). Test-retest reliability of loudness scaling. *Ear Hear*, *17*(2), 120–123.

Savage, I., Dillon, H., Byrne, D., and Bächler, H. (2006). Experimental evaluation of different methods of limiting the maximum output of hearing aids. *Ear Hear*, *27*, 550–562.

Schmidt, E. (2006). *Hearing aid processing of loud speech and noise signals: Consequences for loudness perception and listening comfort.* PhD thesis, Technical University of Denmark.

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., Beaulac, S., and Pumford, J. (2005). The Desired Sensation Level multistage Input/Output algorithm. *Trends Amplif*, *9*, 159–197.

Seewald, R., Ramji, K., Sinclair, S., Moodie, K., and Jamieson, D. (1993). *Computer-assisted implementation of the Desired Sensation Level method for electroacoustic selection and fitting in children: Version 3.1. Users Manual.* London, Ontario: The University of Western Ontario.

Shi, L.-F. and Doherty, K. A. (2008). Subjective and objective effects of fast and slow compression on the perception of reverberant speech in listeners with hearing loss. *J Speech Lang Hear Res*, *51*(5), 1328–1340.

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

Smeds, K. (2004b). *Less is more? Loudness aspects of prescriptive methods for nonlinear hearing aids*. PhD thesis, Signals, Sensors and systems, Royal Institute of Technology (KTH), Stockholm, Sweden.

Smeds, K., Keidser, G., Zakis, J., Dillon, H., Leijon, A., Grant, F., Convery, E., and Brew, C. (2006a). Preferred overall loudness. I: Sound field presentation in the laboratory. *Int J Audiol*, *45*, 2–11.

Smeds, K., Keidser, G., Zakis, J., Dillon, H., Leijon, A., Grant, F., Convery, E., and Brew, C. (2006b). Preferred overall loudness. II: Listening through hearing aids in field and laboratory tests. *Int J Audiol*, *45*, 12–45.

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

Souza, P. E., Jenstad, L. M., and Boike, K. T. (2006). Measuring the acoustic effects of compression amplification on speech in noise. *J Acoust Soc Am*, *119*, 41–44.

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

Stach, B. (1998). *Clinical Audiology: An Introduction*. Singular Publishing.

Stone, M. A. and Moore, B. C. (1992). Syllabic compression: effective compression ratios for signals modulated at different rates. *Br J Audiol*, *26*(6), 351–361.

Stone, M. A. and Moore, B. C. J. (2008). Effects of spectro-temporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task. *J Acoust Soc Am*, *123*(2), 1063–1076.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J Speech Hear Res*, *28*(3), 455–462.

Sweetow, R. W. and Sabes, J. H. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. *J Am Acad Audiol*, *17*(8), 538–558.

Thorne, D. R., Genser, S. G., Sing, H. C., and Hegge, F. W. (1985). The Walter Reed performance assessment battery. *Neurobehav Toxicol Teratol*, *7*(4), 415–418.

Wickelmaier, F. and Schmid, C. (2004). A matlab function to estimate choice model parameters from paired-comparison data. *Behav Res Methods Instrum Comput*, *36*(1), 29–40.

Widex (1999). Real-life environmental sounds examples. Audio Compact Disc.

Wilson, D. H., Walsh, P. G., Sanchez, L., Davis, A. C., Taylor, A. W., Tucker, G., and Meagher, I. (1999). The epidemiology of hearing impairment in an australian adult population. *Int J Epidemiol*, *28*(2), 247–252.

Xu, J., Cox, R., and Naylor, G. (2008). Do clinical paired comparisons predict real world release time preference? In *Poster presented at the American Auditory Society annual conference in Scottsdale, Arizona.* Published on the world wide web at http://www.memphis.edu/ausp/harl/publications.htm.

Yund, E. W., Roup, C. M., Simon, H. J., and Bowman, G. A. (2006). Acclimatization in wide dynamic range multichannel compression and linear amplification hearing aids. *J Rehabil Res Dev*, *43*(4), 517–536.

Zielinski, S., Rumsey, F., and Bech, S. (2008). On some biases encountered in modern audio quality listening tests - a review. *J Audio Eng Soc*, *56*, 427–451.

Zimmer, K., Ellermeier, W., and Schmid, C. (2004). Using probabilistic choice models to investigate auditory unpleasantness. *Acta Acust United Ac*, *90*, 1019–1028.

A

Definitions of Compression Parameters

The main parameters used for characterizing the properties of a compression system are IEC 60118-2 (1983):

Attack time (AT) the time taken for the output to stabilize to within 2 dB of it's final level after the input to the hearing aid increases from 55 to 80 dB SPL.

Release time (RT) the time taken for the output signal to increase to within 2 dB of its final value following a decrease in input level from 80 to 55 dB SPL.

Compression Ratio (CR) the change in input level needed to produce a 1 dB change in output level.

Compression Threshold (CT) Defined as the input sound pressure level at which the output deviates by 2 dB from the output that would have occurred had linear amplification continued to higher input levels.

Number of compression channels The number of different amplifiers/compressors operating in different frequency bands (Dillon, 2001, p. 38).

B

Field trial interview questions in Swedish

1 Hur har det gått?
(Kommentar)

2 Vilket program föredrog du generellt?
(Försökspersonerna var tvingade att välja mellan program 1 och program 2.)

3a Vilket program har den största komfortnivån?
(Försökspersonerna kunde välja mellan program 1 och program 2, om de kunde höra en skillnad mellan dem.)

3b Hur skulle du beskriva den generella komfortnivån/behaglighetsnivån?
(Svarsalternativen var "Mycket bra", "Bra", "Mitt emellan", "Dåligt" och "Mycket dåligt".)

4a 4. Vilket program hade den bästa tydligheten för talet vid en normal konversation i en tyst miljö?
(Försökspersonerna kunde välja mellan program 1 och program 2, om de kunde höra en skillnad mellan dem.)

4b Hur skulle du generellt beskriva talets tydlighet vid en normal konversation i en tyst miljö?
(Svarsalternativen var "Mycket bra", "Bra", "Mitt emellan", "Dåligt" och "Mycket dåligt".)

5a Vilket program föredrog du när det är tyst och det finns svaga ljud omkring dig (kyl, frys, ventilation, fåglar)?
(Försökspersonerna kunde välja mellan program 1 och program 2, om de kunde höra en skillnad mellan dem.)

5b Tycker du om att höra de svaga ljuden "bra"?
(Svarsalternativen var "Ja" och "Nej" med möjlighet att kommentera.)

5c Vilka svaga ljud vill du höra?
(Kommentar)

6a I vilket program var talet tydligast för svaga röster eller röster på avstånd?
(Försökspersonerna kunde välja mellan program 1 och program 2, om de kunde höra en skillnad mellan dem.)

6b Hur skulle du generellt beskriva talets tydlighet för svaga röster eller röster på avstånd?
(Svarsalternativen var "Mycket bra", "Bra", "Mitt emellan", "Dåligt" och "Mycket dåligt".)

7a I vilket program var talet tydligast när det gäller tal i buller?
(Försökspersonerna kunde välja mellan program 1 och program 2, om de kunde höra en skillnad mellan dem.)

7b Hur skulle du beskriva talets tydlighet i bullriga miljöer?
(Svarsalternativen var "Mycket bra", "Bra", "Mitt emellan", "Dåligt" och "Mycket dåligt".)

8a Hur bullrigt tycker du att bakgrundsljudet är i bullriga situationer?
(Svarsalternativen var "För svagt", "Lite för svagt", "Lagom", "Lite för starkt" och "För starkt".)

8b Vilket program är mest bullrigt i dessa bullriga situationer?
(Försökspersonerna kunde välja mellan program 1 och program 2, om de kunde höra en skillnad mellan dem.)

9 Hur skulle du beskriva ljudkvalitén på din egen röst med hörapparaterna?
(Svarsalternativen var "Mycket bra", "Bra", "Mitt emellan", "Dåligt" och "Mycket dåligt".)

10 Tycker du att hörapparaterna har gett ifrån sig ljud i tysta situationer?
(Svarsalternativen var "Ja" och "Nej".)

11 Hade du problem med återkoppling i något program?
(Svarsalternativen var "Ja" och "Nej".)

C

Recording of running speech material in Swedish

The running Swedish signals were recorded in mono-format (24 bit, 44.1 kHz) using a G.R.A.S. $\frac{1}{2}$- inch condenser microphone placed 40 cm from the speakers mouth in the soundproof booth at ORCA-EU. The speaker was a linguist called Christine Eriksdotter who speaks Standard Swedish (Rikssvenska) with some influence of a dialect from North Sweden (Jämtländska). The speaker told stories as if she was talking to a friend. The recordings were cut into approximately one minute long stories. Within these sound files, the levels of individual sentences were adjusted to get approximately the same level across the whole sound file. After level adjustment, reverberation was added in CoolEdit so that the speech would sound more natural when played back with the ICRA2 (Bjerg and Larsen, 2006) background signals. The reverberation settings are shown in table C.1. Finally, all the sound files were adjusted in gain to a total common level of -35 dB RMS and a 15 second long explanatory pre-amble was added at the beginning of each sound file.

Table C.1: Reverberation settings applied to recordings in CoolEdit.

Parameter	Value
Total Reverberation Length	900 ms
Attack time	10 ms
High frequency absorption time	1100 ms
Echoey Perception	0
Percentage Original Signal	100 %
Percentage Reverb Signal	25 %

Contributions to Hearing Research

Vol. 1: *Gilles Pigasse*, Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials, Dec. 2008.

Vol. 2: *Olaf Strelcyk*, Peripheral auditory processing and speech reception in impaired hearing, Apr. 2009.

Vol. 3: *Eric R. Thompson*, Characterizing binaural processing of amplitude-modulated sounds, Jun. 2009.

Vol. 4: *Tobias Piechowiak*, Spectro-temporal analysis of complex sounds in the human auditory system, Jul. 2009.

Vol. 5: *Jens Bo Nielsen*, Assessment of speech intelligibility in background noise and reverberation, Aug. 2009.

Vol. 6: *Helen Connor Sørensen*, Hearing aid amplification at soft input levels, Jan. 2010.

Vol. 7: *Morten L. Jepsen*, Modeling auditory processing and speech perception in hearing-impaired listeners, Feb. 2010.