The Effect of Amplitude Compression on the Perception of Speech in Noise by the Hearing Impaired

Typeset in LaTeX

Printed by Gildeprint BV, Enschede

The Effect of Amplitude Compression on the Perception of Speech in Noise by the Hearing Impaired

Het effect van amplitude-compressie op het spraakverstaan in lawaai door slechthorenden

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. W.H. Gispen,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 10 oktober 2006 des middags te 2.30 uur

door

Adrian Carina Henri Houben

geboren op 10 september 1972 te Weert

Promotor: Prof.dr. G.F. Smoorenburg

Publication of this thesis was financially supported by: Stichting Atze Spoor Fonds, Makker Hoortoestellen, Oticon Nederland BV, GN Resound BV, Siemens Audiologie Techniek BV, Atos Medical, Schoonenberg Hoorcomfort, EmiD Audiologische Apparatuur, and Beltone Netherlands BV.

Contents

List of abbreviations

ΔSRT	SRT obtained for compression minus SRT for linear amplification [dB]
AGC	automatic gain control, also called automatic volume control (AVC)
ANOVA	analysis of variance
ANSI	American National Standards Institute
APHAB	abbreviated profile of hearing aid benefit
AVC	automatic volume control, also called automatic gain control (AGC)
CF	characteristic frequency [Hz]
CP	compression parameters: $CP(NC, CR, T)$
CR	compression ratio, characteristic of compression: change in input level [dB] divided by the change in output level [dB]
CR_{high}	compression ratio for frequency channels above 1000 Hz [dB/dB]
CR_{low}	compression ratio for frequency channels below 1000 Hz [dB/dB]
DSL [i/o]	desired sensation level. Fitting rationale for linear and compression hearing aids
DR	dynamic range of hearing [dB]
$DR_{x,y,z\,kHz}$	average of the dynamic range of hearing for pure tones at frequencies of x, y, and z kHz [dB]
DR_{speech}	dynamic range of hearing for speech; uncomfortable loudness level of broadband speech minus the speech reception threshold in quiet [dB]
F-ratio	Fisher ratio, ratio of variances here: statistics, $F = MS_{effect}/MS_{error}$
FIG6	fitting rationale for compression hearing aids
HL	dB HL, hearing level: level [dB] relative to the (frequency dependent) standardized equivalent threshold level (ISO 389-2, 1996)
HI	hearing-impaired listeners
HSD-test	statistics, Tukey's honestly significant difference test
GOF	growth of masking (used in forward masking experiments to estimate compression ratio)
IEC	International Electrotechnical Commission
IHC	inner hair cell
L	(sound pressure) level [dB] (SPL)
MS	statistics, mean square

MS_{effect}	statistics, mean square of the effect under investigation
MS_{error}	statistics, mean square of the error term
NAL-NL1	National Acoustics Laboratories nonlinear. Fitting rationale for compression hearing aids
NAL-RP	National Acoustics Laboratories revised profound. Fitting rationale for linear hearing aids for subjects with mild to profound hearing loss
NC	number of channels, characteristic of compression: the number of frequency channels that are independently compressed
NH	normal hearing listeners
OHC	outer hair cell
p	statistics, p-value, probability that the sample could have been drawn from the population being tested given the assumption that the null hypothesis is true
PTA	pure-tone average hearing loss, the average hearing loss for pure tones, expressed relative to normal hearing [dB HL]
$PTA_{x, y, z\ kHz}$	average of the hearing loss for pure tones at frequencies of x, y, and z kHz [dB HL]
rms	root mean square, the rms value of a fluctuating quantity x is given by: $x_{rms} = \sqrt{(x_1^2 + x_2^2 + \cdots + x_N^2)/N}$
r^2	statistics, fraction of the variance in the data that is explained by a regression line
SC	syllabic compression
SNR	signal-to-noise ratio [dB]
SPL	dB SPL, sound pressure level: level [dB] relative to the reference sound pressure (20 μPa)
SRT	speech reception threshold in noise [dB]: the signal-to-noise ratio at which 50% of the presented sentences were repeated correctly
$SRT_{stat,\ linear}$	SRT [dB] for stationary noise and linear amplification
$SRT_{fluct,\ linear}$	SRT [dB] for fluctuating noise and linear amplification
T	time constants, characteristic of compression ($T = T_a/T_r$) [ms]
T_a	attack time [ms]
T_{ANSI}	time constants [ms] as defined in ANSI S3.22 (1996)
T_{exp}	time constants [ms] based on an exponential decay
T_{IEC}	time constants [ms] as defined in IEC 60188-2
T_r	release time [ms]
TM	temporal masking (used in forward masking experiments to estimate compression ratio)
UCL	uncomfortable loudness level
WDRC	wide dynamic range compression

“This golden age of communication means everybody speaks at the same time”

Lyrics from ‘225’,

Justin Sullivan and Robert Heaton,

New Model Army, ‘Thunder and Consolation’, 1988

1 Introduction

1.1 Introduction

Since the earliest days of electronic sound reproduction, people have tried to develop hearing aids to reduce hearing disability. The first "practical" electronic hearing aid was based on a carbon microphone, and was made as early as 1902. Since this first device, hearing aid amplification has improved steadily: from vacuum tubes (1920s), via transistors (1950s), to integrated circuits (1960s). And finally, in the last decades we have seen the rise of powerful digital signal processing. Meanwhile, production techniques drastically improved and electronic components were miniaturized. New hearing aids became ever smaller, while still containing powerful processing capabilities. Currently, hearing aid technology has advanced to the point where it is possible to incorporate more than simple and straight forward amplification in wearable devices: the new technology allows the utilization of sophisticated audiological knowledge for improving hearing ability.

Despite these technological and audiological advances, not every hearing aid user is completely satisfied with his/her hearing aid. For the Netherlands it was estimated that only 21% of the hearing-impaired population (i.e., people reporting a hearing difficulty) owned a hearing aid (calculated from Duijvesteijn, 1999). An American study reported a similar situation: in 1999 the market penetration of hearing aids was limited to only 20% of the total American hearing-impaired population (Kochkin, 1999). Moreover, of the hearing aid owners roughly 16% did not use their aids at all (Kochkin, 2000). The primary reason was small benefit, especially in noisy situations.

Hearing disability is most significantly and most frequently present in daily-life communication in noisy situations, such as conversations in restaurants and at parties. A conversation in a background of other speakers is notoriously difficult to follow for hearing-impaired listeners, especially for people with sensorineural hearing loss. This type of hearing loss involves a dysfunction of the sensorineural structures of the inner ear; or in other words, it is often caused by 'defective hearing sensors'. Sensorineural hearing loss generally leads to lower speech intelligibility in a background noise, even if hearing aids are used. The prevalence of this hearing loss is quite extensive and common causes are for instance ageing, excessive noise exposure, genetic predisposition, and ototoxic medication.

It is important to further reduce hearing handicap, especially in noisy situations. One means is to develop hearing aids in which sophisticated audiology-based signal processing is applied. This thesis investigates several (compressive) amplification strategies to determine which strategy provides the best speech intelligibility in noisy situations for listeners with a sensorineural hearing loss.

1.2 The basics of hearing

In daily life we often depend on hearing to interact with our surroundings. A close look at the hearing mechanisms reveals that our hearing ability, which comes so easily, stems from an amazing and dynamic process. The next three sections explain the basic anatomy and sound processing that occurs in the outer and middle ear, the inner ear, and the auditory nerve. Each section is divided into A) anatomy, B) sound transmission, and C) hearing loss.

1.2.1 Outer and middle ear

A. Anatomy

An incoming sound wave is first diffracted by the torso, the head and the outer ear of the listener. It then propagates through the middle ear. Figure 1.1 illustrates the stages in the outer ear and onwards.

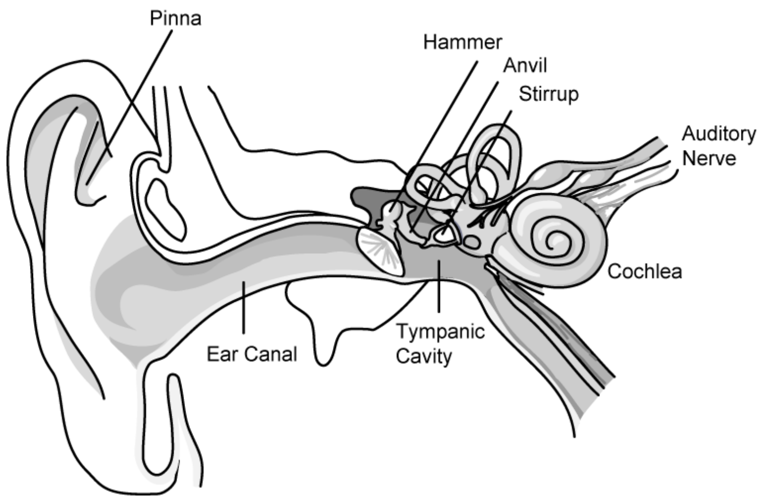


Figure 1.1: *Schematic coronal section of the human ear (adapted from The Capital Region Otitis Media Project (2000), with permission).*

After reaching the pinna, the sound wave continues trough the ear canal to the tympanic membrane (eardrum), which separates the tympanic cavity from the ear canal. Inside the tympanic cavity, the three ossicles (hammer, anvil, and stirrup) convey the sound from the tympanic membrane to the oval window. This oval window is the entrance of the inner ear.

B. Sound transmission in the outer and middle ear

The eardrum, ossicles, and oval window form an impedance transformer; they change the incoming sound pressure wave into a fluid pressure wave in the cochlea. Without this impedance transformation about 99.9% of the incoming sound energy would be reflected back from the fluid surface (Pickles, 1988). The influence of the torso, the pinna, and the entire middle ear can be approximated by a band pass filter (Palmer and Shamma, 2004). The transduction behaves linearly, i.e., the sound transmission depends on frequency only and is unaffected by the intensity of the incoming sound wave. However, this approximation is only accurate for moderate sound levels. At higher sound levels additional nonlinear processes occur. For instance, the small muscles which are connected to the ossicles can protect the inner ear by contracting at high sound pressure levels (middle ear reflex). The contraction predominantly attenuates the low frequency components. For high sound levels these muscles therefore introduce a nonlinear response. Moreover, at very high sound levels the tympanic membrane and the ossicles themselves can introduce nonlinear distortion (Moore, 1998). Nevertheless, at moderate sound levels ($<$90 dB SPL), these nonlinear effects do not occur and the transmission through outer and middle ear can be regarded as a linear system. In the experiments in this thesis (Chapters 2, 3 and 4), all stimuli were presented at moderate levels for which it is safe to assume that the middle ear behaves linearly.

C. Conductive hearing loss

Damage to the outer and middle ear will impair the hearing abilities of the listener. Such hearing impairment manifests itself mostly as a frequency dependent loss of hearing sensitivity, i.e., a conductive hearing loss. Surgical procedures or the application of a hearing-aid can often lead to substantial improvement of hearing ability.

1.2.2 The cochlea

A. Anatomy

The stirrup conveys the incoming sound wave to the cochlea, which converts the sound wave into nerve impulses. The cochlea consists of three fluid filled canals which together are coiled in the form of a snail's shell, hence the name (from Greek "kochlos" meaning "spiral shell"). Figure 1.2 shows a schematic cross section of a human cochlea. Sound waves enter the cochlea via the oval window. At the top of the spiral (apex), the scala vestibuli is connected to the adjacent scala tympani. This connection is called helicotrema. At the far end, the scala tympani is closed by the round window, which is a membrane that allows the incompressible fluid to displace in accordance to the motion forced on the fluid through the oval window (Nedzelnitsky, 1980).

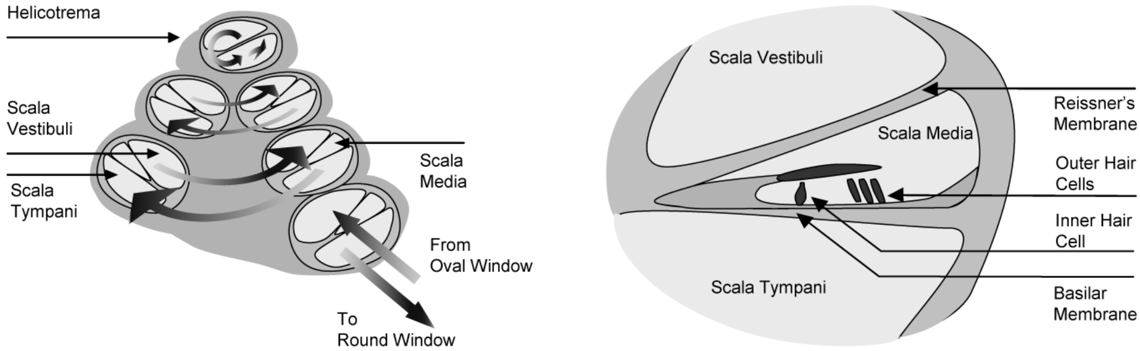


Figure 1.2: *Schematic cross section of the cochlea. The arrows represent the incoming and outgoing sound waves. The right panel shows a cross section of a turn of the cochlea.*

Between the scala vestibule and scala tympani, another duct is present: the scala media. This scala is not connected to the other two scalae, it ends blindly at the helicotrema. The scala media is separated from the scala vestibuli by Reissner's membrane and from the scala tympani by the basilar membrane. The actual sensory organ, the organ of Corti, is situated on the basilar membrane (inside the scala media). Sound pressure waves force the basilar membrane to vibrate and the organ of Corti converts this movement to electro-chemical activity that elicits nerve impulses. The organ of Corti contains two types of receptor cells: inner hear cells (IHCs) and outer hair cells (OHCs). These hair cells are arranged in four rows; one row of IHCs and three rows of OHCs. The function of IHCs and OHCs differs. The IHCs are the actual transduction cells which convert the incoming sound waves to outgoing nerve impulses. The OHCs are thought to amplify incoming sound to facilitate detection by the IHCs (Pickles, 1988).

B. Sound transmission in the cochlea

An incoming sound wave moves over the basilar membrane as a travelling wave (Von Békésy, 1960). The mechanical compliance (e.g., the stiffness and mass) of the basilar membrane changes systematically along the membrane. This change evokes a maximal displacement of the sound wave to occur at a specific position on the basilar membrane. This place on the basilar membrane corresponds directly to the frequency of the incoming sound wave: the characteristic frequency CF of this place. The base of the cochlea responds well to high frequencies, the apex to low frequencies. The passive system, that is the system without OHC function, has poor frequency selectivity. Passive frequency tuning has a shallow slope on the low-frequency side and a steeper slope on the high-frequency side (Geisler, 1998). In fact, the basilar membrane can be regarded as a low-pass filter (Pickles, 1988). Thus, if we disregard OHC function, the response of the basilar membrane is frequency dependent and approximately linear.

However, this passive system is a gross simplification since it is thought that the OHCs act as an active bio-mechanical amplifier. The OHCs amplify incoming sound waves when their frequencies are close to the characteristic frequency. The amplification is highly nonlinear and it strongly influences threshold, dynamic range of hearing, and frequency selectivity, see section 1.3.

C. Cochlear hearing loss

Damage to the IHCs leads to an increased threshold of hearing. If the extent of IHC loss is limited, amplification provided by hearing aids can restore hearing thresholds to near normal. Hearing loss due to loss of OHCs is much more common than hearing loss due to loss of IHCs, since OHCs are more susceptible to damage than IHCs. A dysfunction of the OHCs leads to less bio-mechanical amplification, which results in a loss of sensitivity, a smaller dynamic range, and impaired frequency resolution. The consequences of extensive OHC loss are far reaching and complex, and are difficult to treat. Surgical procedures and efficacious medication hardly exist (Dobie, 1997; Murugasu, 2005), while assistive devices such as hearing aids can not restore the impaired hearing to normal.

Hearing loss that originates from damaged OHCs is the central theme of this thesis and it will be discussed in more detail in sections 1.3 to 1.5.

1.2.3 The auditory nerve

A. Anatomy

The IHCs transform the incoming fluid pressure wave into nerve impulses. Each IHC is connected to about twenty afferent nerve fibres (Pickles, 1988). Spectral content is encoded in two ways, which are both illustrated in Figure 1.3. First, each nerve fibre corresponds to a specific cochlear location that is most sensitive to a specific frequency (place coding). Second, the timing of the nerve spikes does not occur randomly in time. The discharges occur at specific moments in time corresponding to the phase of the waveform at the innervation site (phase locking). Phase locking is limited to the lower frequencies. For instance, in the cat, phase locking occurs for frequencies up to 5 kHz (Rose et al., 1967; Palmer and Shamma, 2004).

Most nerve fibres (about 80%) respond to low intensity sounds (20 dB SPL or less). These nerve fibres have a high spontaneous discharge rate (40–120 discharges/s) and a relatively small dynamic range (30 dB). Saturation at such low levels suggests extreme compression. However, another type of nerve fibres with a low spontaneous rate of firing (<20/s), has higher thresholds (up to 80 dB SPL) and larger dynamic ranges

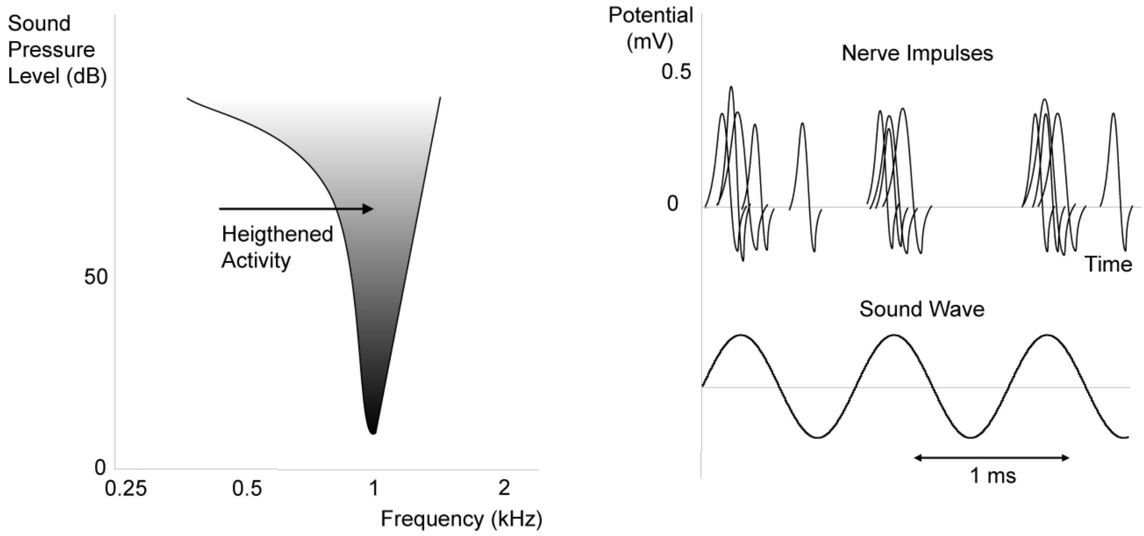


Figure 1.3: *Spectral coding in the auditory nerve. The left panel shows place coding, the right panel phase locking.*

(up to 70 dB). Although the dynamic range of the individual nerve fibres might be small, phase locking still occurs across a much larger range of sound levels. In fact, it has been found from about 20 dB below the threshold of an individual fibre, up to levels well above saturation (at 90 dB SPL) of individual fibres (Palmer and Shamma, 2004). When the information from place coding is lost due to saturation of the nerve fibers, the temporal information is still present in the phase locking. The auditory nerve can thus convey information over a large range of input levels, without the need for compression.

B. Sound transmission in the auditory nerve

Due to the frequency selectivity of the basilar membrane, the individual fibres of the auditory nerve are only sensitive to specific frequency regions (Rouiller, 1997). The response (firing rate and phase locking) in the auditory nerve strongly resembles the response of the basilar membrane. This suggests that the auditory nerve does not substantially alter the dynamics of the incoming signal. Smoorenburg (1974) showed that specific nonlinear phenomena such as two-tone suppression (see section 1.3.2) and intermodulation distortion (see section 1.3.4) could originate from the cochlear nonlinearity. Indeed, it is now believed that these specific nonlinear phenomena originate from the mechanical nonlinearity of the cochlea and not from the auditory nerve (Geisler, 1998, Robles and Ruggero, 2001).

C. Hearing loss of neural origin

In contrast to cochlear hearing loss, which typically begins at high frequencies, neural hearing loss often occurs evenly throughout the cochlea (e.g., neural presbycusis). Mild injury to the auditory nerve mainly affects the timing of the neural discharges (Møller, 2000). Consequently, pure tone sensitivity remains unaffected until the hearing loss is quite severe (up to 90% of neurons may be lost; Roeser et al., 2000) and the dynamic range of hearing is not affected. However, speech discrimination at moderate levels is often disproportionately lower than suggested by the pure tone sensitivity (Jerger and Jerger, 1981), and it can even decrease as the speech level is increased. Since neural hearing loss affects the timing of the neural impulses, it has the strongest effect on timing-related perceptions and tasks, such as temporal masking, gap detection, temporal integration, and low-frequency pitch perception (Zeng et al., 2005). Moderate neural hearing loss does not affect the dynamic range of hearing, and therefore compressive amplification (section 1.6) is not suitable for improving speech intelligibility for this type of hearing loss.

1.3 Perceptual consequences of cochlear compression

The cochlear nonlinearity influences auditory perception considerably, and behaves in an essentially nonlinear way (Goldstein, 1967; Eguíluz et al. 2000). This means that the nonlinearity is already present at moderate sound levels and is not the result from distortion caused by high sound pressure levels. Figure 1.4 illustrates the effect of cochlear amplification. The figure is based on a famous in vivo experiment by Ruggero and Rich (1991), and shows the velocity of the basilar membrane as a function of the input sound level.

The figure shows amplification at low levels resulting in a compressive input-output function between 30 and 90 dB SPL (solid line). For these levels, the velocity of the basilar membrane increases with about 0.5 dB for each increase of 1 dB in stimulus level; the response is nonlinear. In contrast, for low (<30 dB SPL) and high (>90 dB SPL) stimulus levels, the response is linear (response growth of 1 dB per 1 dB increase in input level). The dashed line in Figure 1.4 shows the situation where OHC amplification is temporarily suppressed; the response is linear. The response to tones with a frequency far away from the characteristic frequency, i.e., far away from the recording location on the basilar membrane, also behaves linearly.

1.3.1 Improved sensitivity and dynamic range

Figure 1.4 illustrates the influence of the compressive nonlinearity at low thresholds. The high amount of amplification for low input levels improves hearing sensitivity. For human hearing, it has been estimated that the outer hair cells can improve the threshold by as much as 40 to 60 dB. The measurement from Ruggero and Rich in a chinchilla shows an increase of about 30 dB (see Figure 1.4).

The lower threshold of hearing greatly increases the dynamic range of the system. The

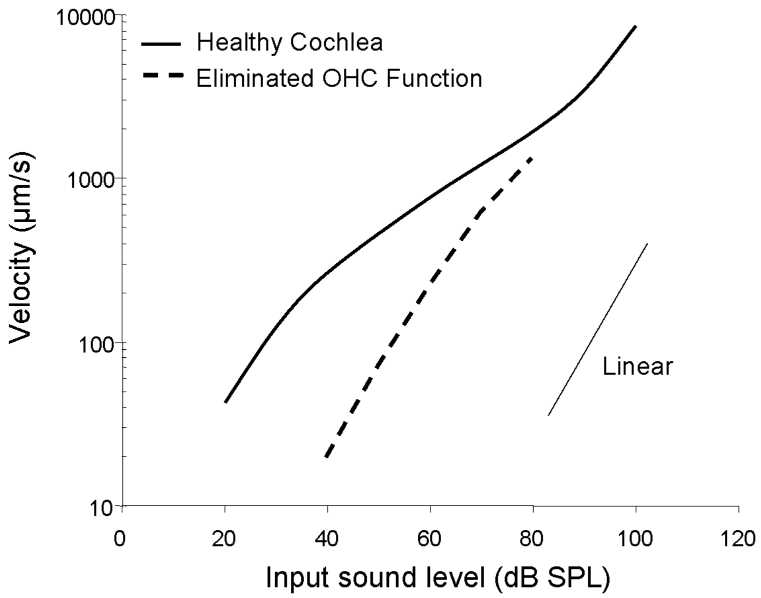


Figure 1.4: *Illustration of the input-output function of the basilar membrane of a chinchilla. The solid line represents the response of a healthy cochlea at the characteristic frequency, the dashed line the response of a cochlea with eliminated OHC function. Redrawn from Ruggero and Rich, 1991. Additionally, the thin line indicates the curve for a linear response. A factor 10 change in velocity (ordinate) corresponds to 20 dB.*

compressive nonlinearity therefore directly improves the human dynamic range from 60–80 dB HL to about 120 dB HL. In other words, the nonlinearity effectively reduces (compresses) the incoming sound waves by about 2 dB per 1 dB.

1.3.2 Improved frequency selectivity

The frequency tuning on the basilar membrane is generally thought to stem from two components. The first arises from the hydro-mechanical properties of the cochlea (such as the compliance, friction, and mass of the basilar membrane; the volume of the scalae; the viscosity of the cochlear fluids etc.). This component acts only passively and its behaviour is linear for all stimulus levels. The second component is thought to arise from the active influence of the outer hair cells on the basilar membrane. This effect is markedly nonlinear: at low input levels the bio-mechanical amplification is higher than at higher input levels. The sharpness of tuning therefore depends on the stimulus level (Robles and Ruggero, 2001). For low input levels it is dominated by the active OHC component, at high input levels by the passive tuning of the basilar membrane. Figure 1.5 illustrates the sensitivity of basilar membrane responses as a function of frequency and stimulus level.

The curves for low stimulus levels (e.g., 40 dB SPL), show a sharp drop in sensitivity for frequencies which are further away from the characteristic frequency. For higher stimulus levels, the sensitivity changes less with frequency; the tuning is much broader.

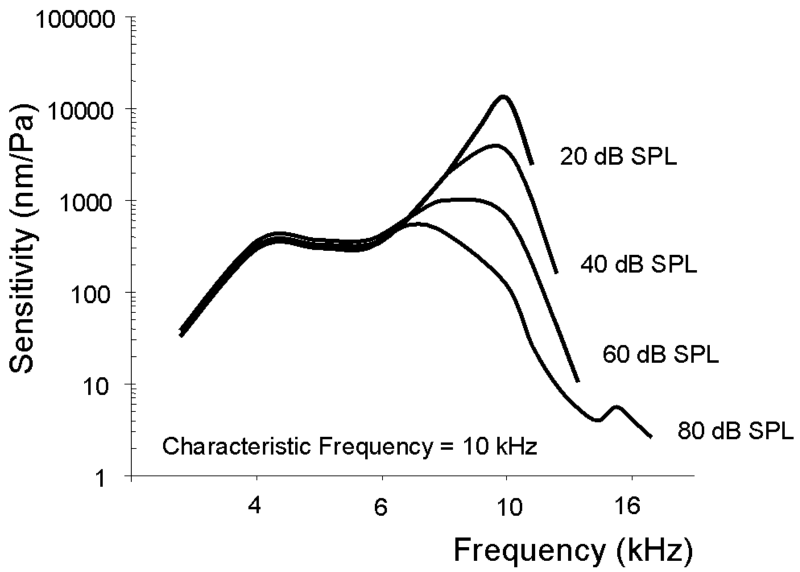


Figure 1.5: *Sensitivity (displacement divided by stimulus sound pressure) of basilar membrane responses to tones as a function of frequency. The different lines (isointensity curves) show the response for different input sound pressure levels (in dB SPL). Redrawn from Robles and Ruggero (2001).*

The presence of a second stimulus can change the response to the primary stimulus. For instance, if two tones are presented at the same time, the response to the second tone can be less than the response for that tone presented alone. This suppression effect has been found in all auditory nerve fibres (Rouiller, 1997), and it was previously attributed to synaptic inhibition (Sachs and Kiang, 1968). Based on psychophysical data Smoorenburg (1972; 1974) argued that suppression could be explained by a cochlear nonlinearity. Physiological measurements of basilar membrane responses (Ruggero et al., 1992), and computational models (Giguère and Smoorenburg, 1997) have confirmed that two-tone suppression originates in the cochlea.

1.3.3 Improved temporal resolution

The cochlear nonlinear component is extremely fast. Based on psychophysical data Smoorenburg (1972) reported that the nonlinearity adapted well within 20 ms from onset of a stimulus. Modern measurements of basilar membrane vibration (Recio et al.,1998) found that input-output functions became nonlinear within 100 μs from onset.

Fast compression can influence the amount of forward masking, which is a measure that shows the recovery speed from stimulation by a previous masker (i.e., temporal resolution). If we present a masker, and after this masker a test signal, the signal will be masked if it falls in the temporal "shadow" of the masker. This effect is illustrated in the left panel of Figure 1.6.

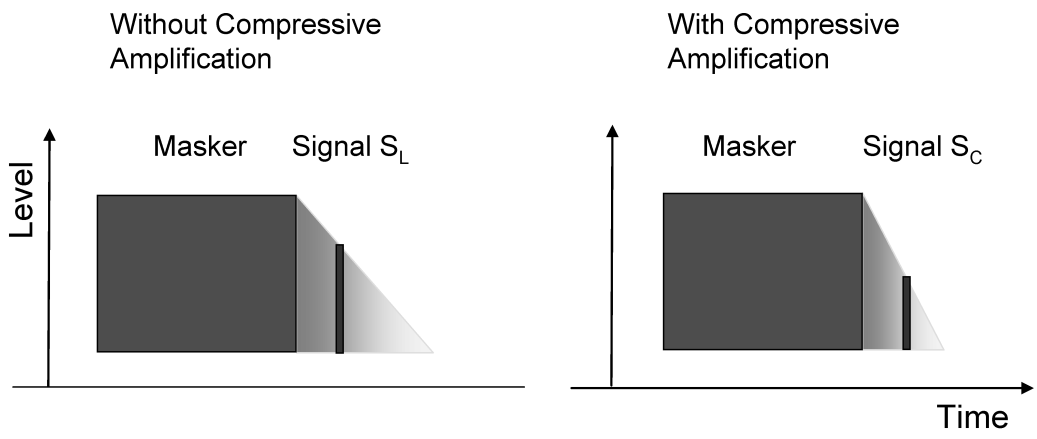


Figure 1.6: *Illustration of the influence of a fast-acting compressive nonlinearity on temporal masking. The left panel shows the situation without compressive amplification (i.e., linear amplification), the right panel with compressive amplification. S_L and S_C represent the signal level that is just masked for the system without and with compression, respectively. Compression results in more gain for S_C and therefore $S_C < S_L$.*

Because the cochlear amplifier operates very fast, there will be no interaction between the amplification of the two sounds. Usually, the signal is presented a short time after the masker (forward masking), and therefore needs a lower level for detection than the level of the masker. Due to independent compressive amplification for masker and signal, this leads to more amplification for the lower-level signal than for the higher-level masker. Compressive amplification will therefore improve detection of the signal. Thus the effective shadow of the masker is diminished by the fast amplification, see the right panel of Fig. 1.6. In effect, the cochlear amplifier decreases temporal masking and thus increases temporal resolution. Because forward masking is strongly influenced by compression, it can be used to estimate the amount of cochlear compression, see section 1.5.

1.3.4 Distortion (combination tones)

The cochlear amplifier doesn't only greatly improve many aspects of hearing, it also introduces distortion. In fact, it has been known for a long time that the human ear is subject to intermodulation distortion (Helmholtz, 1875). When the ear is presented with two tones (for instance f_1 and f_2), it produces other tones which were not present in the acoustic stimulus. These new tones are called combination tones because their frequencies correspond to the original stimulus frequencies (e.g., $f_2 - f_1$, $2f_2 - f_1$, etc.). It was this distortion that led to the early hypotheses of the existence of an active non-linear amplification mechanism (Goldstein, 1967) of an amplitude-compressing type (Smoorenburg, 1972). Nowadays, distortion products can be used for screening of hearing ability, since the presence of distortion products indicates that the cochlear nonlinearity is functional. Distortion products can also be used to estimate the amount of cochlear compression (see section 1.5).

1.4 Perceptual consequences of outer hair cell loss

1.4.1 Lower sensitivity and reduced dynamic range

One of the most noticeable effects of cochlear damage is an impaired ability to hear low-level sounds. The solution, of course, is to amplify the incoming sounds to the point where they become audible for the hearing impaired. However, simple amplification can quickly lead to uncomfortably loud sounds, and even to more hearing loss.

Normal cochlear amplification is largest for low input levels, and it decreases for higher level sounds. The pathological absence of cochlear amplification therefore results in inaudible low-level sounds, while the loudness of high-level sounds stays (near) normal, i.e., the dynamic range of hearing becomes smaller. Figure 1.7 schematically illustrates this reduced dynamic range due to loss of cochlear compression.

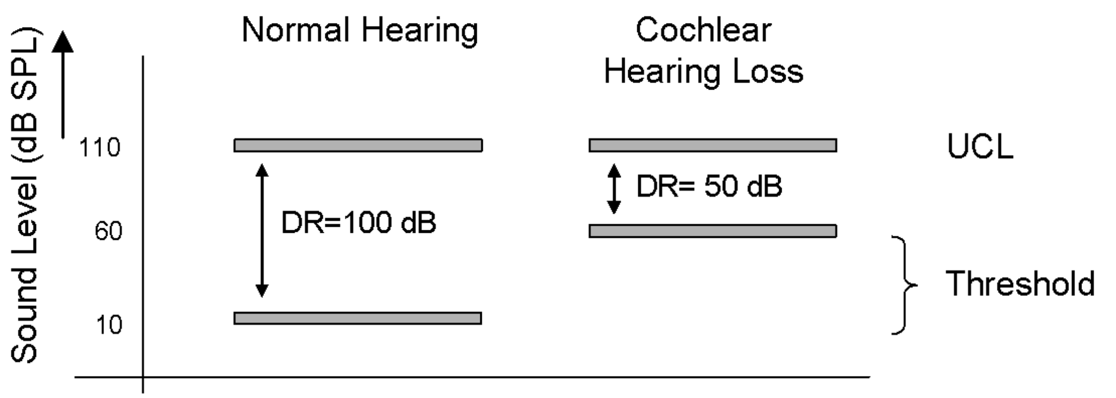


Figure 1.7: *Schematic illustration of a reduced dynamic range (DR) of hearing due to cochlear hearing loss. The ordinate denotes the input sound level. Although cochlear hearing loss leads to higher thresholds, the uncomfortably loud levels (UCL) can stay unaffected. This can easily lead to the classic scenario of an individual saying "Speak up, I can't hear you", and in the next breath: "There's no need to shout, I'm not deaf".*

Elevated thresholds combined with near normal uncomfortable levels (UCLs) cause an abnormal loudness growth (recruitment). Recruitment can be very awkward; a little more vocal effort can turn inaudible speech into uncomfortably loud speech (see Fig. 1.7). A solution lies in hearing aids with compressive amplification: low-level sounds should receive more amplification than high-level sounds. Figure 1.8 illustrates the benefit of compressive amplification for alleviating cochlear hearing loss. Note that the effect of compression on the speech signal depends on the type of compression, see section 1.6 and Fig. 1.16.

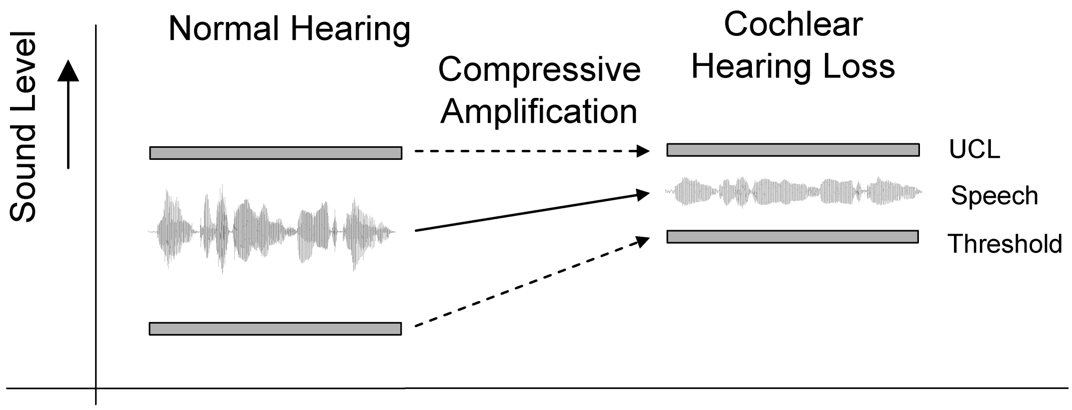


Figure 1.8: *Speech hardly fits into a small dynamic range. Compression of the speech signal can place speech within the residual dynamic range of the hearing-impaired listener.*

1.4.2 Reduced frequency resolution

Section 1.3 described that the cochlear amplification component is tightly linked to the sharp frequency tuning on the basilar membrane. Damage to cochlear amplification will therefore inevitably lead to deteriorated tuning. This has been confirmed by auditory nerve and basilar membrane measurements which show that damage to the outer hair cells results in a deterioration of the sharp tuning peak. Unfortunately, it is not possible to fully compensate for this decrease in frequency resolution by signal processing in hearing aids, as was shown by computer simulations of outer hair cell damage (Giguère and Smoorenburg, 1999). As a result, experiments that try to enhance spectral contrast of speech have shown only small improvements in speech recognition (Lyzenga, 2002).

1.4.3 Temporal resolution

If cochlear filtering would behave as if it were a simple resonant filter system, we would expect a trade-off between frequency and temporal resolution. A decrease in frequency resolution would then result in better temporal resolution (with respect to the fine structure of a speech signal) (Smith, 1997). However, this is generally not the case. The most likely reason is that the temporal abilities of the auditory pathway beyond the cochlea are matched to normal temporal resolution (Oxenham, 2003). An increase in cochlear temporal resolution might be limited by the rest of the pathway. Experiments showed that sensorineurally hearing-impaired listeners had worse recovery of forward masking than normal hearing listeners. In the experiments, the deterioration could be explained completely by the loss of fast cochlear compression (Oxenham, 2003).

1.5 Estimates of cochlear compression

This section gives an indication of the amount of cochlear compression. The amount of compression is expressed as a "compression ratio" (CR), i.e., the change in input gain (in dB) divided by the change in output gain (in dB). For an input-output graph such as Figure 1.4, where both axes have a logarithmic scale, the compression ratio is represented in the slope of the curves. See section 1.6.1 for a more elaborate description of compression ratio. We adopt CR because this is a common measure in the hearing aid industry. In contrast, many physiological investigations express the amount of compression as a "growth rate", which is the reciprocal of the compression ratio (growth rate = 1/CR).

1.5.1 In vivo measurements

Robles and Ruggero (2001) reviewed available data on basilar membrane responses that were measured at basal sites of the cochlea. They reported response growths that were averaged over the total input range (i.e., one value for the entire range, their Table 1). Figure 1.9 shows a graphical representation of their findings; the presentation is analogous to Figure 1.4.

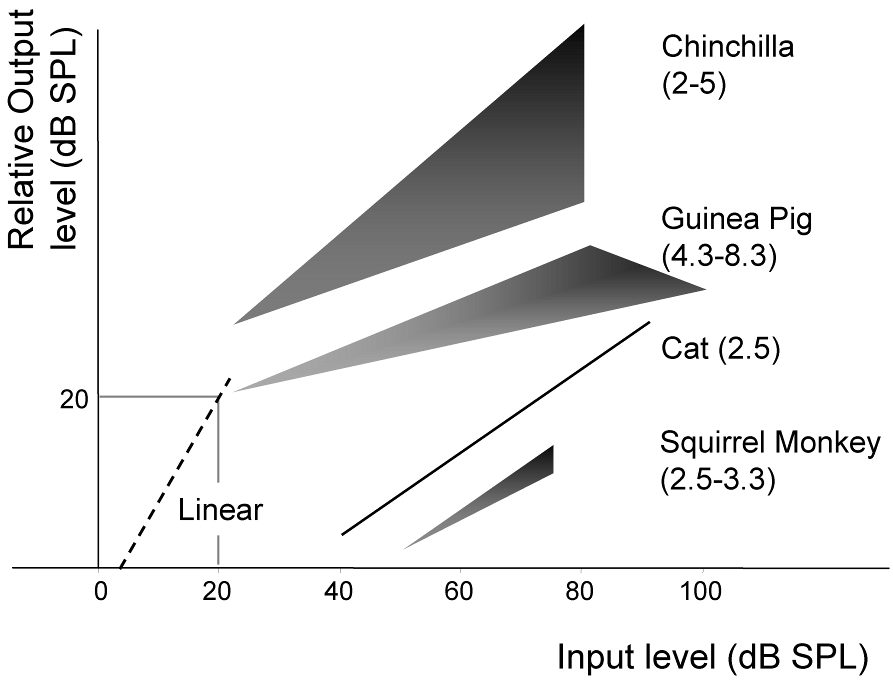


Figure 1.9: *Range of experimentally found basilar membrane input-output curves, for different species. The figure is an adaptation from the literature review by Robles and Ruggero (2001, data from Table 1). The slope of the curves and the corresponding values to the right show the average compression ratio. Data was obtained from in vivo experiments at basal sites of the cochlea. The line in the lower left corner represents a linear response (CR = 1). For clarity, data for the different animals are shifted along the ordinate.*

Overall, the figure shows that the measured compression ratios ranged between 2 and

8. Most compression ratios centered around 3, with the exception of the guinea pig for which much higher compression ratios were found. Generally, the amount of compression was less at more apical sites on the basilar membrane (i.e., for lower characteristic frequencies). It is very difficult to measure apical basilar membrane responses, and inflicted cochlear damage quickly leads to linear responses. Therefore, data for sites on the basilar membrane near the apex is scarce. Nevertheless, a few papers reported compression ratios of about 1.25–2 for the chinchilla and a linear response (CR = 1) for both the guinea pig and squirrel monkey (Robles and Ruggero, 2001).

1.5.2 Dynamic range

Besides animal models, it is also possible to obtain a rough estimate of the overall amount of compression by using its influence on the dynamic range. For instance, it is often assumed that complete loss of outer hair cells causes a threshold elevation of about 50 ($\pm$10) dB HL, while the upper limit of hearing remains normal (110 $\pm$ 10 dB HL). This implies that the OHCs reduce the dynamic range with a CR of about 2 ($\pm$0.5).

1.5.3 Forward masking

Smoorenburg (1974) used forward masking (pulsation technique of Houtgast, 1972) to obtain an estimate for the compression ratio in humans. He measured a compression ratio of about 2 for a CF of 1 kHz. Another approach is to compare the forward masking results for on-frequency and off-frequency maskers. The two most used methods are growth of masking (GOF, Oxenham and Plack, 1997; Oxenham and Bacon, 2003) and temporal masking (TM, Nelson et al., 2001). Both methods assume that a masker with a frequency far away from the characteristic frequency yields a linear response. However, it is not clear that this assumption is valid, especially for low CFs (Lopez-Poveda et al., 2003). This makes the results somewhat difficult to interpret. Moreover, for both methods it is not yet established if the observed compression can be completely ascribed to cochlear compression or to other mechanisms. Estimates based on GOF and TM show compression ratios from about 2 to 5.5 (Oxenham and Plack, 1997; Lopez-Poveda et al., 2003; Plack and Drga, 2003; Rosengard et al., 2005).

1.5.4 Distortion products (combination tones)

Cochlear distortion products can be used to obtain an estimate of the amount compression in the human cochlea. Smoorenburg (1972, 1974) used the level of the combination tone $2f_1$-f_2 in relation to level of the primary tones f_1 and f_2. He obtained an estimate of compression ratio of about 2 (f_1 = 1 kHz). Neely et al. (2003) re-evaluated previously obtained data from several recent distortion product measurements, and found CRs ranging from 1 to 4, for low (0–20 dB SPL) to high (50–70 dB SPL) input levels, respectively. A similar range (1.7–5) has been obtained from tone-burst and transient evoked oto-acoustic emissions (Harte and Elliott, 2005).

Evaluating the different estimates, we can assume that cochlear compression in the human ear is, on average, about 3 ($\pm$2) dB).

1.6 Compression in hearing aids I. Basic properties

Modern hearing aids counteract consequences of sensorineural hearing loss with compressive amplification. Such a compression system is an amplifier with high gain for low input levels, and with low gain for high input levels. The amount of gain is controlled by the envelope of a signal, leaving the fine structure unaffected.

Compression in hearing aids can be characterised by three main parameters, namely compression ratio, time constants and number of channels. The next three sections explain these parameters.

1.6.1 Compression ratio

Compression ratio (CR) is defined as the change in input gain (in dB) divided by the change in output gain (in dB). Compression ratio is thus a measure of the reduction in gain. A compression ratio of 2, for instance, indicates that an increase in sound level of 2 dB will lead to an increase in output level of only 1 dB. This process is shown in Figure 1.10 which illustrates the static properties of compression.

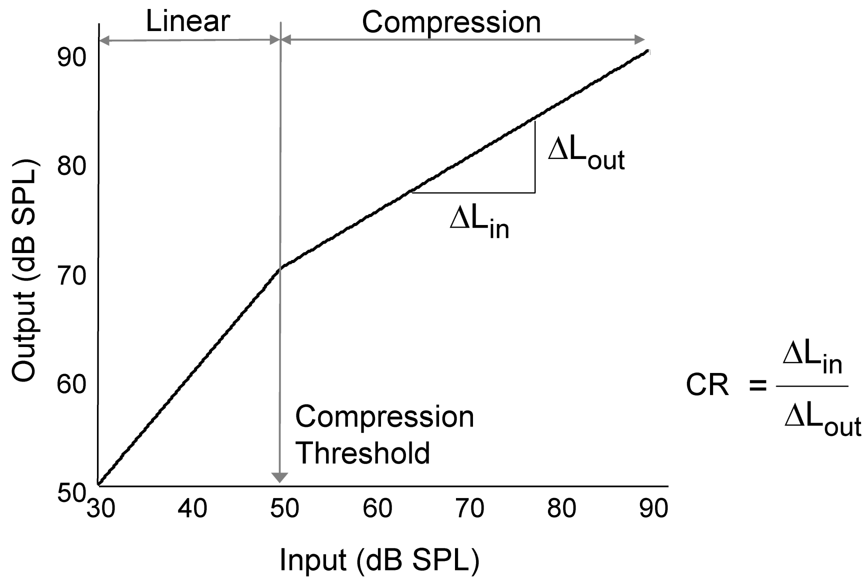


Figure 1.10: *Input-output curve of a compression system. Compression ratio (CR) is defined as the change in input level (ΔL_{in}) relative to the change in output level (ΔL_{out}).*

Compression is only active above the compression threshold. Figure 1.10 indicates that for input levels below threshold, the transduction is linear. As a side note, in our experiments (Chapters 2, 3 and 4) we ensured that all input levels were above the compression threshold. Some commercial hearing aids use additional expansion for extremely low input levels (<10 dB SPL) to block low level microphone noise, and severe compression for high input levels (>90 dB SPL) to prevent the amplifier from peak clipping.

1.6.2 Time constants

The speed at which a compression hearing aid reacts to changes in input level is characterised by its time constants. Figure 1.11 clarifies the definition of attack and release time. Attack time is the time it takes the gain to adjust to a sudden increase in input level (for instance caused by a door slamming shut). Similarly, release time is the time it takes the gain to adjust to a sudden decrease in input level (for instance caused by an engine switching off). Attack and release times can be varied independently. Due to non-zero time constants it takes the compressor some time to adjust the gain. During the adjustment the compression system is in a state of overshoot (too much gain), or undershoot (not enough gain).

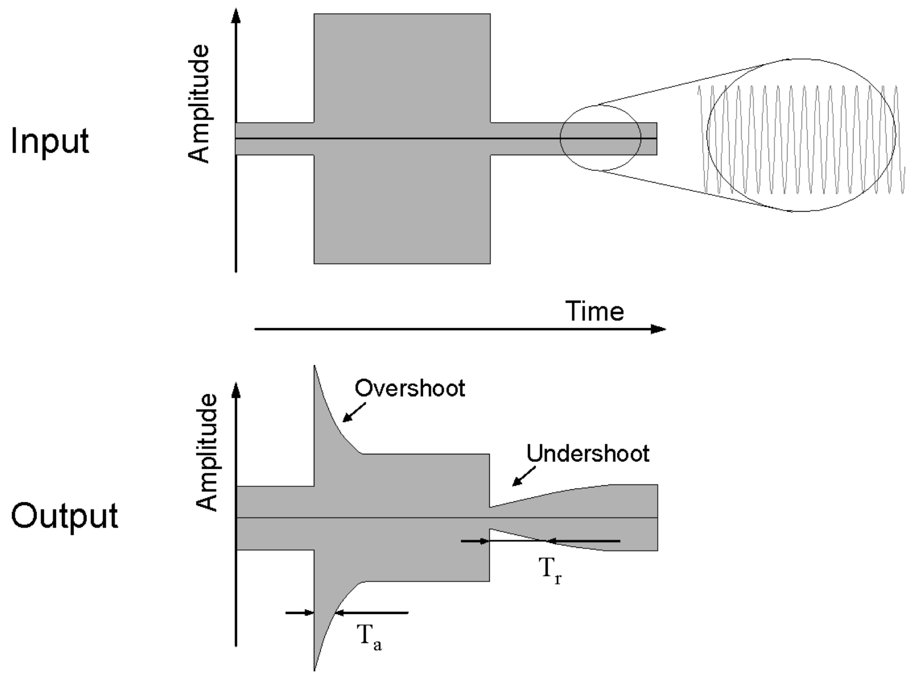


Figure 1.11: *Dynamic characteristics of compression. The upper panel shows the envelope of an input signal, in this example a sine wave which is modulated by a square wave. Compression in hearing aids typically reacts to level changes of the envelope of the input signal, and not to the fine structure (the sine wave). The lower panel shows the corresponding output signal.* T_a *= attack time;* T_r *= release time.*

1.6.3 Several definitions of time constant

When comparing the effects different compression systems, it is important to know how attack and release time are defined, since they describe the reaction of a compressor to changes in signal amplitude. Currently, several definitions of attack and release time are in use. The two standards that are most frequently used are an international/European (IEC 60118-2, 1983) and an American (ANSI S3.22, 1996) standard. Both the IEC 60118-2 (1983) and the previous (obsolete) ANSI S3.22 from 1987, define attack and release time according to a change in input level from 55 to 80 dB (and vice versa) of a sine wave with a frequency of 2 kHz. Both attack and release time are defined with respect to a level 2 dB off the steady state output level. However, the new ANSI S3.22 (1996) defines attack and release time for a change in input level from 55 to 90 dB (and vice versa) for a sine wave of arbitrary frequency. Attack time is defined as the time it takes the output level to reach 3 dB above steady state output level, release time is the time it takes to reach 4 dB below steady state output level.

Both IEC and ANSI define time constant as the time it takes the output level to reach a certain distance from the final steady state level. This type of definition is ambiguous because the value of time constant is not only influenced by the speed of the peak detector (the RC constant), but also by the amount of overshoot. This implies that the behaviour of two systems with the same IEC/ANSI time constant can be very different (Kates, 1993). For instance, one system with a low threshold and fast compression, and another system with a high threshold and slow compression, might have the same release time constant when measured over the entire input range. Of course, this is remedied by IEC and ANSI through the specification of fixed input levels for time constant measurements.

In contrast to IEC and ANSI, we define time constant mathematically as the time it takes for the output level to reach 1/e ($\approx$ 37%) of the ultimate change in gain. Figure 1.12 illustrates the difference between our and the IEC definition.

From the figure it can be seen that the definitions are fundamentally different. Both the IEC and ANSI definitions are based on the tail of the response, while our definition is based on the initial part. Consequently, the ratio of time constants according to the IEC/ANSI definition and ours is not constant, but depends on the amount of overshoot (and therefore on the combination of input signal, compression ratio and threshold). A change in input from 55 to 80 dB with CR = 2 results in an overshoot of the output level of 12.5 dB. Using this overshoot and assuming logarithmic decay for an IEC system, the ratio between our (T_{exp}) and the IEC definition (T_{IEC}) is $T_{exp}/T_{IEC} = 0.4$ for both attack and release time. For an ANSI (1996) system the ratio for the same input rise is $T_{exp}/T_{ANSI} = 0.5$ for attack time and $T_{exp}/T_{ANSI} = 0.6$ for release time. Since these ratios depend on overshoot, they differ for different input signals and compression ratios. An overshoot of 6.25 dB (25 dB input change with CR = 4) and 16.7 dB (25 dB input change with CR = 1.5) results in $T_{exp}/T_{IEC} = 0.7$, and $T_{exp}/T_{IEC} = 0.3$, respectively. In comparing our results to those specified in terms of T_{IEC} or T_{ANSI} we suggest to use the

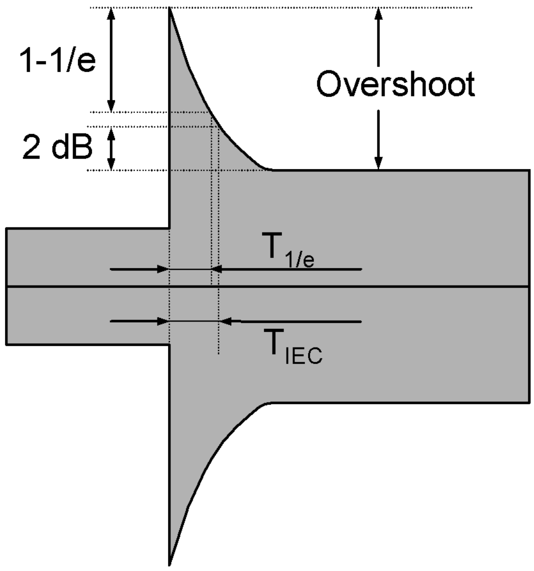


Figure 1.12: *Different definitions of time constant.* T_{IEC} *represents time constant as defined by IEC 60188-2;* $T_{1/e}$ *is our mathematical definition.*

factor of 0.4. See section 1.7 for a description of time constants in common compression systems.

1.6.4 Number of channels

Instead of applying compression to the entire signal (single-channel or wide-band compression), it is also possible to compress spectral parts of the signal in smaller frequency bands (multi-channel compression). Figure 1.13 illustrates the difference between single-channel and multi-channel compression.

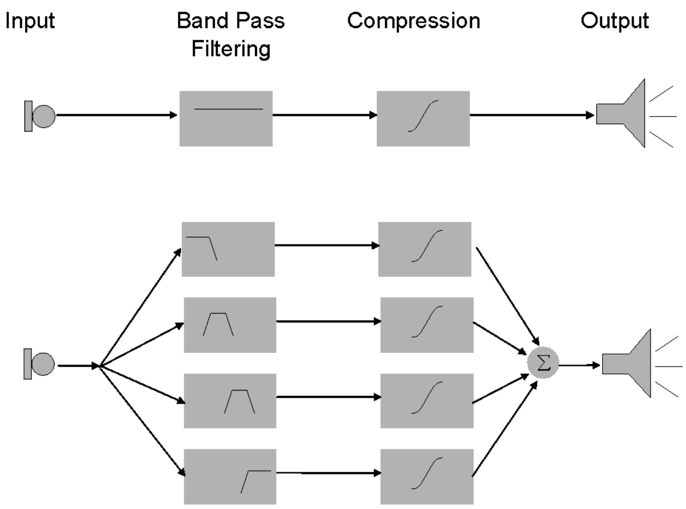


Figure 1.13: *Schematics of compression. The top panel shows single-channel compression, the bottom panel shows multi-channel compression.*

In multi-channel compression the input signal is band-pass filtered, and the signal in each channel is independently compressed. After compression the outputs of the channels are summed and presented to the listener. A major advantage of multi-channel compression is the possibility to use fitting methods based on the individual's frequency dependent hearing loss. Multi-channel compression also leads to less temporal distortion because it acts on small frequency bands and not on the broad frequency spectrum as is the case for single-channel compression. For instance, single-channel compression of a speech signal implies that the low-frequency components which dominate the spectral energy distribution will control the compressor. High-frequency speech components will therefore receive amplification that is controlled by the energy of the low-frequency components of the speech. Figure 1.14 illustrates this effect for an input signal which consists of two modulated tones (f_1 and f_2).

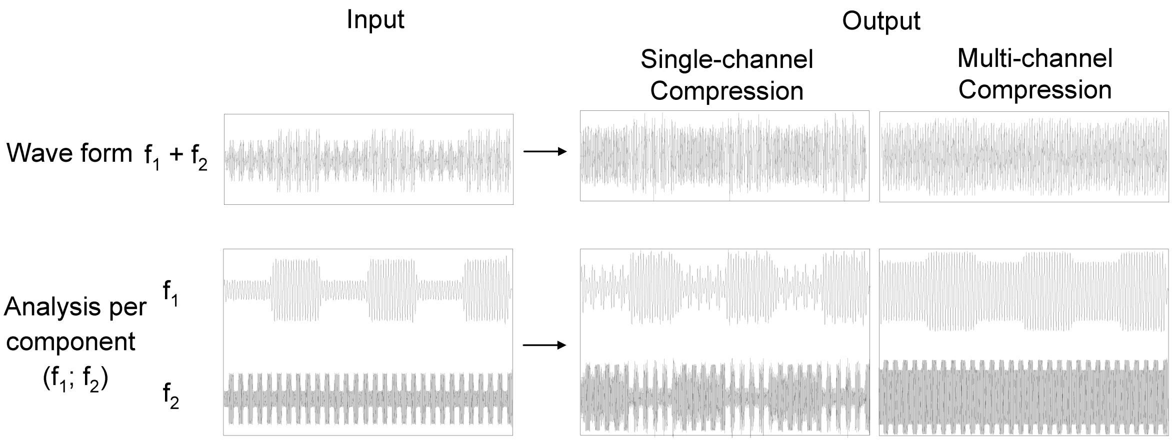


Figure 1.14: *Example of temporal interaction due to compression of a single input signal which consists of two modulated tones (f_1 and f_2). The panels show signal amplitude (ordinate) versus time (abscissa). The left panels show the input signal: two sine waves (375 and 3000 Hz) that are both square wave modulated (10 and 100 Hz, respectively). The middle and right panels show the output of single-channel, and multi-channel compression, respectively. The top panels show the actual signals. The bottom panels show an analysis in which each component is displayed separately (two-channel analysis). Overshoot and undershoot are absent due to instantaneous compression (i.e., $T_a = T_r = 0$).*

The output of the single-channel compressor (Fig. 1.14, middle panels) shows that f_1 became modulated by the original modulations of f_2 (and vice versa). Multi-channel compression suffers much less from this temporal interaction, since the signal in each channel receives its own gain, independently from the signals in the other channels (see the signals in the right panels of Fig. 1.14).

However, the use of many channels can introduce a degradation of spectral contrast. Take for instance two speech components (e.g., the frequencies f_1 and f_2 from Fig. 1.14, or similarly, the first and second formant of a vowel F_1 and F_2, which both happen to fall in the same compression channel of our multi-channel system. In that case, the two components will receive the same amount of gain, thus keeping the intensity difference (contrast) between them in tact. If these same components each fall in separate

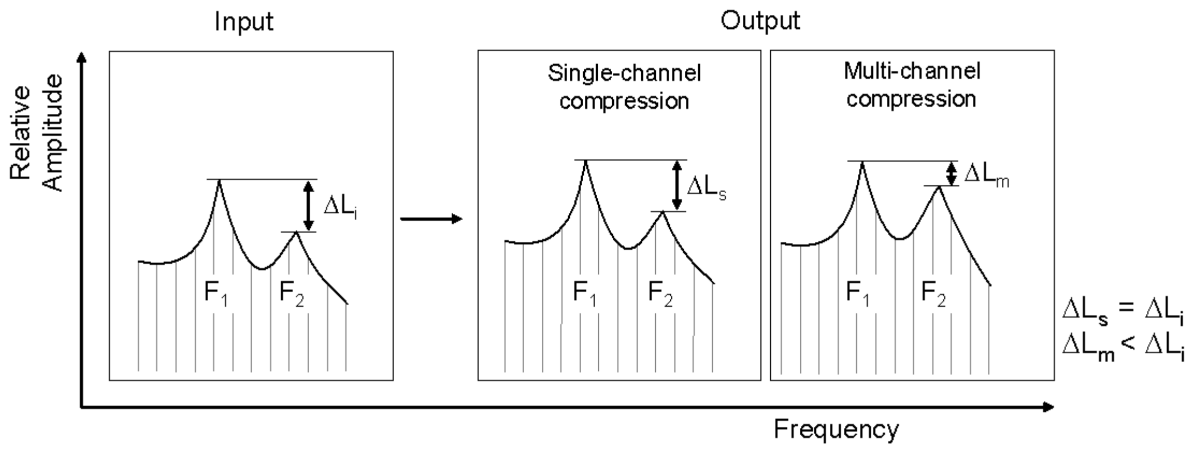


Figure 1.15: *Example of reduced spectral contrast due to multi-channel compression. Panels display the relative level (ordinate) of two formants (F_1 and F_2). The left panel shows the input signal. The middle panel shows the output after single-channel compression. Spectral contrast after single-channel compression (ΔL_s) is equal to the spectral contrast of the original input signal (ΔL_i). The right panel shows the effect of multi-channel compression in which F_1 and F_2 each receive a different amount of gain because they fall into different channels. The spectral contrast for multi-channel compression (ΔL_m) is lower than the contrast for single-channel compression (ΔL_s).*

channels, they most likely will receive a different amount of gain, appropriate for their respective channel, and this will reduce spectral contrast. Figure 1.15 illustrates this process.

An alternative way of looking at this is the following. Instead of thinking about the influence of multi-channel compression as a reduction in spectral contrast, one may realize that multi-channel compression in a hearing aid restores contrast to (near) normal. The hearing aid replaces the effect of the normal (multi-channel) cochlear nonlinearity. In fact, compensation of hearing loss by single-channel compression will lead to an increase in spectral contrast relative to normal hearing. This increased spectral contrast might actually be beneficial for speech intelligibility, and might to some extent alleviate the reduced frequency selectivity. On the other hand, increasing the number of channels might decrease temporal interaction (see Fig. 1.14) and this could compensate the decrease in spectral contrast. Compensating for reduced cochlear compression by multi-channel compression in hearing aids therefore leads to a trade-off between the two effects. The optimum number of channels for alleviating hearing loss should be determined experimentally, see Chapters 2, 3, and 4.

1.7 Compression in hearing aids II. Implementations

Compressive amplification systems are very versatile, and many different types of systems have been designed. This section gives a short overview of different rationales underlying the application of compression in hearing aids. For the general description in this section, the mentioned values of time constants are appropriate for the IEC and ANSI definitions (see Section 1.6.3).

1.7.1 Compressive limiting

Compression can be used for output limiting to prevent the ear from being stimulated by too high (peak) sound levels, while minimizing distortion. Compressive limiting is usually characterized by a short attack time (< 5 ms), high compression thresholds (> 90 dB SPL) and a high compression ratio (> 5). These characteristics imply linear amplification for most input levels, including normal speech levels.

1.7.2 Automatic gain control

Automatic gain control (AGC, also called automatic volume control, AVC) is a form of compression that is used to keep the long-term average presentation level near the level corresponding to maximum intelligibility. This type of compression is typically used for people with only a slightly reduced dynamic range. AGC can adjust the overall level of speech such that it is at the preferred level and not too soft or too loud for optimal speech understanding. An AGC system reduces the need of a volume wheel since the gain is automatically changed over a wide range of input levels. AGC is characterized by low compression thresholds (< 50 dB SPL), moderate compression ratios (< 5) and attack and release times that are comparable to the length of a phrase or sentence (between approximately 150 ms and several seconds). In practice, many AGC systems use a short attack time to allow the system to quickly react to sudden increases in input. Some AGC systems obtain the same effect by incorporating an additional compressive limiter.

1.7.3 Syllabic compression

In contrast to AGC, syllabic compression (SC) is used to alter the short-term intensity relations among speech elements to improve intelligibility. The objective of syllabic compression is to compress the dynamic range of the speech signal in order to present the speech within the reduced dynamic range of the patient. Figure 1.16 shows the different effect of AGC and SC on a speech signal.

Because syllabic compression is fast, it can fit all relevant speech aspects (syllables, or even phonemes like consonants and vowels) into the dynamic range of the hearing-impaired listener and it can restore an abnormal loudness growth (recruitment) to normal. Syllabic compression is characterized by short attack and release times (< 150 ms), a low compression ratio (< 4) and a low threshold (< 50 dB SPL). If the time constants are comparable to the duration of phonemes (< 25 ms), this type of compression is also called phonemic compression.

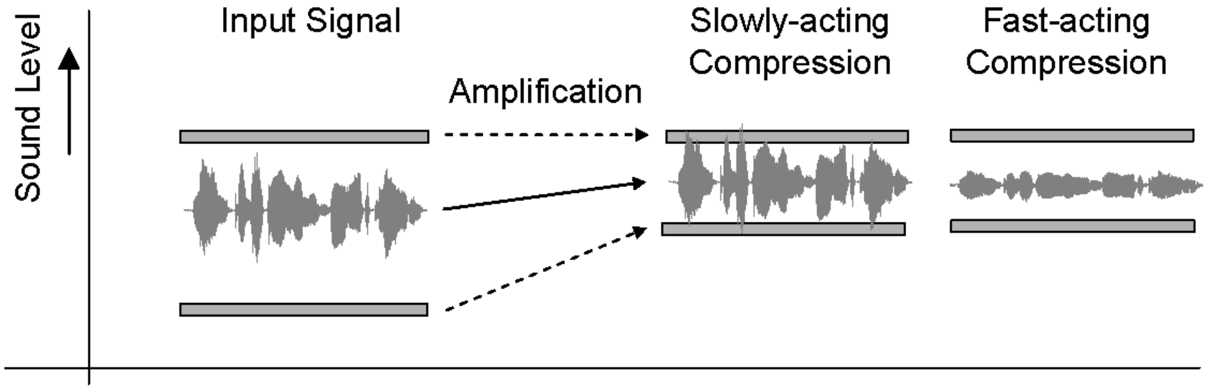


Figure 1.16: *Both slowly-acting and fast-acting compression lift the speech above the elevated threshold. Fast-acting compression additionally reduces the dynamic range of speech itself.*

Of the described compression types, syllabic (or phonemic) compression is most similar to the physiological compressive nonlinearity. Both act fast, and have a compression ratio of about 2 to 3. Syllabic compression is therefore a prime candidate for restoring hearing abilities. However, as indicated in section 1.4.2, compression can not restore impaired frequency selectivity, and this inability might limit the possible benefit from syllabic compression.

The compression used in the experiments of Chapters 2, 3, and 4, can be characterized as ranging from syllabic compression to automatic volume control. All stimuli were presented within the dynamic range of hearing of the individual subject and compression limiting was not needed.

1.7.4 Other types of compression

Several other types of compressive amplification have been designed. Such systems often incorporate aspects of the types of compression mentioned above, or use different techniques. For reference, some examples are given below:

- Some (commercial) hearing aids have an adaptive release time. If the compression threshold is exceeded for a longer period of time (not just momentarily), the release time will be lengthened. For signals above the compression threshold, these hearing aids act as automatic gain control (Dillon, 1996).

- Hearing aids can also incorporate different compression ratios, depending on the signal level (curvilinear amplification). An example of this is the “K-amp”, developed by Killion (1993). A K-amp consists of four amplification stages: linear with high gain for low signal levels, compression for moderate signal levels, linear gain for high signal levels and compression limiting for the highest signal levels. In addition, the K-amp uses high-frequency emphasis for low signal levels.

- Specific systems have been designed to reduce overshoot. In a compression system an abrupt increase in input level causes a signal overshoot because the system needs time to adjust the gain. The overshoot is partly caused by a delay of the gain signal relative to the audio signal (due to low-pass filtering of the gain signal). Some compression systems synchronise the gain and audio signals by introducing a delay to the audio signal, thus reducing signal overshoot.

- Hearing aids can apply compression to specific frequency ranges only. For example, 'treble increase for low levels' (TILL) uses more compression at high frequencies than at low frequencies, or 'bass increase for low levels' (BILL) uses more compression for the low frequencies than for the high (Killion, 1990; Dillon, 1996).

- Moore and Glasberg (1988) have developed a dual time constant hearing aid that incorporates two sets of time constants. A slow-acting control voltage (release time $= 5$ s) keeps the average signal level at the output constant, regardless of the input level. A fast-acting control (release time $= 150$ ms) reduces the gain in response to sudden high intensity sounds. When the sudden sound has diminished, the gain will return to the value set by the slowly-acting component.

1.8 Scope and overview of this thesis

Cochlear compression is extremely important for hearing acuity. Consequently, loss of cochlear compression often leads to severely impaired speech intelligibility. To compensate for this loss, hearing aids can (and should) incorporate compressive amplification. However, due to the nature of cochlear hearing loss it is not sufficient to replace lost cochlear compression by compressive amplification in a hearing aid. Cochlear compression is inextricably linked to several perceptual phenomena, and although compression in hearing aids can counteract impaired thresholds and decreased dynamic range, it can not compensate for reduced frequency resolution. The choice of hearing aid compression parameters leads to a trade-off between different effects. Currently, the optimal compression characteristics remain unclear. This thesis investigates the effect of combinations of compression characteristics on speech intelligibility.

The main research question of this thesis is "What is the influence of number of channels, compression ratio, and time constants on speech intelligibility in noise for listeners with a moderate sensorineural hearing loss?". The focus lies on the effect of compression on speech intelligibility and not on speech audibility. In the study we will simultaneously investigate the effect of the different compression parameters. The advantage of such a complete parametric design is the possibility to determine interaction effects between number of channels, compression ratio, and time constants.

Chapter 2 investigates the effect of compression on speech intelligibility in stationary noise. Due to the temporal aspects of compression, we expect that the effect of compression will be different for a fluctuating background. Therefore, Chapter 3 extents the research to speech in fluctuating noise. Chapter 4 investigates possible predictors for the effect of compression for an individual user. Finally, in Chapter 5 the thesis concludes with a summary and a discussion on clinical applicability of the results.

The effect of compression on speech intelligibility in stationary noise

2

Rolph Houben and Guido F. Smoorenburg

Manuscript in revision for Audiology & Neurotology

2.1 Introduction

Cochlear hearing loss results in a reduced dynamic range of hearing (Chapter 1), and this loss can be alleviated by compression hearing aids. Compression can lead to better speech intelligibility when it is used to amplify low-level (inaudible) speech (Souza, 2002), or when it is used to attenuate high-level (uncomfortably loud) speech (Dillon, 1996). As a result, many modern compression hearing-aids use compression over a wide range of input levels.

In quiet, compression only slightly influences speech intelligibility for speech that is presented at a comfortable listening level (Dillon, 1996), for instance in a one-on-one conversation in a quiet room. If speech is already audible, compression does not improve or degrade speech intelligibility substantially. However, communication in a noisy background is generally much more difficult and in this situation compression often leads to worse speech intelligibility than linear amplification. The effect of compression characteristics on speech intelligibility is still under debate, especially for the situation when the speech levels are comfortable and compression is not needed to make the speech audible.

In this chapter we focus on the effects of compression on speech intelligibility (and not on speech audibility) in a stationary background noise. Previous studies focused mainly on one or two parameters per experiment. This study focuses on the combination of four parameters, namely the number of channels, the compression ratio, and the attack and release time. The study is designed to investigate possible interaction effects of the parameters with respect to speech intelligibility in stationary noise.

2.1.1 Previous research on amplitude compression

The body of research on compression is quite large, therefore we will first look at several literature reviews.

Reviews

Rintelmann (1972) presented a chronological review of experiments on compression for the hearing impaired. At the time of his review, compression was predominantly used in amplitude limiting circuits. Only few studies had used compression over a wide dynamic range. These studies showed a large variability between subjects; in quiet most of the subjects showed improved speech intelligibility with compression relative to linear amplification. Rintelmann stressed that future evaluations should not only test speech intelligibility in quiet, but also in various types of background noise.

Villchur (1978) focused on both single- and multi-channel compression. He looked only at fast systems with an attack time of ≤ 15 ms and a release time of ≤ 40 ms. Unfortunately, he did not make a distinction between speech intelligibility in quiet and in background noise. Villchur concluded that the experiments showed seemingly contradictory results. Many compression experiments had failed to improve speech intelligibility relative to linear amplification. He proposed several possible explanations for this poor result. Next to improper speech material, improper selection of subjects or poor post-compression frequency equalization, he argued that multi-channel compression might be better than single-channel compression.

Braida et al. (1979) differentiated only between slow (attack time between 10 and 200 ms, release time between 200 and 3000 ms) and fast compression systems (attack time ≤ 25 ms; release time ≤ 50 ms). Although both categories included single- and multi-channel compression, they did not perform a systematic analysis of the effect of the number of channels. Furthermore, they did not evaluate the effect of compression ratio. Most included studies used moderate compression ratios ($CR \leq 5$,) with the exception of amplitude limiting systems ($CR > 5$). Finally, they made no clear distinction between results obtained in quiet and in noise. They concluded that there was hardly any data available regarding slowly-acting compression. With respect to fast-acting compression Braida et al. reached the same conclusion as Villchur: the experimental results seemed to be inconsistent. The reasons for the discrepancies were not thoroughly understood. At the time of this review, they concluded that no experiment had demonstrated clear advantages of either slowly-acting or fast-acting compression for hearing-impaired listeners.

Walker and Dillon (1982) compared the results of several studies that used commercially available hearing aids. They also evaluated a set of laboratory studies. They made a distinction between single-channel and multi-channel compression, and slow and fast compression systems. Throughout the review no systematic differentiation between experiments in quiet and in noise was made. The evaluation indicated that compression hearing aids failed to show any consistent advantages in terms of speech intelligibility or wearer acceptance above non-compressive hearing aids. The hearing aids that yielded the highest speech discrimination scores were predominantly the fast systems. Evaluating the laboratory studies they found some minor advantage for fast-acting single-channel compression. With fast-acting multi-channel compression the inter subject variability was large. Some subjects had shown a clear improvement in speech intelligibility as compared to linear amplification. Overall, they concluded that the acquired results provided only minor support for the use of fast-acting compression. They indicated that more work was needed in this area. With respect to slowly-acting compression, Walker and Dillon sided with Braida et al.: not enough data was available to reach a conclusion. Furthermore, they concluded that there had been very little attention paid to subjective measures like acceptability and pleasantness.

Preves (1991) found that several authors had shown little, if any, benefit when comparing fast-acting compression to linear amplification. He emphasized that fast-acting compression increases consonant energy and decreases vowel energy. In his view, more research was needed to determine whether this enhancement of the consonant to vowel energy ratio results in improved speech perception in both quiet and noise.

Hickson (1994) concluded that there was only minor support for the use of fast-acting single channel compression in hearing aids. Results when using slowly-acting single-channel compression in noise were seemingly contradictory. The benefit from multi-channel compression was not clear. Hickson loosely concluded that for multi-channel systems generally the best results had been obtained with a maximum of three channels. She didn't advise on preferable compression ratios or time constants. In her view, optimal compression characteristics were still to be determined.

Dillon (1996) categorized experiments according to frequency dependency of compression (compression for all, low, or high frequencies only) and amplitude dependency (compression of all, small, or large amplitudes only). He limited his evaluation to these two categories and did not differentiate on the basis of the number of frequency channels, compression ratio, or time constants. All papers included in Dillon's evaluation had a maximum of three channels. Compression ratios varied between about one and five. Most systems used fast-acting compression. Some systems used slowly-acting compression or compression with adaptive time constants (see section 1.7.4). Omitting amplitude limiting in Dillon's classification system, six categories remained, which included a total of 12 experiments. Dillon compared the results of the different systems by converting all data (i.e., percentage points or dBs) into rationalized arcsine units. For listening in quiet, Dillon found that the difference between the systems was quite small. The various compression schemes did not have much effect on speech identification in quiet if speech in the reference condition was presented at a comfortable listening level. In quiet, compression is beneficial if the absence of compression would result in speech at long-term levels that are lower than optimal. This means that both slowly- and fast-acting compression will improve intelligibility of low-level speech without degrading it for speech at a higher level. Thus, compression can be important for speech intelligibility if the acoustic environment contains widely varying sound levels. For intelligibility in noise, Dillon again found the differences in intelligibility between the different systems to be small, except for two systems that outperformed the others. The first system (Moore and Glasberg, 1986) consisted of a slowly-acting single-channel compressor (attack/release time = 8/150 ms), followed by fast-acting two channel compression (attack/release time = 3/10 to 3/50 ms). The second system (Moore and Glasberg, 1988) used a slowly-acting single-channel compressor (release time = 5000 ms) preceding a fast-acting single-channel compressor (release time = 150 ms). This dual time constant compression system was then followed by a high-frequency syllabic compressor. Moore and Glasberg (1988) attributed the high performance to the facts that the slowly-acting single-channel compressor kept the signal at a comfortable level so that the fast high-frequency compressor had to act over only a small dynamic range, thus minimizing distortion. Moreover, the number of channels was only two. They thought that this prevented the loss of spectral contrast. Dillon wondered why compression in the high-frequency channel could provide much better results than (frequency shaped) linear amplification, especially since the compressor was most of the time noise controlled (the signal-to-noise ratio was around -7 dB) and the speech-shaped noise would have kept the gain at a rather constant level. He concluded that these two systems were promising but that their results were not fully understood.

Souza (2002) concluded that wide dynamic range compression could lead to an advantage over linear amplification, especially for low-level speech in quiet. In a background noise, the benefit of compression and linear amplification was similar. However, she indicated that the benefit of compression might depend on the signal-to-noise ratio and on the noise type. Based on research that was published after Hickson wrote her review (Hickson 1994) Souza stated that compression with a maximum number of channels of four (and not three, as was suggested by the research available to Hickson) might give similar speech intelligibility as linear amplification (but not better than linear amplification). Souza described some studies that showed improved speech intelligibility for compression systems with many channels. However, she attributed this positive result on increased audibility and not on the large numbers of channels itself. Souza suggested using the lowest possible compression ratio that will maximize audibility for a broad range of input sound levels. For this compression ratio, the increased audibility of speech might be larger than the detrimental effects of compression. For a quick response to sudden increases in input level, attack time should be short. For release time the situation is more complicated since it might interact with compression threshold and ratio. Souza found no consensus for the best release time. For multi-channel compression with many channels, she suggested to use longer release times because this might prevent a degradation of speech intelligibility caused by temporal and spectral loss of contrast.

In summary, the reviews show that many experimental results were inconclusive or seemingly contradictory. The results seem to imply that a small number of channels (≤ 3 or 4) gives the best results. However, since most systems under investigation had only a few channels, the effect of many channels on speech intelligibility is still not clear. Most systems evaluated had compression ratios smaller than 5. Higher compression ratios seem detrimental. The reviews differentiated between fast and slow systems, but none appeared superior to the other. A clear distinction between compression in quiet and in noise was made by Hickson (1994) and Dillon (1996). This distinction is needed since in a noisy background it is possible that the noise, and not the speech, controls the compressor. From the reviews it can be concluded that compression in quiet does not have to be detrimental for speech intelligibility, and that it might even be beneficial. For speech in noise, the effect of compression on speech intelligibility was mainly detrimental. Dillon (1996) concluded that the effect on speech intelligibility in noise of the systems under investigation did not differ much from each other.

Since the effect of compression on speech intelligibility tends to be more detrimental in noise, speech in noise may be a better research tool than speech in quiet if one is interested in the effects of compression parameters on speech intelligibility. Moreover, speech in noise is closer to everyday listening conditions. According to a survey of hearing aid owners on what improvements they would like for their hearing aids, improvement of speech intelligibility in noise was the most desired one (Kochkin, 2002).

In conclusion, speech in noise can be a useful tool for investigating the effects of compression. However, from the reviews quoted above, it is not clear which parameter values should be used in the current study, including possible interaction effects. The next section expands on the reviews and classifies papers according to the number of frequency channels (NC) and the compression ratio (CR). Secondly, the results are arranged according to the time constants namely, attack and release time in ms (T_a/T_r, defined according to IEC 60118-2 or ANSI S3.22, see section 1.6.3).

Previous research classified parametrically

From the reviews it was concluded that many results with compression were seemingly contradictory. This might have been aggravated by the comparison of experiments which spanned too broad a range of relevant parameters, such as hearing loss, compression settings, types of noise, etc.

Research included in the evaluation given below concerned experiments with speech signals in a stationary background noise, unless stated otherwise. Compression thresholds were well below the input signal levels, which ensured full dynamic range compression. All research was conducted with *moderately* hearing-impaired subjects. Subjects had a dynamic range of hearing larger than 30 dB, i.e., larger than the typical dynamic range of speech. If the dynamic range was unknown, the inclusion criterion of a pure tone average hearing loss (PTA) smaller than 60 dB at 0.5, 1, and 2 kHz was applied. For research with profoundly hearing-impaired subjects we refer to other studies such as Boothroyd et al. (1988), Busby et al. (1988), Crain and Yund (1995), De Gennaro et al. (1986), Drullman and Smoorenburg (1997), Moore and Glasberg (1986), Neuman et al. (1994), Verschuure et al. (1993), Verschuure et al. (1998), Villchur (1987), and Yund et al. (1987).

For clarity, the compression parameters are indicated by $CP(NC, CR, T_a/T_r)$. Linear amplification is indicated by CR = 1. A question mark indicates that the parameter values were not specified by the authors. In addition, two symbols are used: - and +. The - symbol refers to a range used within a single experiment. If used for compression ratios, it indicates that frequency dependent compression ratios within that range were used. This can also be subject dependent. If used for time constants the - symbol represents a single experiment in which time constants differ across channels (multi-channel compression) or change over time (adaptive release time). The + symbol refers to individual parameter values. If used for several parameters it implies that all combinations were studied. For instance *CP(1 + 2, 2 + 3, 10/10)* gives 4 different experiments: *CP(1, 2, 10/10)*, *CP(2, 2, 10/10)*, *CP(1, 3, 10/10)*, and *CP(2, 3, 10/10)*. Whereas *CP(3, 1-3 + 3-5, 10/50-500)* represents only two experiments: *CP(3, 1-3, 10/50-500)* and *CP(3, 3-5, 10/50-500)*, with frequency dependent compression ratios ranging from 1 to 3, and 3 to 5, respectively and release times ranging from 50 to 500 ms.

The compression parameters and the results of the studies described below are summarized in Table 2.1. Additionally, the table gives the number of subjects that were included in the experiments.

Table 2.1: *Results of previous research with moderately hearing-impaired subjects and stationary noise. The reference is linear amplification. Research that has not used linear amplification is indicated by $\otimes$. Equal or worse scores than the reference are indicated by $\odot$, and $\ominus$, respectively. Better scores than the linear reference did not occur. Research that could not be classified in this table is not numbered but indicated by $\star$ in the list below. The number of subjects included (n) is given to the right of the name of the authors.*

	Single-channel (NC = 1)		Multi-channel (NC > 1)	
	CR ≤ 3	CR > 3	CR ≤ 3	CR > 3
$T_r \leq 25$ ms	$2\odot$	$4\ominus$	$2\odot\odot\odot$	$4\ominus$
	$3\odot$		$4\odot\ominus\ominus$	
	$4\odot$		$9\odot\ominus$	
$25 < T_r \leq 100$ ms	$1\odot$			
	$5\otimes$			
	$7\odot\odot\odot$			
$T_r \geq 100$ ms	$5\otimes$	$6\ominus\ominus$		
	$6\odot\odot\odot$	$8\ominus$		
	$7\ominus\ominus\ominus\ominus\ominus\ominus$			

1	Dreschler et al. (1984)	n = 12	8	Tyler and Kuk (1989)	n = 16
2	Moore et al. (1999)	n = 18	9	Barfod (1978)	n = 5
3	Maré et al. (1992)	n = 18	$\star$	Yund et al. (1987) CR = 1-7	n = 20
4	Van Buuren et al. (1999)	n = 26			
5	Neuman et al. (1995)	n = 20	$\star$	Yund and Buckles (1995; 1995a) CR = 1-7	n = 16
6	Neuman et al. (1998) *CP(1, 1+1.5+2+3+5+10, 5/200)*	n = 20	$\star$	Van Toor and Verschuure (2002) CR = unspecified	n = 38
7	Neuman et al. (1998) *CP(1, 1.5+2+3, 5/60+5/200+5/1000)*	n = 20			

Single-channel compression

CR ≤ 3

Dreschler et al. (1984) *CP(1, 1 + 3, 8/50)* did not find a significant difference in speech intelligibility between linear amplification and fast-acting single-channel compression.

Maré et al. (1992) *CP(1, 1 + 2, 10/20)* used compression with a delay to reduce overshoot (see section 1.7.4). Compression threshold was chosen at the rms of speech. Both speech reception threshold and nonsense CVC scores were slightly worse for compression than for (non frequency-shaped) linear amplification. However, it is not known if these differences were significant.

Van Buuren et al. (1999) *CP(1 + 4 + 16, 1 + 2 + 4, 0/0)* obtained speech reception thresholds for instantaneous compression. For single-channel compression and a compression ratio of 2, the speech reception threshold did not differ significantly from results obtained with linear amplification.

Moore et al. (1999) *CP(1 + 2 + 4 + 8, 1 + 1-2.9, 7/7)* used compression with a delay to reduce overshoot. Compression ratio was subject and channel dependent. The background noise had spectral gaps. No significant differences in speech reception thresholds were found between linear amplification and single-channel compression.

Neuman et al. (1995) *CP(1, 1.5 + 2 + 3, 5/60 + 5/200 + 5/1000)* used paired-comparison judgements to investigate the effect of release time on sound quality. They did not find a significant preference for release time with respect to speech in noise. However, this might be due to (ventilation) noise that was presented 14 dB below compression threshold. Moreover, the signal-to-noise ratio was very high, +35 dB. Their experimental design did not allow an analysis with compression ratio as main effect.

Neuman et al. (1998) described two experiments. For both experiments they used the same subjects and noise as Neuman et al. (1995). In the first experiment *CP(1, 1 + 1.5 + 2 + 3 + 5 + 10, 5/200)* they varied compression ratio with only one T_a/T_r-setting. Quality ratings with all types of noise were lower with CR = 3 than with CR = 1, 1.5, and 2. The subjective rating of the amount of background noise increased with increasing compression ratio for CR > 1. In their second experiment *CP(1, 1.5 + 2 + 3, 5/60 + 5/200 + 5/1000)*, Neuman et al. conducted a parametric study of the effect of CR and T_r on sound quality with the same subjects. In this study, the highest compression ratios (CR = 5, and 10) of the previous study were omitted in favour of two extra time constants ($T_a/T_r = 5/60$, and 5/1000 ms). The results confirmed previous results that higher compression ratios lead to lower subjective quality ratings (clarity, pleasantness, and perceived amount of background noise). Across all noise types they found a significant effect of CR * T_r. For CR = 3, a larger release time resulted in better scores of clarity, of pleasantness, and of overall quality. For CR = 2, the influence of release time was less (except for the perceived amount of background noise). For all compression ratios, $T_r = 1000$ ms was better than $T_r = 60$, and 200 ms, but results were still worse than the linear scores obtained in the previous experiment.

CR > 3

Tyler and Kuk (1989) *CP(1, 5, 6/50-6/550)* used compression with an adaptive release time (see section 1.7.4). Release time depended on the duration of the input signal. For short input sounds the release time was short (50 ms). For long input sounds the release time was long (550 ms). With compression nearly all subjects showed a decrease in nonsense syllable recognition in low-frequency noise, relative to linear amplification.

The experiment by Van Buuren et al. (1999) *CP(1 + 4 + 16, 1 + 2 + 4, 0/0)* quoted above for single-channel compression with CR = 2 (with results similar to those for linear amplification) showed, for CR = 4, significantly worse results than linear amplification.

In the experiments of Neuman et al. (1998) *CP(1, 1 + 1.5 + 2 + 3 + 5 + 10, 5/200)* with fixed time constants ($T_a/T_r = 5/200$), quoted above, compression ratios higher than 3 resulted in lower quality scores.

In summary, when applying single-channel compression in stationary noise for subjects with moderate sensorineural hearing impairment, speech intelligibility was at best equal to that obtained with linear amplification. No improvement relative to linear amplification has been found. However, all experiments used fast-acting compression. The effect of slowly-acting single-channel compression on speech intelligibility is still not clear.

Sound quality experiments showed that a compression ratio higher than 2 degraded quality. Long release times (T_r up to at least 1000 ms) decreased the negative effect of compression on sound quality, but they did not restore it to the level achieved with linear amplification.

Multi-channel compression

CR ≤ 3

Barfod (1978) *CP(2 + 3 + 4, 1 + 1-3, 6-24/6-24)* used fast-acting two-, three- and four-channel compression at several positive signal-to-noise ratios (0, +5, +10, +15, and +20 dB). Compression ratios were subject and frequency dependent. The lowest channel always had linear amplification. Four-channel compression showed no significant difference in speech intelligibility relative to linear amplification, but three- and two-channel compression degraded performance. The four-channel scores were significantly better than those for two or three channels.

The experiment of Moore et al. (1999) *CP(1 + 2 + 4 + 8, 1 + 1-2.9, 7/7)*, quoted above for single-channel compression, also included multi-channel compression (NC = 1 + 2 + 4 + 8). Again, all compression conditions resulted in speech reception thresholds worse than achieved with linear amplification, although none differed significantly.

Van Buuren et al. (1999) CP*(1 + 4 + 16, 1 + 2 + 4, 0/0)*, quoted above, found the same results with multi-channel compression (NC = 4 + 16) with CR = 2, as with single-channel compression. Speech reception thresholds were worse than those achieved with linear amplification, but the differences were not significant.

CR $>$ 3

Van Buuren et al. (1999) CP*(1 + 4 + 16, 1 + 2 + 4, 0/0)* also used CR = 4. As with single-channel compression, quoted above, speech intelligibility with multi-channel compression (NC = 4, 16) was significantly lower than that achieved with linear amplification. It is noteworthy that for CR = 4 the scores decreased with increasing number of channels. This was not found for CR = 2.

Yund et al. (1987) CP*(8, 1 + 1-7, 4/4)* used fast eight-channel compression as described by Robinson and Huntington (1973). In this type of system, the average power of the input signal is measured continuously and is used to determine the gain for the signal at the center of each interval. For the system of Yund et al., the resulting attack and release times were both approximately 4 ms in terms of the IEC definition (see section 1.6.3). Compression ratio, varying between one and seven, was subject and channel dependent. Yund et al. measured nonsense syllable recognition with several signal-to-noise ratios, ranging from -5 to +15 dB. Compression resulted in higher nonsense syllable recognition than linear amplification for all signal-to-noise ratios.

Yund and Buckles (1995) CP*(4 + 8 + 12 + 16, 1-7, 4/4)* extended the previous experiment of Yund et al. (1987) to include four-, twelve- and sixteen-channel compression. Again several signal-to-noise ratios (from -5 to +15 dB) were used. The average nonsense syllable score was lower for four channels than for eight channels. The scores for eight, twelve, and sixteen channels were roughly equal. This is a remarkable finding, since most experiments showed deterioration in speech intelligibility with increasing number of channels. No interaction between signal-to-noise ratio and the number of channels was found.

Yund and Buckles (1995a) CP*(8, 1 + 1-7, 4/4)* compared the eight-channel results from Yund and Buckles (1995) to results obtained with linear amplification. Using all signal-to-noise ratios, a quarter of the subjects showed worse speech intelligibility with compression, whereas half of the subjects had better speech intelligibility compared to linear amplification.

Van Toor and Verschuure (2002) CP*(4, ?, 2/16-64 + 2-16/64-512 + 32-64/1024-2048)* compared three wearable four-channel compression systems. They used a fast, an intermediate, and a slow system. Time constants changed over channels, the high frequency channels being the fastest. All systems used a short signal delay to reduce overshoot. Compression was curvilinear, i.e., the compression ratio depended on the input level. The systems were fitted according to the DSL [i/o] algorithm (Cornelisse et al., 1995) with both a flat response, and a fitting with high-frequency emphasis. They used three types of noise, stationary speech-shaped noise, fluctuating noise (speech-shaped as well), and low-frequency car noise. No significant effect of time constants on speech reception threshold was found in both the stationary speech-shaped noise and in the low-frequency car noise. Only for fluctuating noise did time constants have a significant effect, see Chapter 3. With the flat frequency-response the slow system gave significantly less improvement than the intermediate or fast system. Across all three noise types (stationary, fluctuating, and low-frequency noise) and all time constants, compression with high-frequency emphasis gave significantly better results than compression with a flat frequency-response. A subjective performance test (APHAB) yielded no consistent preferences for any set of time constants.

In summary, for multi-channel compression some experiments showed better speech intelligibility with compression than with linear amplification. However, generally the results with multi-channel compression were worse than those obtained with linear amplification. The effect of an increasing number of channels was not clear.

Interaction between compression parameters

Only few studies have investigated interaction effects of number of channels, compression ratio, and time constants.

Neuman et al. (1998) $CP(1,\ 1.5 + 2 + 3,\ 5/60 + 5/200 + 5/1000)$ quoted before, have reported a significant interaction of compression ratio and release time ($CR * T_r$) based on subjective quality assessments. The positive effect of longer release times (increasing from 60, to 200, to 1000 ms) was larger with increasing compression ratio (going from 1.5, to 2, to 3).

Van Buuren et al. (1999) $CP(1 + 4 + 16,\ 1 + 2 + 4,\ 0/0)$ did not report significant interactions. However, for $CR = 4$ their data consistently showed a decrease in speech intelligibility and quality (only tested in quiet) with increasing number of channels ($NC * CR$).

Plomp (1994) $CP(1 + 2 + 4 + 8 + 16,\ 1 + 2 + 4 + \infty,\ 0/20)$ reported data of Van Dijkhuizen (J.N. van Dijkhuizen, unpublished data, 1993) in which a larger number of channels had a detrimental effect on speech intelligibility in noise for twelve moderately hearing-impaired subjects (Plomp did not specify the subjects' hearing loss or the type of background noise). This detrimental effect for larger number of channels, became more profound for higher compression ratios ($CR = 2$, 4, or ∞).

Festen and Van Dijkhuizen (1999) $CP(1 + 2 + 4 + 8 + 16,\ 1 + 2 + 4 + \infty,\ 0/20)$ investigated the effect of amplitude compression on speech intelligibility in quiet. They used a parametric design in which they changed compression ratio and the number of channels. Their sixteen subjects had an average hearing loss of 55 dB and a dynamic range of about 45 dB. The results showed a gradual decline of sentence intelligibility with increasing compression ratio and with an increasing number of channels. The decline in intelligibility for higher compression ratios was larger for a larger number of compression channels.

2.1.2 Conclusions from previous research

Our evaluation of previous studies on compression only included research with moderately hearing-impaired subjects and stationary noise. For these criteria, compression with a low compression ratio (≤ 3) did not significantly degrade speech intelligibility relative to linear amplification. For higher compression ratios, some studies showed a slight (mostly insignificant) improvement in speech intelligibility, whereas some studies showed a reduction. For a small number of channels (≤ 3) the effect of compression was limited, whereas experiments with a larger number of channels gave seemingly conflicting results. Scarcely any studies on the influence of time constants were available. The effect of time constants on speech intelligibility for moderately hearing-impaired listeners and stationary noise is therefore largely unknown. Systematic investigations into the effect of different parameter values on speech intelligibility in noise are scarce. Van Buuren et al. (1999) and Plomp (1994) investigated the effect of compression ratio and number of channels, but they did not change the time constants in their experiments.

2.1.3 The present research

In view of the inconclusive results from the literature quoted above we conducted a full parametric investigation with respect to the combination of number of channels, compression ratio and time constants. This allows for the analysis of possible interaction effects. The choice of appropriate values of the compression parameters is based on previous research, and is given in Table 2.2. First, since multi-channel compression with many channels gave seemingly contradictory results we included $NC = 1$, 2, and 6. Second, the literature showed that the compression ratio should be small. The best results were obtained with a CR around 2. Since many hearing losses are frequency dependent, we included channel dependent compression ratios. This yielded five compression ratio settings ranging from $CR = 1$ to 3, see Table 2.2. Third, it is generally accepted to use a short attack time to prevent hearing damage and discomfort, and the effect of this choice on speech intelligibility should be verified. Thus, we used two different attack times, 4 and 40 ms. Most research quoted has been performed with short release times. The effect of long release times on speech intelligibility in stationary noise is still not clear. Therefore, we used three different release times, 4, 40, and 400 ms. The experimental design will be explained in more detail in section 2.2.4.

Table 2.2: *Processing characteristics of the full parametric investigation. In brackets the total number of conditions with that specific value is given. For example, for two-channel compression 20 different conditions were measured (4 different compression ratios, each with 5 time constants). The compression ratios (CR = 1, 2, or 3) were divided in two separate ratios for low and high frequencies. The cross-over frequency was 1 kHz.*

NC		CR_{low}/CR_{high}		T_a/T_r (in ms)	
linear	(1)	linear	(1)	linear	(1)
1	(10)	1/2	(10)	4/4	(10)
2	(20)	2/2	(15)	4/40	(10)
6	(20)	2/3	(10)	40/40	(10)
		3/3	(15)	4/400	(10)
				40/400	(10)

2.2 Methods

2.2.1 Subjects

Twenty adult subjects with moderate sensorineural hearing impairment participated in this study. All included subjects completed the entire experiment; none of them quit early. The inclusion criterion for our subjects was a sensorineural hearing loss of approximately 40–60 dB HL at 4 kHz. All subjects were native Dutch speakers. Relevant subject data are given in Table 2.3. The average pure-tone hearing loss at 0.5, 1, and 2 kHz ($PTA_{0.5/1/2\,kHz}$) was 45 dB (ranging from 35 to 65 dB). The dynamic range averaged over the same frequencies ($DR_{0.5/1/2\,kHz}$) ranged from 43 to 83 dB. All but one subject had used compression hearing aids prior to the experiment. The average age of the participants was 65 ($\pm$10) year. The first session started with audiometric measurements. Pure-tone thresholds and uncomfortable loudness levels were measured for each subject in an unaided condition. Subjects had a sloping, a flat, or a helmet-shaped hearing loss. The ear used was the one that satisfied the inclusion criteria. If both ears were suitable, the ear with the smallest hearing loss was used. For most subjects the hearing loss was only slightly asymmetric between the two ears. The $PTA_{0.5/1/2\,kHz}$ of the included ear was maximally 12 dB worse than the non-included ear. Subjects participated on a voluntarily basis. After completion of all sessions, they received a reimbursement of travelling expenses and a small financial compensation of €45.

This research was approved by the Medical Ethical Committee of University Hospital Utrecht, protocol number METC 01/165, dated 21 December 2001.

2.2.2 Signal Processing

Signal processing was carried out off-line using a personal computer with MATLAB Release 12 (The Mathworks, 2001). The noise masker and the speech signal were

Table 2.3: *Relevant data of the 20 subjects. For each subject, the hearing thresholds are given on the first row and their uncomfortably levels on the second. When the maximum output level of the equipment was not sufficient the values are preceded by a > symbol. The ear measured was the one that satisfied the inclusion criteria. If both ears were suitable, the ear with the smallest hearing loss was used.*

Subject	Sex	Age (yr)	Etiology	Threshold or UCL (dB HL) 125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
s1	F	72	Unknown	65	65	55	55	65	60	85
				>90	>110	110	105	105	95	90
s2	F	80	Age	30	25	30	35	45	65	80
				>90	>110	120	120	>120	>120	>110
s3	F	41	Unknown	25	20	25	35	55	65	85
				>90	110	115	120	120	>120	>110
s4	M	64	Shotgun noise	20	25	25	45	50	60	80
				>90	>110	105	100	105	105	>110
s5	F	53	Hereditary	25	20	20	45	55	75	80
				>90	95	85	95	95	95	85
s6	F	73	Unknown/Age	20	15	25	35	45	55	75
				>90	110	115	120	>120	>120	>110
s7	M	69	Noise	25	25	30	35	45	55	90
				>90	>110	120	110	120	>120	>110
s8	F	66	Unknown	25	30	35	55	50	55	55
				>90	>110	120	120	115	115	>110
s9	M	63	Hereditary	20	20	25	50	45	60	65
				>90	>110	110	110	115	>120	>110
s10	F	62	Hereditary	50	55	50	40	50	50	50
				90	100	105	105	110	>120	>110
s11	M	71	Noise	20	20	30	50	70	70	85
				90	95	100	110	115	>120	>110
s12	F	69	Unknown/Age	35	35	35	35	45	30	65
				>90	110	110	105	105	>120	>110
s13	M	83	Age	45	40	40	55	55	60	85
				>90	>110	110	110	110	120	>110
s14	F	57	Unknown	65	65	70	65	60	65	95
				80	90	100	115	110	105	105
s15	F	67	Hereditary	45	50	50	55	60	70	75
				>90	100	100	100	95	90	90
s16	F	61	Unknown	20	15	25	45	60	45	45
				85	90	95	110	>120	115	110
s17	F	62	Hereditary	53	40	40	40	50	65	65
				>90	>110	110	110	115	120	>110
s18	F	49	Unknown	40	40	50	50	60	50	65
				>90	105	100	95	100	85	90
s19	M	66	Noise	45	30	25	30	50	75	70
				>90	105	110	115	115	>120	>110
s20	M	76	Noise	30	35	45	50	65	65	80
				>90	110	110	110	110	120	>110

summed before processing. For all conditions, including the linear condition, the speech-plus-noise signal passed through six elliptical band filters (see Table 2.4). The band filters were applied twice; after the first filtering the filters were applied again to the time reversed signals to remove any phase distortion introduced by the first filtering stage (Smith, 1997). Depending on the experimental condition, the signal was compressed independently in one, two, or six channels. With six-channel compression, each channel corresponded to one frequency band. With two channels, prior to compression, the low-frequency channel received a summation of the outputs of bands 1, 2, and 3, and the high-frequency channel received a summation of bands 4, 5 and 6. For single-channel compression the outputs of all bands were summed into one channel before compression.

Table 2.4: *Characteristics of band filtering.*

Band	Low cut-off frequency (kHz)	High cut-off frequency (kHz)	Filter order
1	-	0.25	5
2	0.25	0.5	4
3	0.5	1	3
4	1	2	3
5	2	4	3
6	4	22	7

After filtering and summation of the outputs of the bands, the amplitude envelope in each channel was calculated for an entire stimulus (full sentence) by means of a Hilbert transform (Oppenheim and Schafer, 1989). These envelopes were then logarithmically compressed (Braida et al., 1979). A gain signal was constructed for the signal in each channel by calculating the ratio of the instantaneously compressed envelope and the original envelope. In order to introduce attack and release times, each gain signal passed through two first-order low-pass filters. One filter acted only on rising signals (attack time filter), and resulted in a logarithmic decay. The other filter acted only on falling signals (release time filter), and resulted in a gain that rose logarithmically. It should be noted that our definition of time constant differs from the current and previous IEC and ANSI standards for non-linear hearing aids (see Section 1.6.3). After subjecting the gain signals to the attack and release filters, the signals in the channels were multiplied sample by sample by their respective gain signal. In each channel the compressed signals were then amplified to the long-term rms values of the original input signals in that channel. Finally, the outputs of all channels were summed.

The digital stimuli were converted to the analogue domain using a 'Creative Sound-Blaster Live Platinum 5.1' 16-bit sound card with a separate front-panel connection box. The analogue signal was adjusted to the individual audiograms using a Boss GE-131 one-third octave band graphic equalizer. The maximal individual adjustment of each of the one-third octave bands was 30 dB. The frequency response of the equalizer was measured after each change in setting and was adjusted until the desired response was obtained. During the presentation of all stimuli the frequency response remained fixed for each subject. This was possible because the compressor output was amplified to the original (uncompressed) rms for each channel. Frequency shaping was applied according to the half gain rule (Dillon, 2001). After frequency shaping the signal could be attenuated with Shallco attenuators type 2511 and 2513 by 1 and 10 dB/step, respectively. The signal was then amplified by a Sony TA-F470 amplifier, and presented monaurally through Sennheiser HDA200 circumaural headphones. The overall gain was such that the stimuli were presented around the most comfortable loudness level of each subject. Thus only the frequency response of the half gain rule (and not the prescribed overall level) was used for the final fitting. On request we slightly decreased the high frequency gain for one subject (s15).

2.2.3 Speech Material and Procedure

Speech materials consisted of Dutch sentence material for measurement of the Speech Reception Threshold (SRT) in noise, developed by Versfeld et al. (2000). This material consisted of two sets (a male and a female speaker) of 39 lists of 13 sentences each. The total number of sentences was 1014. The long-term average speech spectrum (LTASS) of each speaker was used as background noise. Speech was embedded in this background noise. The noise extended from 2 seconds before speech (onset), to 1 second after speech (offset). The onset was used to allow any initial gain adjustments of the compressor to even out before speech was presented and to attract subjects' attention to the stimulus. The offset prevented an abrupt ending after the speech had finished, thus giving the compressor time to fully change the gain in accordance with its time constants. The dynamic range of the original speech material was 26 ± 6 dB (the standard deviation was calculated from the dynamic range of all the separate sentences); the dynamic range of the background noise was approximately 4 dB (both wide band measurements between the 1^{st} and 99^{th} percentile, applying an integration time constant of 125 ms). Within each list, sentence order was randomized. This was done to minimize possible recognition effects for a repeated list. For each subject a particular list was used only once or twice. The set of sentences per list was fixed because the lists were balanced with respect to intelligibility.

Speech intelligibility was measured by the Speech Reception Threshold (SRT) test in noise. The SRT in noise is defined as the signal-to-noise ratio at which a subject can reproduce 50% of the sentences entirely correctly. The SRT was measured with the adaptive procedure developed by Plomp and Mimpen (1979). The noise was kept at a fixed average level; the signal-to-noise ratio was varied by changing the level of the speech signal. For each experimental condition, one list (consisting of 13 sentences) was presented to the subject. The first sentence was presented at a low signal-to-noise ratio. This sentence was repeated until the subject had reproduced it correctly. Each repetition was presented at a 4 dB higher signal-to-noise ratio. The subsequent sentences (2–13) were not repeated. After a correct response, the signal-to-noise ratio of the next sentence was decreased by 2 dB; after an incorrect response it was increased by 2 dB. After the last sentence, the SRT was estimated by averaging the last ten signal-to-noise ratios. The first three values were discarded from this calculation.

The experiments were conducted in a soundproof room. The subjects removed (both) their hearing aids prior to the experiment. Since the headphones provided high passive attenuation of external noise (approximately 29 dB at 1 kHz), masking at the contra lateral ear was not used. Subjects received written and oral instructions in which they were told to repeat the sentences, or part of the sentences if they did not understand them completely. It was made clear that subjects should try to give every sentence approximately the same amount of attention during the whole session, i.e., to listen to all sentences in a comfortable, not too tiring manner. Since the experiment consisted of many experimental conditions, the task was divided over several sessions on different days. Every session started with a practice session to familiarize subjects with the experimental procedure. A session lasted from 2.5 to 4 hours, including breaks. The total number of sessions (including those with fluctuating noise) varied from 3 to 6 per subject, depending on availability of the subjects and their fitness.

2.2.4 Design

The compression threshold level remained fixed for the entire experiment at −30 dB *re* speech rms. Only signals above the compression threshold were compressed. Since the dynamic range of the speech material was approximately 26 dB (see section 2.2.3) this choice of threshold level brought the entire range of relevant speech levels within the compression range, ensuring full dynamic range compression.

In total, 50 compressive processing conditions were used. The full parametric study consisted of three numbers of channels, four compression ratios, and five time constants (see Table 2.2). The numbers of channels were 1, 2, and 6 channels. The compression ratios (CR = 1, 2, or 3) were divided in two separate ratios for low and high frequencies. The cross-over frequency was 1 kHz. Note that although the compression ratio in some channels might be equal, all channels were independently compressed. The time constants consisted of combinations of attack (4 to 40 ms) and release times (4 to 400 ms). In order to limit the number of conditions, and following common hearing aid design, release times were chosen to be greater than or equal to attack times. With single-channel compression it is, of course, not possible to use different compression ratios for low and high frequencies. Therefore single-channel compression included only two compression ratios (CR = 2, and 3), resulting in 10 single-channel compression conditions in total.

In order to estimate test-retest variance, measurements were made in duplicate. Due to time constraints the duplicate measurements were made for only 13 predetermined conditions. These duplicate conditions were evenly distributed over the experiment (2^{nd}, 5^{th}, 7^{th}, 10^{th} measurement, etc.).

At the time of measurement, all conditions were presented with two types of background noise: speech-shaped stationary noise (the results of which are presented here) and fluctuating noise (Chapter 3). Ten subjects selected at random performed the experiment with stationary noise first and with fluctuating noise second; the other ten subjects vice versa. The speech materials included speech of both a male and a female speaker. The frequency spectrum of the speech-shaped noise was changed accordingly. For each type of noise (stationary or fluctuating) each subject was presented with materials from one speaker only.

2.3 Results

Using the duplicate measurements (see section 2.2.4) of all subjects, a t-test for dependent samples (paired by condition) showed a significant improvement of the retest score relative to the test score (improvement was 0.5 dB; $t = 3.7$; $p < 0.001$). No significant interaction between this learning effect and subject was found ($F = 0.7$; $p = 0.8$) and the learning effect did not depend on compression condition ($F = 0.5$; $p = 0.9$). As stated above, the duplicate measurements were distributed over the entire experiment. Some test-retest pairs were measured during the same session, whereas others spanned 2 or 3 sessions. Test-retest measurements which took place in the same session did not differ significantly from test-retest pairs in two different sessions ($F = 0.2$; $p = 0.7$). Regression analysis over all subjects showed that the learning effect depended on only the number of previous measurements ($r^2 = 0.02$; $p < 0.05$), not on session. Figure 2.1 gives the difference between the retest and the test score for all subjects as a function of presentation number, irrespective of session. Although the learning effect was small, it was significant and therefore a correction was applied. This correction depended on the number of previous encountered sentences and was +0.019 dB per presentation number. The maximum correction at the end of the experiment was +1.3 dB.

After correction for the learning effect, the average variance of the duplicate pairs was 2.3 dB^2. It was independent of compression condition ($F = 0.6$; $p = 0.9$). Before further analysis, all results for the duplicate pairs were averaged and the average values were included in the analyses as one data point. Subsequent analyses of variance did therefore not include the duplicate variance in error estimation.

As described in Design (section 2.2.4), for stationary noise half of the subjects used only material uttered by a male speaker and half used material uttered by a female. An analysis of variance with noise type (stationary, fluctuating) and speaker (male, female) as independent factors showed no significant difference in speech reception thresholds between the male and the female lists ($F = 0.9$; $p = 0.3$; $SRT_{female} - SRT_{male} = -1.0$ dB). The interaction between noise type and speaker was also insignificant ($F = 0.2$; $p = 0.7$).

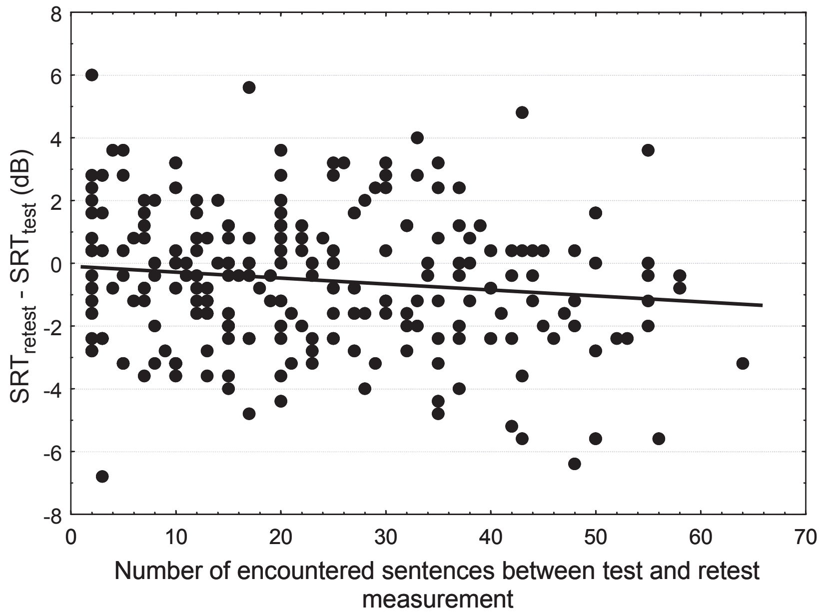


Figure 2.1: *SRT of the retest measurement (SRT_{retest}) minus SRT for the test measurement (SRT_{test}) for all subjects. The test and the corresponding retest conditions were distributed at random over the entire experiment. The ordinate gives the difference in speech reception threshold of the retest and the test measurements, irrespective of session. A negative value represents a better retest. The abscissa represents the difference in number of previously encountered sentences between the two measurements.*

Since the effect of speaker was not significant, this factor is omitted in further analyses. Possible interactions between compression parameters and male/female material were assumed to be negligible and were therefore not tested.

Across all subjects, the average speech reception threshold for linear amplification in stationary noise was 0 dB, which is 4 dB worse than the average scores of normal-hearing subjects using the same speech material (Versfeld et al., 2000; Lyzenga et al., 2002). Figure 2.2 gives the SRT as a function of the subjects' pure tone average hearing loss at 0.5, 1, and 2 kHz. The figure suggests that SRT deteriorates with increasing hearing loss. However, linear regression analysis restricted to our present data yielded no significant correlation and explained only a small amount of variance ($r^2 = 0.11$; $p = 0.15$). Comparing the speech reception thresholds with PTAs for other frequencies yielded correlation coefficients explaining even less variance ($r^2 \leq 0.003$ for $PTA_{1,2,4\,kHz}$, $PTA_{2,4\,kHz}$, and $PTA_{2,4,8\,kHz}$). Apparently, speech perception of our subjects in stationary noise after linear amplification could not be predicted accurately by pure tone hearing loss measured in silence. Since this speech material was developed recently (Versfeld et al., 2000), only few results pertaining to this material have been reported. Included in Fig. 2.2 are data for both normal hearing subjects (Versfeld et al. 2000; Lyzenga et al., 2002) and for sensorineurally hearing impaired subjects (Lyzenga et al., 2002). Our data relate well to the added data, see Fig. 2.2.

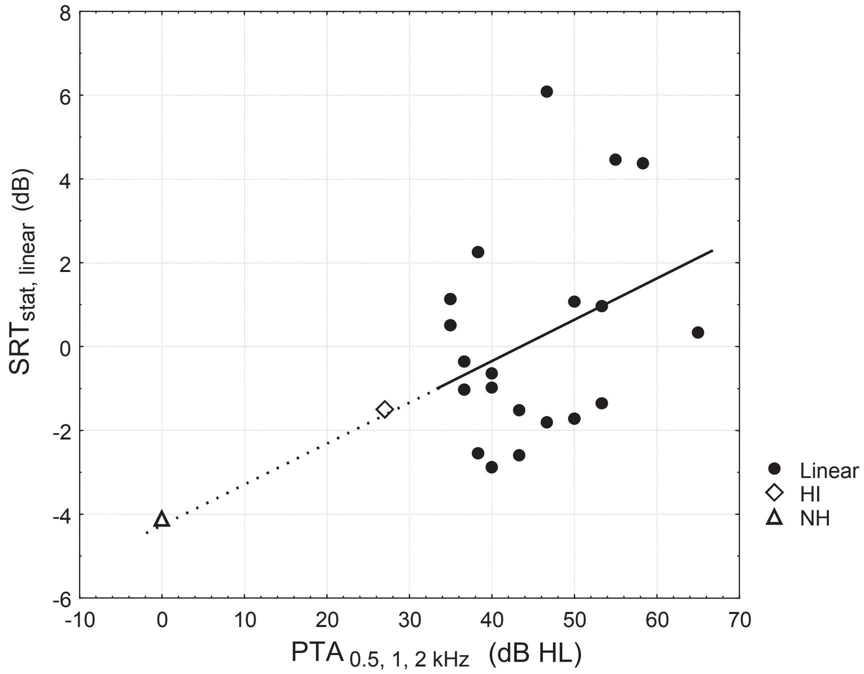


Figure 2.2: *SRTs obtained with linear amplification ($SRT_{stat,\,linear}$) as a function of the subjects pure-tone average hearing loss at 0.5, 1, and 2 kHz ($PTA_{0.5,\,1,\,2\,kHz}$). Regression analysis of our data (filled circles, solid line) yielded no significant correlation ($r^2 = 0.11$, $p = 0.15$). The diamond represents additional data for hearing-impaired (HI) subjects from Lyzenga et al. (2002). The Triangle represents data for normal hearing subjects (NH) from both, Versfeld et al. (2000) and Lyzenga et al. (2002). These additional data coincide with the insignificant regression line of our data.*

2.3.1 Concurrent analysis of single-, two-, and six-channel compression

The measured SRT for linear amplification ranged between -2.9 dB and +6.1 dB (average SRT $= 0.2 \pm 2.5$ dB). The SRT for compression conditions will be presented relative to these linear scores. For this purpose, ΔSRT is defined as the SRT for a particular condition minus the SRT for the linear condition for that subject. Negative values of ΔSRT indicate a better speech reception threshold for compression than for linear amplification.

A univariate[1] repeated measures analysis of variance was performed on ΔSRT. Number of channels (3 levels), compression ratio (2 levels) and time constants (5 levels) were all treated as independent within subject factors. Single-channel compression does not allow for frequency dependent compression ratios (section 2.2.4), therefore CR $= 1/2$ and CR $= 2/3$ were not included in this analysis. The results are given in Table 2.5. Figure 2.3 shows ΔSRT as a function of time constants. The panels represent number of channels and the two curves represent compression ratios.

[1]Mauchley's test of sphericity did not yield any significant results. Univariate analysis was thus appropriate.

Table 2.5: *Results of a repeated measures analysis of variance. Single-channel compression was included and therefore split-frequency compression ratios were omitted. (NC = 1, 2, 6; CR = 2/2, 3/3; T_a/T_r = 4/4, 4/40, 40/40, 4/400, 40/400).*

Effect	Degrees of freedom	MS	F	p-value
NC	2	11.1	7.0	< 0.005
CR	1	73.5	24.6	< 0.0001
T	4	21.2	10.0	< 0.00001
NC * CR	2	4.8	2.6	0.08
NC * T	8	7.9	3.9	< 0.0005
CR * T	4	2.9	1.3	0.3
NC * CR * T	8	4.7	2.4	< 0.05

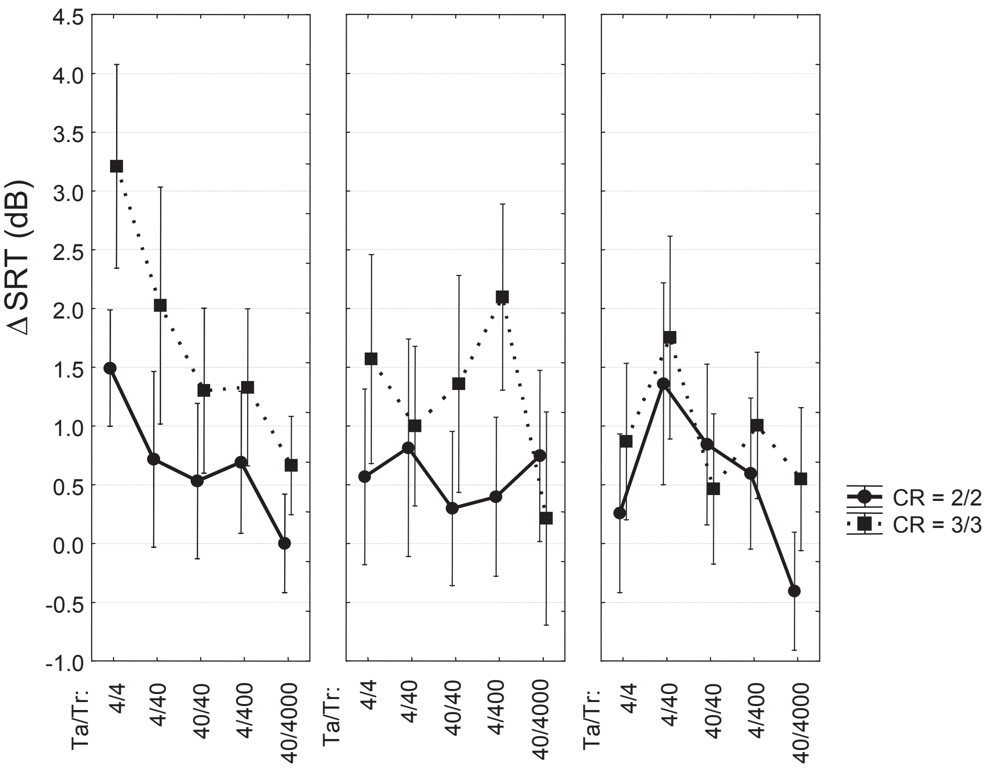


Figure 2.3: *ΔSRT averaged over all subjects. ΔSRT is the SRT for compression minus the subjects' SRT for linear amplification. Lower values of ΔSRT represent better speech intelligibility with compression. ΔSRT is plotted as a function of time constants (T_a/T_r), and the ordering of T_a/T_r is primarily based on release time. Each curve represents a compression ratio setting. The panels represent different number of channels. Split-frequency compression ratios are not included. Vertical bars denote 0.95 confidence intervals.*

All main effects (NC, CR, and T_a/T_r) were highly significant ($F \geq 7.0$; $p < 0.005$). First, the effect of number of channels showed that the average results improved when going from single-, to two-channel, to six-channel compression. However, post-hoc analysis (Tukey's Honestly Significantly Difference test, HSD) gave only one significant contrast: six-channel compression resulted in better speech intelligibility (lower ΔSRT) than single-channel compression ($\Delta SRT_{NC=6} - \Delta SRT_{NC=1} = -0.5$ dB, $p < 0.05$).
Second, a compression ratio of 2/2 yielded consistently lower ΔSRTs than CR = 3/3. Post-hoc comparison confirmed this finding ($\Delta SRT_{CR=2/2} - \Delta SRT_{CR=3/3} = -0.7$ dB, $p < 0.001$). Both compression ratios resulted in significantly higher SRTs than those obtained with linear amplification ($\Delta SRT_{CR=2/2} = +0.6$ dB, $p < 0.001$; $\Delta SRT_{CR=3/3} = +1.3$ dB, $p < 0.00001$).
Third, larger time constants gave better results than smaller constants. This finding was confirmed by post-hoc comparisons which showed that $T_a/T_r = 40/40$ gave better results than $T_a/T_r = 4/4$ ($p < 0.05$), and that $T_a/T_r = 40/400$ was better than 4/4 ($p < 0.001$), 4/40 ($p < 0.001$), and 4/400 ($p < 0.05$). The improvement for $T_a/T_r = 4/400$ over 4/40 and 4/4 was insignificant.

Besides the main effects, the effect of the interaction NC * T was also significant ($p < 0.0005$, see Table 2.5). This interaction is apparent in Figure 2.4, which shows ΔSRT as a function of time constants for each number of channels. For NC = 1, the smallest time constants ($T_a/T_r = 4/4$) resulted in the highest ΔSRTs. Whereas for NC = 2, or NC = 6, the results for $T_a/T_r = 4/4$ were not worse than for larger time constants.

Finally, the effect of the highest interaction term (NC * CR * T) was also significant. Results with single-channel compression were best for the largest time constants. Although this effect was present for both compression ratios, the improvement was larger for CR = 3/3 than for CR = 2/2 (Fig. 2.3). With two-channel compression the effect of time constants was different for CR = 2/2 and CR = 3/3. With CR = 2/2 speech intelligibility did not change for larger time constants, it was rather constant. In contrast, for CR = 3/3, speech intelligibility was lowest with $T_a/T_r = 4/400$, and best with $T_a/T_r = 40/400$ ($\Delta SRT_{T_a/T_r=40/400} - \Delta SRT_{T_a/T_r=4/400} = -1.9$ dB, $p < 0.01$). For six-channel compression the effect of time constants was roughly the same for both compression ratios.

For the highest interaction effect (NC * CR * T), compression ratio seems to have the least influence. Moreover, both second order effects with compression ratio (NC * CR and CR * T) were not significant. The significance of both NC * T and NC * CR * T indicates that interactions are important and that full parametric investigations might be important for understanding the effects of compression.

2.3.2 Analysis with the inclusion of split-frequency compression ratios (1/2 and 2/3)

In the previous analysis, no frequency dependent compression ratios were included be-

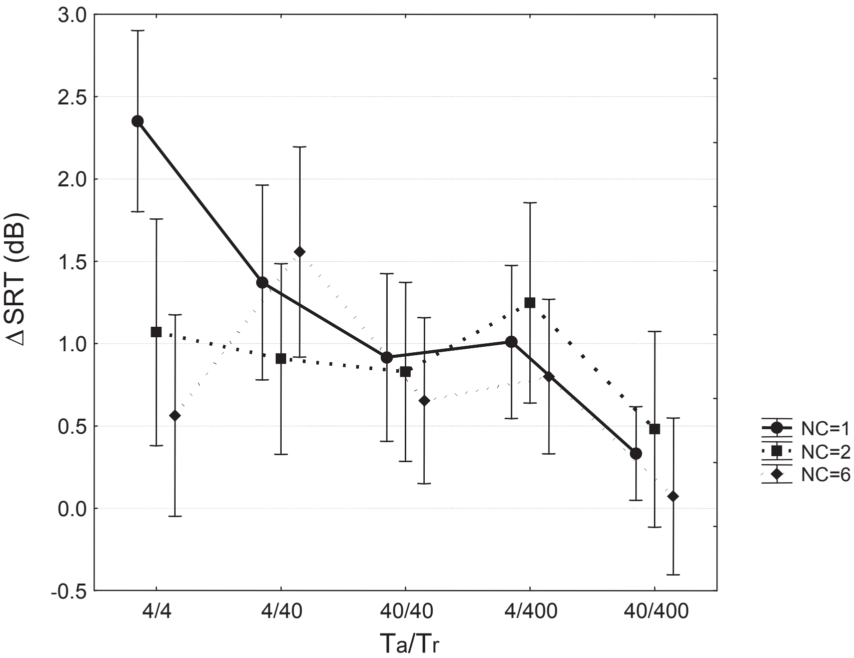


Figure 2.4: *ΔSRT averaged over all subjects. ΔSRT is given as a function of time constants (T_a/T_r) and number of channels (NC) for CR = 2/2 and 3/3 only. ΔSRT is the SRT for compression minus the subjects' SRT for linear amplification, and lower values represent better speech intelligibility with compression. The ordering of T_a/T_r is primarily based on release time. Each curve represents a different number of channels. Split-frequency compression ratios were not included. Vertical bars denote 0.95 confidence intervals.*

cause of the inclusion of single-channel compression. An additional univariate analysis[2] was conducted in which single-channel compression was omitted in favour of the inclusion of the split-frequency compression conditions CR = 1/2 and 2/3. Number of channels (NC = 2, 6), compression ratios (CR = 1/2, 2/2, 2/3, and 3/3), and time constants (5 levels) were all treated as independent within subject factors. The results of the analysis are presented in Table 2.6 and Figure 2.5.

In this second analysis, the main effect of number of channels was not significant, due to the exclusion of single-channel compression which caused the significant effect in the previous analysis.

The main effect of compression ratio still was significant. Post-hoc comparison (HSD) of the effect of compression ratio showed that, in accordance with the previous analysis, CR = 3/3 resulted in significantly worse scores than those obtained for CR = 2/2. Moreover, the scores for CR = 3/3 were worse than those obtained for all other compression ratios (CR = 1/2, 2/2, and 2/3, $p < 0.01$). The differences between the other compression ratios (CR = 1/2, 2/2, and 2/3) were small (0.1 dB) and insignificant

[2] After inclusion of the split compression ratios, Mauchley's test of sphericity again did not yield any significant results. Univariate analysis is thus appropriate.

Table 2.6: *Results of a repeated measures analysis of variance. Split-frequency compression conditions were included, and single-channel compression was omitted. NC = 2, 6; CR = 1/2, 2/2, 2/3, 3/3; T_a/T_r=4/4, 4/40, 40/40, 4/400, 40/400.*

Effect	Degrees of freedom	MS	F	p-value
NC	1	7.0	2.7	0.1
CR	3	15.8	6.8	< 0.0005
T	4	13.6	5.9	< 0.0005
NC * CR	3	2.1	0.9	0.5
NC * T	4	7.0	2.9	< 0.05
CR * T	12	2.4	1.1	0.4
NC * CR * T	12	4.4	1.8	< 0.05

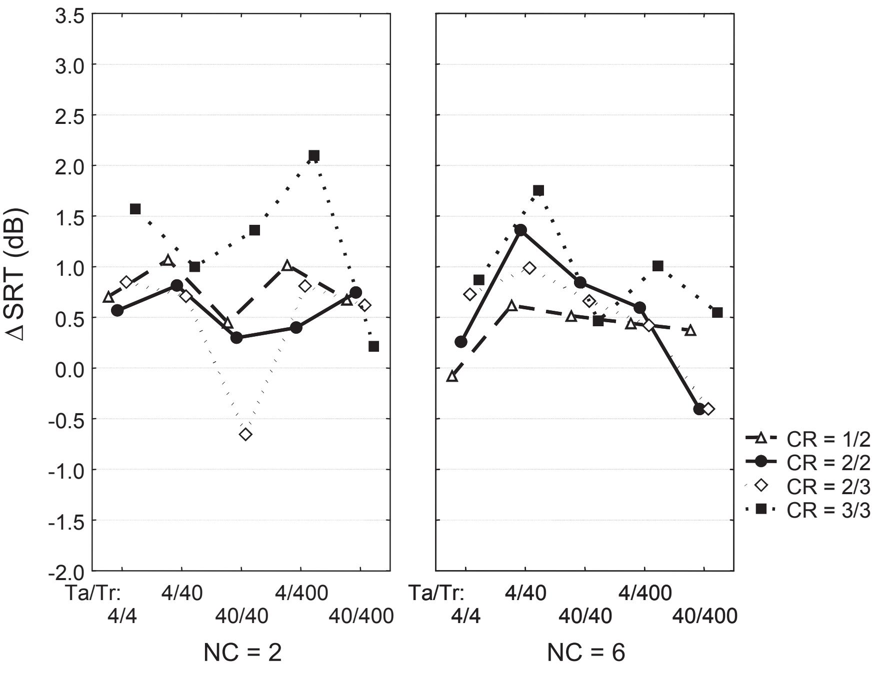


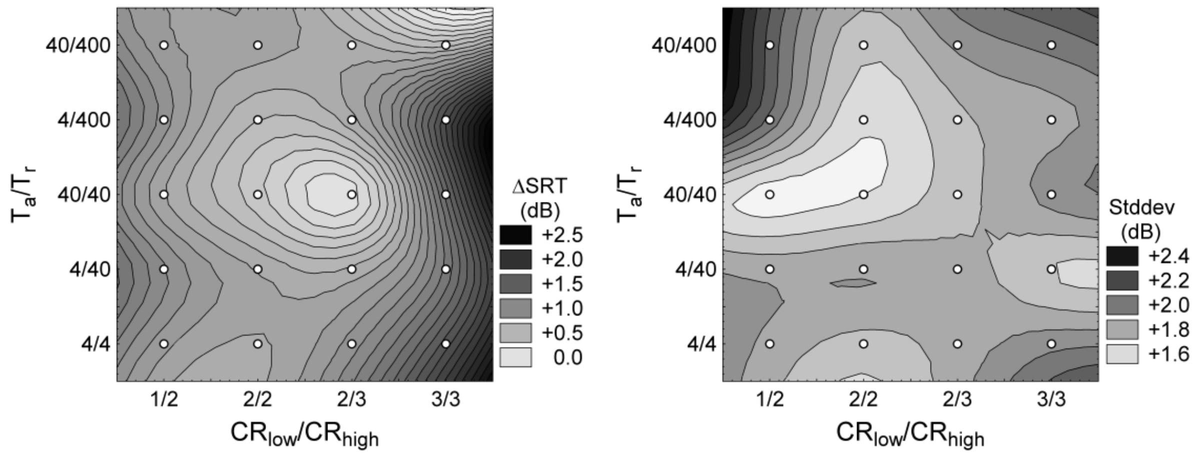
Figure 2.5: *ΔSRT is averaged over all subjects and shown as a function of time constants (T_a/T_r), compression ratio (CR) and number of channels (NC). Split-frequency compression ratios (CR = 1/2 and 2/3) were included; single-channel compression was omitted. ΔSRT is the SRT for compression minus the subjects' SRT for linear amplification, and lower values represent better speech intelligibility with compression. The ordering of T_a/T_r is primarily based on release time. Confidence intervals are omitted for clarity. The 0.95 confidence intervals were approximately 1.5 dB.*

($p \geq 0.9$). All four compression ratios resulted in significantly higher ΔSRTs than those obtained with linear amplification (ΔSRT $\geq$ +0.5 dB; $p < 0.05$).

After inclusion of the split-frequency compression conditions, the main effect of time constants still exhibited the same trend as in the previous analysis, i.e., larger time constants yielded better results. The exclusion of NC = 1 resulted in a less detrimental effect of compression on ΔSRT for all time constants, except for the largest ($T_a/T_r = 40/400$). For both analyses, $T_a/T_r = 40/400$ gave the best results of all time constants (ΔSRT = +0.3 dB, for both analyses). Post-hoc comparisons of the results for the main effect of T_a/T_r gave three significant results. First, $T_a/T_r = 40/40$ resulted in a 0.5 dB better ΔSRT than $T_a/T_r = 4/40$ ($p < 0.05$). Second, $T_a/T_r = 40/400$ gave better ΔSRTs than both $T_a/T_r = 4/40$ and 4/400 (0.7 dB; $p < 0.05$, and 0.6 dB; $p < 0.0001$, respectively). The difference between $T_a/T_r = 40/400$ and $T_a/T_r = 4/4$, which was significant in the previous analysis, was insignificant: the omission of NC = 1 diminished the detrimental effect of $T_a/T_r = 4/4$.

The inclusion of split-frequency compression conditions gave the same significant interaction effects as the previous analysis: both NC * T and NC * CR * T were significant. For NC = 2, a contour plot of the average ΔSRT score as a function of compression ratio and time constants is shown in Figure 2.6(a). The standard deviation between the subjects is shown in Figure 2.6(b). Experimental data was available for every grid point. The intermediate values, represented by contour lines and colours, were obtained by a distance-weighted least-squares calculation. While the main effect of compression ratio is clearly discernible in the figure (CR = 3/3 resulted in the highest ΔSRTs), the main effect of time constants (better scores for increasing T_a/T_r) is less pronounced. Looking only at data obtained with NC = 2, the interaction CR * T was significant ($F = 1.8$; $p < 0.05$). As stated before, overall (NC = 2+6), this interaction was not significant. For NC = 2 the best SRT was −0.7 dB and was achieved with CR = 2/3 and $T_a/T_r = 40/40$. The improvement relative to linear amplification was not significant ($p = 0.1$). However, Fig. 2.6(b) shows that the standard deviation of the individual scores was relatively low for this setting (1.7 dB). This indicates that for two-channel compression this setting is a safe choice for most subjects, since the difference in scoring between the subjects is relatively small.

Contour plots for NC = 6 are shown in Figure 2.7. The main effect of compression ratio, which was significant for NC = 2, was not significant for NC = 6 ($p = 0.06$). The main effect of time constants (decreasing ΔSRT with increasing T_a/T_r) is clearly visible in Fig. 2.7, and was significant ($p < 0.0001$). Without the data for NC = 2, the interaction CR * T was no longer significant ($F = 1.1$; $p = 0.4$). For NC = 6, the best result (ΔSRT = −0.4 dB) was obtained for the longest release time and two compression ratios: *CP(6, 2/2 + 2/3, 40/400)*. This improvement relative to linear amplification was not significant ($p = 0.1$). However, the standard deviations for these best settings were relatively low (1.1 and 1.5 dB, respectively).



(a) *ΔSRT* (b) *Standard deviation*

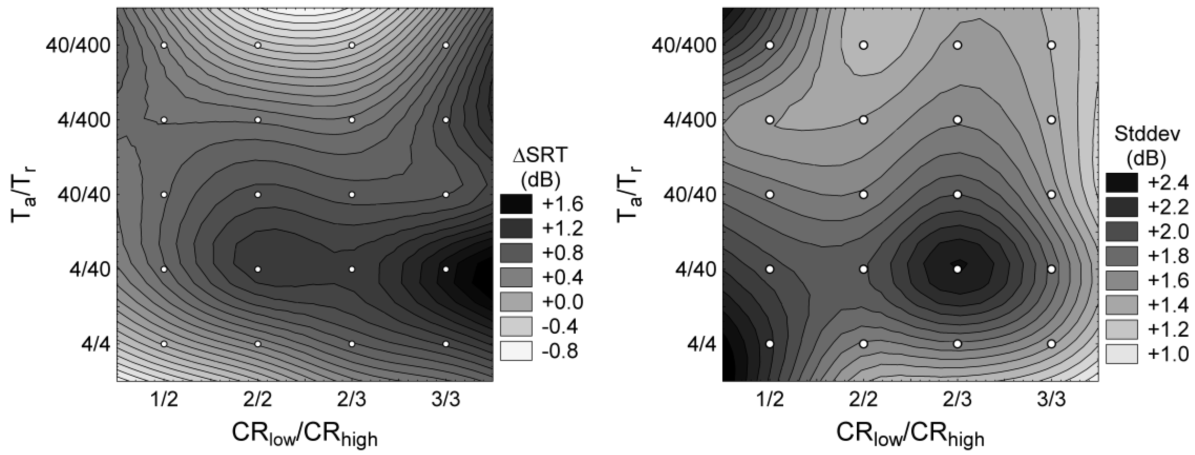
Figure 2.6: *Contour plots for two-channel compression. Experimental data was available for every grid point (○). The intermediate values were calculated by a distance weighted least squares interpolation.*

The results for NC = 2 show that the interaction between CR and T is important if one seeks the best setting with respect to speech intelligibility. This interaction was not significant when the analysis was limited to data for NC = 6, thus indicating that the interaction of CR * T changes with number of channels. This is in agreement with the significant interaction of NC * CR * T. Unfortunately, all improvements relative to linear amplification were insignificant.

2.3.3 Analysis with attack and release time

In order to analyse the results of attack and release time individually, a univariate repeated measures ANOVA was performed with T_a and T_r as independent within variables (all combinations of $T_a = 4$, 40 ms; $T_r = 40$, 400 ms). All compression ratios (CR = 1/2, 2/2, 2/3, 3/3) and number of channels (NC = 1, 2, 6) were included. The analysis showed that the effect of both T_a ($p < 0.000001$) and T_r ($p < 0.05$) was significant. Post-hoc analysis with Tukey's HSD test showed that an attack time of 40 ms yielded better results than an attack time of 4 ms ($\Delta SRT_{T_a = 40} - \Delta SRT_{T_a = 4} = -0.6$ dB; $p < 0.0001$). A release time of 400 ms gave better ΔSRTs than a release time of 40 ms ($\Delta SRT_{T_r = 400} - \Delta SRT_{T_r = 40} = -0.2$ dB; $p < 0.05$). Thus larger time constants (both attack and release time) resulted in better scores.

Since release time spanned a larger range (4 to 400 ms) than attack time (4 to 40 ms), the ordering of T_a/T_r in Fig. 2.3 and subsequent figures is primarily based on release



(a) Δ*SRT* (b) *Standard deviation*

Figure 2.7: *Contour plots for six-channel compression. Experimental data was available for every grid point (○). The intermediate values were calculated by a distance-weighted least-squares interpolation.*

time.

2.4 Discussion

2.4.1 Main effects of NC, CR, and T_a/T_r

Number of channels

The effect of number of channels was significant in the analysis which included single-channel compression. In this analysis, six-channel compression gave better results (0.5 dB) than single-channel compression. This significant difference was severely influenced by a detrimental score for single-channel compression with small time constants and $CR = 3/3$ (see Fig. 2.3). However, for larger time constants $NC = 1$ gave lower scores than $NC = 2$ and 6 as well

In contrast to our result, the available literature often states that a larger number of channels (< 3 or 4) leads to lower speech intelligibility (section 2.1.1: Hickson, 1994; Souza, 2002). However the (limited number of) available studies with moderately hearing impaired subjects and low compression ratios ($CR \leq 3$) did not show detrimental results for $NC > 3$ with $CR \leq 3$ (section 2.1.1: Barfod, 1978; Moore et al., 1999; Van Buuren et al., 1999). These results relate well to our finding that speech intelligibility did not decrease for compression with an increasing number of channels.

Compression ratio

A compression ratio of 3/3 resulted in consistently worse speech intelligibility than all lower compression ratios (CR = 1/2, 2/2, and 2/3). This detrimental effect of CR = 3/3 was confirmed by post-hoc analysis, and was very consistent and virtually independent of NC and T_a/T_r. The differences between the other three compression ratios were small (< 0.1 dB) and insignificant. Thus, no overall best compression ratio was found. The experiments of Van Buuren et al. (1999) also showed no significant difference in speech intelligibility for CR = 1 and 2, and worse results for CR = 4. The experiments of Neuman et al. (1998), which spanned compression ratios from 1 to 10, showed that the subjective quality of compressed speech declined for CR > 2.

Previous research with *severely* hearing-impaired subjects did not show such a clear detrimental effect for CR = 3 in stationary noise (Dreschler et al., 1984; King and Martin, 1984). In those experiments the results of CR = 3 might have been less detrimental because compression allowed the entire speech signal to fit within the small dynamic range of the subjects. In contrast, our subjects only had a moderate hearing loss, which allowed us to present the uncompressed speech signal entirely within their dynamic range, without the need for compression. A compression ratio of 3/3 was thus not required, and the detrimental effects of compression were dominant.

Time constants

The main effect of time constants was significant in both the analysis with and without split-frequency compression ratios. With our choice of parameter values, the effect of time constants on speech intelligibility was larger and more significant than the effect of compression ratio and number of channels. Overall, larger time constants resulted in better scores. The best results were obtained for the largest time constants (T_a/T_r = 40/400). Analysis per number of channels (NC = 1, 2, or 6) showed that the effect of time constants on speech intelligibility was largest for single-channel compression. For speech quality, Neuman et al. (1998) also found a large effect of time constants for single-channel compression.

2.4.2 Attack and release time

Attack time

Most commercial hearing aids use a short attack time to prevent sudden loud noises from overloading the aid or, more importantly, the ear. In the present experiment no stimuli with sudden large increases in sound level were used, since the experiment was designed to investigate the influence of compression on speech intelligibility and not on comfort or speech audibility. Thus, all stimuli were presented within the dynamic range of hearing of the subjects and a short attack time was not required to protect the subject from discomfort (or damage) caused by sudden high-level noises. Our results indicated that an attack time of 40 ms resulted in better speech intelligibility than 4 ms ($\Delta SRT_{T_a=40} - \Delta SRT_{T_a=4} = -0.6$ dB). This benefit of a longer attack time for speech intelligibility was consistently present for all number of channels and all compression ratios.

Nábělek performed an experiment in which he changed attack time and his results are in agreement with ours (Nábělek, 1983, *CP(1, 1 + 2.5, 1/10 + 1/30 + 3/30 + 3/90 + 10/90 + 42/370)*). He found that an attack time of 10 ms gave better speech intelligibility than an attack time of 3 ms with nine severely hearing-impaired subjects and noise added after compression. His attack time of 10 and 3 ms were based on the IEC definition, and would be roughly equivalent to $T_{exp} = 4$ and $T_{exp} = 1$ ms according to our definition (see section 1.6.3).

Envelope fluctuations of speech mostly occur in a range from about 0.1 to 40 Hz (Verschuure et al., 1996) in each frequency band. For speech intelligibility the modulations between 2 and 8 Hz are most important (Chi et al., 1999). Verschuure et al. (1996) measured the effective compression in quiet as a function of release time: *CP(1, 4, 5/15 + 5/30 + 5/60 + 5/120)*. Their results showed that for $T_{r,\,IEC} = 120$ ms (about $T_{exp} = 50$ ms) only modulations up to about 5 Hz were compressed. For a release time of 15 ms (about $T_{exp} = 6$ ms) modulations up to 40 Hz were compressed. Extrapolating these results for release time to attack time, our longest attack time (40 ms) might have left important modulations (> 5 Hz) intact, while the short attack time (4 ms) might have introduced temporal distortion.

Release time

Results indicated that on average longer release times gave better speech intelligibility. This effect of long release times is typically attributed to less temporal distortion due to less effective compression (see the previous section). Since our stimuli were audible without the need for compression, temporal distortion caused by compression with short release times might have dominated a possible beneficial effect of increased audibility.

Attack time confounded with release time

Results showed that attack time had more effect on speech intelligibility than release time: the maximal effect of attack time ($\Delta SRT_{T_a=40} - \Delta SRT_{T_a=4} = -0.6$ dB, $p < 0.00001$) was larger than the maximal effect of release time ($\Delta SRT_{T_r=400} - \Delta SRT_{T_r=40} = -0.2$ dB, $p < 0.05$). This might be related to two possible effects.

First, compression with small time constants distorts the wave form. The amount of distortion can be diminished by choosing large time constants. If one time constant (T_a or T_r) is small, the effect of compression can be reduced by increasing the other constant. Thus, in order to limit the amount of distortion, either T_a or T_r (or preferably both) should be long. However, in the present experiment attack and release times were confounded. No combinations with $T_a > T_r$ were used (Table 2.2). Changing the smallest of T_a and T_r might have a more profound effect on speech intelligibility than changing the largest.
Second, T_a might have more effect on syllable audibility than T_r. A short T_a causes a gain reduction during the occurrence of a (high intensity) vowel. If T_r is large, the gain will be restored only slowly and it will still be lower than optimal during the following consonant, making this consonant less audible. In contrast, a long T_a (or a short T_r) will not decrease audibility of consonants following a vowel.
Both arguments lead to the assumption that T_a should be long, albeit for different reasons. They are in agreement with the finding that $T_a = 40$ ms resulted in better speech intelligibility than $T_a = 4$ ms.

2.4.3 Interactions

The repeated measures analyses showed that the interaction terms NC * T and NC * CR * T both had significant effects. To our knowledge, both significant effects have not been reported before.

NC * T: slowly-acting single-channel compression resulted in better speech intelligibility than fast-acting single-channel compression

Speech intelligibility with single-channel compression improved with increasing time constants (see Fig. 2.3). The frequency spectra of our materials (both male and female speech) contained the most energy at low frequencies (roughly 100-600 Hz). With single-channel compression these dominant low frequencies mainly controlled the compressor. The gain at higher frequencies (> 600 Hz) was thus controlled by the sound energy at the dominant low frequencies. This caused intermodulation distortion of speech envelopes. Increasing the time constants decreased this distortion, which might have resulted in lower ΔSRTs. Another way of suppressing the (low-frequency controlled) high-frequency gain can be the use of high-frequency emphasis prior to single-channel compression, as is done in several commercially available hearing aids. Note that in this study the negative effect of fast-acting compression was not counteracted by the advantage of an increased audibility of speech, since the dynamic range of hearing of the subjects was large enough to accommodate the stimuli, especially at the dominant low frequencies.

NC * T: fast-acting two-channel compression gave better results than fast-acting

single-channel compression

In contrast to single-channel compression, for two-channel compression the results with small time constants were relatively good, see Fig. 2.3. With two-channel compression, the gain was independently controlled below and above 1 kHz. By using two channels, the control of dominant low frequency signals over high-frequency speech information was reduced. Two-channel compression thus resulted in less intermodulation distortion than single-channel compression, which might explain the improved results relative to fast-acting single-channel compression.
The choice of the frequency bands might have influenced this result. As stated before, our stimuli contained most energy between 100 and 600 Hz. Lowering the cross-over frequency of the two-channel system might even further reduce the intermodulation distortion. A two-channel compression system with a cross-over frequency based on speech characteristics might prove beneficial. At present no research into multi-channel compression is known to systematically cover the effect of filter cross-over frequencies.

NC * T: fast-acting six-channel compression resulted in better speech intelligibility than fast-acting single- and two-channel compression

For six-channel compression the energy distribution of the speech materials resulted in the lowest channel (0-250 Hz) being dominant. Using six-channels allowed low-level signals in the other five channels to be amplified independently from the lowest channel, decreasing intermodulation distortion. The dynamic range of our stimuli was about the same for all frequencies (approximately 26 dB, section 2.2.3). In contrast, the average dynamic range of hearing of our subjects was much larger at low frequencies (72 ± 17 dB at 500 Hz) than at high frequencies (53 ± 17 dB at 4 kHz). Our fitting procedure let the subjects determine the preferred overall sound level which ensured that subjects did not experience the discomfort of too loud stimuli during the experiment. Although the dynamic range of our subjects was large enough to accommodate the entire speech signal, low-level high frequency components of the stimuli were sometimes presented close to or even below the threshold of hearing. Fast-acting compression in the high frequency channel(s) could lift these low-level elements to audible levels, thus improving speech intelligibility relative to two-channel compression.

NC * CR * T: difference in scores between CR = 2/2 and CR = 3/3 changed with T_a/T_r

Generally, slowly-acting compression results in less effective compression than fast-acting compression (Verschuure et al., 1996). Therefore an increase in compression ratio from CR = 2/2 to 3/3 will have more influence on the speech signal for a fast-acting system than for a slowly-acting system. One would thus expect that the effect of compression ratio will be smaller for slowly-acting compression than for fast-acting compression. Indeed, for our single-channel compression, the effect of compression ratio was larger for fast-acting compression ($\Delta SRT_{CR=3/3} - \Delta SRT_{CR=2/2} = +1.7$ dB) than for slowly-acting compression ($\Delta SRT_{CR=3/3} - \Delta SRT_{CR=2/2} = +0.7$ dB), see Fig. 2.3. This effect was not present for NC=2, and 6. Similarly, Neuman et al. (1998) reported a larger influence of release time for CR = 3 than for CR = 2 on subjective quality assessments for their single-channel compression system.

2.4.4 Best settings

Unfortunately, nearly all obtained speech reception thresholds were higher than those measured with linear amplification. Overall the best result ($\Delta SRT = -0.7$ dB) was achieved with two-channel compression, *CP(2, 2/3, 40/40)*. Besides this setting, only three others resulted in better scores than those obtained with linear amplification. These three were obtained with six-channel compression: *CP(6, 2/2 + 2/3, 40/400)* which both resulted in $\Delta SRT = -0.4$ dB, and *CP(6, 1/2, 4/4)* which resulted in $\Delta SRT = -0.1$ dB. The best result with single-channel compression was equal to that with linear amplification.

ΔSRT was averaged across subjects, which implies that the settings resulting in the best ΔSRTs might not be optimal for each individual subject. However, the small standard deviations between the scores at the 3 best results (1.7, 1.1, and 1.5 dB, respectively) indicate that these settings vary relatively little over subjects. Furthermore, for 11 subjects compression with *CP(2, 2/3, 40/40)* resulted in the best 25% of scores (and for 1 subject it resulted in the worst 25%). Results for compression with *CP(6, 2/2, 40/400)* and *CP(6, 2/3, 40/400)* were among the top 25% for 8 and 12 subjects, respectively. Both were among the worst 25% for 2 subjects only.

Our results could be used as a general guideline in hearing aid fitting. If one wishes to optimize an individual fit the present results demonstrate that such an optimization is difficult to conduct, particularly within the time constraints in everyday practice. The effects of the compression parameters were small. One may consider to simply use the general result for each individual.

2.5 Conclusions

In the absence of audibility improvements, compression did not significantly improve speech intelligibility of moderately hearing-impaired subjects in stationary noise. In fact, most compression settings resulted in lower scores than those obtained with linear amplification. Main effects showed that compression ratio should be smaller than 3, time constants should be large ($T_a/T_r = 40/400$ ms), and the number of channels larger than 1. Six-channel compression resulted in approximately the same speech intelligibility as two-channel compression. Two interaction effects (NC * T and NC * CR * T) were significant, indicating that specific combinations of parameter values can lead to different results. Finding the best settings for an individual based on speech intelligibility will be difficult if not almost impossible. One may consider relying on the best result from this experiment; for two-channel compression CP*(2, 2/3, 40/40)* and for six-channel compression CP*(6, 2/2 + 2/3, 40/400)*.

The effect of compression on speech intelligibility in fluctuating noise

3

Rolph Houben and Guido F. Smoorenburg
Manuscript in revision for Audiology & Neurotology

3.1 Introduction

In real-life communication, background noise often consists of one or more interfering speakers. For normal-hearing listeners, speech intelligibility is higher in such a fluctuating noise than in stationary noise, at least when the noise is presented at the same rms signal level and with the same spectral distribution. This difference in speech intelligibility between fluctuating and stationary noise is called the release of masking, and it is generally much smaller in listeners with a sensorineural hearing loss than in normal-hearing subjects (Festen and Plomp, 1990; Bronkhorst and Plomp, 1992; Gustafsson and Arlinger, 1994; Eisenberg et al., 1995; Stuart and Phillips, 1996; Hagerman, 1997). Amplitude compression might improve speech intelligibility of the hearing-impaired listeners, by lifting the speech signal to higher levels during the gaps in fluctuating noise (Souza, 2002). However, the effect of different compression parameters on speech intelligibility can be quite complicated and compression parameters may interact. Previous studies focused on varying one or two parameters per experiment. This study focuses on the combination of four parameters, namely the number of channels, compression ratio and attack and release time. The study is designed to investigate possible interaction effects of the parameters with respect to speech intelligibility in fluctuating noise.

3.1.1 Previous research on amplitude compression in fluctuating noise

Many publications have addressed the effects of amplitude compression on speech intelligibility. Their conclusions, focusing on stationary noise, were summarized in chapter 2. Here, an overview is presented with respect to speech intelligibility in fluctuating noise. For clarity the compression parameters are abbreviated as follows: NC = number of frequency channels, CR = compression ratio, T_a = attack time, and T_r = release time.

Reviews

Rintelmann (1972) concluded that almost all research prior to that time had been performed in a quiet surrounding. He specifically suggested the use of various types of background noise for investigating speech intelligibility with compression.

In Villchur's review of amplitude compression hearing aids (1978), he did not distinguish between speech intelligibility in quiet and in background noise. He included only one experiment in fluctuating noise in his review. This experiment (Yanick, 1976a) showed positive results in cafeteria noise with two-channel compression (CR > 1, $T_a < 1$ ms, $T_r = 20$ ms) relative to single-channel compression and linear amplification. Yanick's subjects had a moderate to severe sensorineural hearing loss. In Villchur's view, the gain in the single-channel compressor might have been kept at a rather constant level, since the cafeteria noise was dominated by low-frequency components and the signal-to-noise ratio was low (0 dB). With two-channel compression the noise might have saturated only the low-frequency channel, while the high-frequency channel could actually compress the signal, thus resulting in improved speech intelligibility.

Braida et al. (1979) expressed the same reservations with respect to the results of Yanick (1976a) as Villchur did. They also evaluated results of another study by Yanick (1976b), in which he measured speech intelligibility of subjects with a mild to moderate hearing loss in a speech babble. Compression showed no advantage over linear amplification. Braida et al. concluded that the effects of compression on speech intelligibility (and on comfort and annoyance) had to be studied in a wide variety of realistic environments.

In their reviews on compression, Walker and Dillon (1982), Preves (1991), and Hickson (1994) did not evaluate research in fluctuating noise. Dillon (1996) included some research which used fluctuating noise. However, he did not discern amongst results obtained with different types of background noise. Souza (2002) suggested that compression might increase audibility of speech during gaps in a modulated noise. This could lead to a larger benefit of compression for modulated noise than for unmodulated noise.

In summary, none of the reviews specifically evaluated research with amplitude compression in fluctuating noise. Therefore, the next section evaluates individual studies on fluctuating noise.

Previous research in fluctuating noise classified parametrically

This section evaluates individual studies on fluctuating noise. Since it can be expected that results for fluctuating noise are different from results obtained for stationary noise, experiments included in the evaluation given below concerned fluctuating noise only. The inclusion criteria are chosen equal to those of Chapter 2: compression thresholds were well below the input signal levels, which ensured full dynamic range compression, and all research was conducted with moderately hearing-impaired subjects.

The systematic description is analogous to that of Chapter 2. Compression parameters are indicated by $CP(NC, CR, T_a/T_r)$; this notation is explained in section 2.1.1. The studies described below are summarized (with the included number of subjects) in Table 3.1.

Single-channel compression

CR $\leq$ 3

Tyler and Kuk (1989) $CP(1, 2 + 5, 6/36)$ found better recognition of nonsense syllables in speech babble (signal-to-noise ratio of +5 dB) for compression than for linear ampli-

Table 3.1: *Results of previous research in a fluctuating background noise and for moderately hearing-impaired subjects. The reference is linear amplification. The symbol* ⊗ *indicates that linear amplification was not investigated. Better, equal or worse scores are indicated by* ⊕, ⊙, *and* ⊖, *respectively. Some research could not be classified and is indicated with* ⋆ *instead of a number in the list below. The number of subjects included (n) is shown after the name of the authors.*

	Single-channel (NC = 1)		Multi-channel (NC > 1)	
	CR ≤ 3	CR > 3	CR ≤ 3	CR > 3
$T_r \leq 25$ ms	2⊙	2⊙	2⊙⊙	2⊖⊖
	10⊙⊖		10⊙⊕⊕⊕⊕⊕	
			15⊙⊖	
$25 < T_r \leq 100$ ms	1⊕	1⊙	12⊗	14⊖
	4⊙	5⊙⊖	13⊗	17⊖
	5⊙		14⊖	
	6⊗		16⊖	
	8⊖⊖⊖			
	9⊕			
$T_r \geq 100$ ms	3⊖⊙	7⊖⊖	12⊗	17⊖
	7⊙⊙⊖	11⊖	16⊖⊖	
	8⊖⊖⊖⊖⊖⊖			

1	Tyler and Kuk (1989)	n = 16	11	Peterson et al. (1990)	n = 30
2	Van Buuren et al. (1999)	n = 26	12	Hansen (2002)	n = 6
3	Hickson et al. (1995)	n = 15		*CP(15, 2.1, 1/40+10/400+1/4000+100/4000)*	
4	Kam and Wong (1999)	n = 20	13	Hansen (2002)	n = 6
5	Boike and Souza (2000)	n = 9		*CP(15, 2.1+3, 1/40)*	
6	Neuman et al. (1995)	n = 20	14	Lippmann et al. (1981)	n = 5
7	Neuman et al. (1998)	n = 20	15	Verschuure et al. (1998)	n = 20
	CP(1, 1+1.5+2+3+5+10, 5/200)		16	Moore et al. (2004)	n = 5
8	Neuman et al. (1998)	n = 20	17	Olsen et al. (2004) n = 20	
	CP(1, 1.5+2+3, 5/60+5/200+5/1000)		⋆	Humes et al. (1999)	n = 55 (⊙⊕)
9	Dillon et al. (1998)	n = 140	⋆	Novick et al. (2001)	n = 10 (⊗)
10	Moore et al. (1999)	n = 18	⋆	Van Toor and Verschuure (2002)	n = 38 (⊗)
			⋆	Humes et al. (2004)	n = 53 (⊙)

fication. Although this result was only significant for 2 out of 15 subjects, the trend was found for nearly all subjects.

Van Buuren et al. (1999) *CP(1 + 4 + 16, 1 + 2 + 4, 0/0)* measured speech reception thresholds in a single-talker babble (the disturbing talker had the opposite sex of the speaker). Single-channel compression with CR = 2 resulted in a slightly better speech reception threshold (−0.25 dB) than linear amplification, but this difference was not significant. In stationary noise, this compression resulted in slightly worse intelligibility than obtained for linear amplification, but this was insignificant.

Hickson et al. (1995) *CP(1, 1 + 1.3 + 1.8, 8/50-550)* used a system with an adaptive release time (see section 1.7.4). Release time depended on the duration of the input signal: for short input sounds the release time was short (50 ms), and for long input sounds the release time was longer (550 ms). They measured consonant perception in two different background noises: a multi-talker babble and the sound of dishes in a sink. This last noise was characterized by marked intensity fluctuations. Hickson et al. found a significant interaction between compression ratio and noise type. In the multi-talker babble, compression with CR = 1.8 resulted in the worst consonant score, and it was significantly worse than that obtained for linear amplification. In contrast, this same compression ratio (CR = 1.8) resulted in the best score in the sink noise (significantly better than CR = 1.3, but not significantly different from linear amplification).

Kam and Wong (1999) *CP(1, 1 + 1.1-2.7, 5/30)* found no difference in speech recognition scores between linear amplification and compression in a four-talker babble. Compression ratio was subject and channel dependent, and signal-to-noise ratios ranged from −9 to +9 dB.

Boike and Souza (2000) *CP(1, 1 + 2 + 5 + 10, 3/70)* used a multi-talker babble at a signal-to-noise ratio of +10 dB. Speech recognition scores for CR = 2 did not differ significantly from scores obtained with linear amplification. Subjective quality assessments for CR = 2 were worse than those achieved with linear amplification. However, this difference was insignificant.

Neuman et al. (1995) *CP(1, 1.5 + 2 + 3, 5/60 + 5/200 + 5/1000)* used paired-comparison judgements to investigate the effect of compression on speech quality in three different background noises. They used fluctuating noise (cafeteria, and apartment noise), and stationary noise (ventilation noise, see Chapter 2). A significant interaction was found between noise type and release time. In cafeteria noise (signal-to-noise ratio of +14 dB) subjects preferred long release times (200 and 1000 ms) over a shorter release time (60 ms). In apartment noise no significant difference in release time preference was found. However this last result might have been caused by a high signal-to-noise ratio (+26 dB) and a high compression threshold (6 dB above rms level). In stationary (ventilation) noise, no preference was found for any release time. Their experimental design did not allow for an analysis with compression ratio as main effect.

Neuman et al. (1998) describe two experiments. For both experiments they used the same subjects and noise as Neuman et al. (1995). In the first experiment *CP(1, 1 + 1.5 + 2 + 3 + 5 + 10, 5/200)* they varied compression ratio with only one T_a/T_r-setting. Quality ratings of clarity and pleasantness with all types of noise were worse for CR = 3 than for CR = 1, 1.5, and 2. The perceived amount of background noise increased with increasing compression ratio. In their second experiment Neuman et al. (1998) *CP(1, 1.5 + 2 + 3, 5/60 + 5/200 + 5/1000)* performed a parametric study of the effect of CR and T_r on subjective sound quality with the same subjects and background noises as their previous experiment. In this additional study, the highest compression ratios (CR = 5, and 10) of the previous study were omitted to give way to two extra time constants $T_a/T_r = 5/60$, and 5/1000 ms). The outcome confirmed their previous results that higher compression ratios lead to worse subjective quality ratings (clarity, pleasantness, and perceived amount of background noise). Furthermore, averaging over all three noise types, they found a significant effect of $CR * T_r$. For $CR = 3$ a larger release time resulted in better scores of clarity, pleasantness, and overall quality. For $CR = 2$, the influence of release time was smaller (except for the amount of background noise). For all compression ratios, $T_r = 1000$ ms resulted in better scores than $T_r = 60$, and $T_r = 200$ ms. However, these results were still worse than the linear scores obtained in the previous experiment. The effect of the interaction of release time and noise type (noise $* T_r$) was also significant. For all noise types the ratings of the amount of background noise improved with increasing release time. This effect was more pronounced for cafeteria noise than for apartment noise and the stationary ventilation noise.

Dillon et al. (1998) *CP(1, 1 + 2, 5/50)* performed a field study in which subjects compared compression to linear amplification. Of all subjects, 55% preferred compression to linear amplification, 31% preferred linear amplification, and the remaining 14% did not have a preference.

Moore et al. (1999) *CP(1 + 2 + 4 + 8, 1 + 1-2.9, 7/7)*, used compression in several background noises with spectral and/or temporal dips. Their compression system had a delay to reduce overshoot (see section 1.7.4). Compression ratio was subject dependent. In a single-talker background, speech reception thresholds for single-channel compression did not differ from results obtained with linear amplification. This was also found in stationary noise with spectral dips. Moore et al. also used speech-shaped noise with the modulations of a single speaker. In this background noise, single-channel compression resulted in significantly lower (better) speech reception thresholds than those obtained with linear amplification. In stationary noise with spectral gaps, results for none of the conditions differed significantly form linear amplification or from each other.

CR > 3

Tyler and Kuk (1989) *CP(1, 2 + 5, 6/36)*, quoted above for $CR = 2$, also used $CR = 5$. Whereas for $CR = 2$ results were better than those achieved with linear amplification, for $CR = 5$ no scores differed from those for linear amplification. Compression with $CR = 5$ was also tested in low-frequency stationary noise and this resulted for nearly all subjects in lower nonsense syllable recognition with compression than with linear amplification.

Peterson et al. (1990) *CP(1, 1 + 1-10, 15/180)* measured nonsense syllable scores in cafeteria babble (signal-to-noise ratio of +12 dB) for curvilinear compression, i.e., the compression ratio depended on the input level (see section 1.7.4). Compression ratio ($CR = 1$–10) was subject dependent. Syllable scores obtained with compression were significantly worse than scores for linear amplification.

Neuman et al. (1998) *CP(1, 1 + 1.5 + 2 + 3 + 5 + 10, 5/200)*, quoted above for CR $\leq$ 3, also used CR = 5 and 10. Both these higher compression ratios resulted in worse scores than scores obtained with lower compression ratios.

Van Buuren et al. (1999) *CP(1 + 4 + 16, 1 + 2 + 4, 0/0)*, quoted above for single-channel compression with CR = 2, also used compression with a compression ratio of 4. Speech reception thresholds with CR = 4 were worse than those obtained with CR = 2 and than those obtained with linear amplification, however, the differences were not significant. Note that in stationary noise, single-channel compression with CR = 4 resulted in significantly worse results than those for linear amplification.

Boike and Souza (2000) *CP(1, 1 + 2 + 5 + 10, 3/70)* were quoted above for CR = 2 with speech recognition that did not significantly differ from scores obtained with linear amplification. They also used higher compression ratios and found speech recognition to decrease with increasing compression ratios (going from 2 to 5 to 10). Speech quality also decreased with increasing compression ratio, although only compression with CR = 10 differed significantly from linear amplification.

In summary, when applying single-channel compression in fluctuating noise for subjects with a moderate sensorineural hearing impairment, few experiments showed consistently better speech intelligibility with compression than with linear amplification. Best results were achieved with compression ratios smaller than 3. The effect of time constants on speech intelligibility has not been sufficiently investigated and is still not clear. Sound quality experiments showed that compression ratios larger than 2 degraded quality. Longer release times (up to at least 1000 ms) decreased the negative effect of compression on sound quality, but did not restore it to the level achieved with linear amplification. Results from Moore et al. (1999), Van Buuren et al. (1999) and Tyler and Kuk (1989) suggest that single-channel compression might be slightly less detrimental in fluctuating noise than in stationary noise.

Multi-channel compression

CR $\leq$ 3

Lippmann et al. (1981) *CP(16, 1 + 1-3 + 1-5, $<$ 6/20-6/32)* used two sixteen-channel compression systems in cafeteria noise (signal-to-noise ratio of +10 dB). Compression ratios were subject and frequency dependent. For CR = 1–3, the measured phoneme correct scores were slightly worse than for those obtained with linear amplification.

Verschuure et al. (1998) *CP(2, 1 + 2, 5/15)* evaluated compression with overshoot reduction in four real-life background noises: restaurant noise, noise of an industrial plant, printing office noise, and city background noise. For restaurant and industrial noise, speech intelligibility did not differ between compression and linear amplification. However, for both printing office and city noise, results with compression were significantly lower than with linear amplification.

Moore et al. (1999) *CP(1 + 2 + 4 + 8, 1 + 1-2.9, 7/7)*, quoted above for single-channel compression, also used two-, four, and eight-channel compression. Their compression was subject and frequency dependent. The results showed a significant interaction between background noise and number of channels. In a single-talker background, results obtained for two-channel compression did not differ from those for single-channel compression or linear amplification. However, four- and eight-channel compression resulted in significantly better speech reception thresholds than linear amplification. In modulated speech-shaped noise all compression conditions (NC = 1, 2, 4, and 8) resulted in significantly better scores than achieved with linear amplification. For this noise, the compression conditions did not differ significantly across number of channels. Note that in stationary noise (with spectral gaps) none of the conditions differed significantly from linear amplification or from each other.

Van Buuren et al. (1999) *CP(1 + 4 + 16, 1 + 2 + 4, 0/0)*, quoted above for single-channel compression (with results slightly better than those for linear amplification), also used multi-channel compression. Both four- and sixteen-channel compression with CR = 2 resulted in slightly worse speech intelligibility than with linear amplification; the differences, however, were statistically insignificant. In stationary noise results for this compression (CR = 2; NC = 4, and 16) were also worse than those obtained for linear amplification; these differences were nearly significant at the 5% level.

Hansen (2002) *CP(15, 2.1, 1/40 + 10/400 + 1/4000 + 100/4000)* and *CP(15, 2.1 + 3, 1/40)* investigated the effect of time constants on the subjective quality of speech in several real-life background noises, and music. The longest release time (4000 ms) resulted in higher quality scores than 400 and 40 ms, for both speech in noise and music. With the longest release time, the attack time (1 and 100 ms) had a minor influence. With fast compression ($T_a/T_r = 1/40$), a compression ratio of 2.1 resulted in significantly better quality scores than a compression ratio of 3, for both speech in noise and music.

Moore et al. (2004) compared four compression systems with different time constants and linear amplification. They used three twenty-channel systems: a fast system *CP(20, 1-1.7, 8/32)*, a slow-fast system *CP(20, 1-1.5, 500-100/500-100)* in which T_a and T_r decreased from 500 ms for the low frequencies to 100 ms for the high frequencies, and a fast-slow system *CP(20, 1-1.7, 50-500/50-500)*, in which T_a and T_r increased from 50 ms for the low frequencies to 500 ms for the high frequencies. They also used a ten-channel dual system consisting of a slow compression system followed by a fast system *CP(10, 1-1.7, 20000/20000 + 20-30/100-150)*. Compression ratios were subject and channel dependent, but they were not specified. The values for CR included in the *CP* specification were calculated from the reported average insertion gain at input levels of 80, 65 and 50 dB SPL. Moore et al. measured intelligibility of VCV nonsense syllables in three types of background noise: single-talker babble, multi-talker babble, and cafeteria noise. No significant differences were found between the nonsense syllable scores for the four compression systems. Scores obtained for linear amplification were significantly better than those for compression.

CR $>$ 3

Lippmann et al. (1981) *CP(16, 1 + 1-3 + 1-5,$<$ 6/20-32)* used a system that was fitted according to the pure-tone dynamic range of the subjects. Because this led to compression ratios of up to 5 for the high frequency channels, they also used a fitting with lower compression ratios (CR = 1–3, quoted above). Both systems resulted in worse scores than those for linear amplification, and CR = 1–5 resulted in slightly worse scores than CR = 1–3.

Van Buuren et al. (1999) *CP(1 + 4 + 16, 1 + 2 + 4, 0/0)*, quoted above, used four- and sixteen-channel compression with a compression ratio of 4. They obtained significantly worse speech reception thresholds for compression than for linear amplification. In stationary noise, results for two these systems (NC = 4 and 16, with CR = 4) were also significantly worse than those obtained for linear amplification.

Humes et al. (1999) *CP(2, 1 + 1-4, ?/?)* used a wearable two-channel compression system with an adaptive release time constant. Compression ratio was subject and frequency dependent. Speech was presented at two input levels (60 and 75 dB) and at two signal-to-noise ratios (+5 and +10 dB) in a multi-talker babble. They used a monosyllabic word test, a connected speech test, and a test of subjective ease of listening. At a presentation level of 60 dB, compression gave slightly better results than linear amplification (for both signal-to-noise ratios). This was found for all three tests, but it was only significant for the monosyllabic word test. At 75 dB, results were mixed: with the monosyllabic word test the compression system gave better results (only significant for a signal-to-noise ratio of +10 dB). For the other two tests, the compression system gave worse results than the linear system, but this was insignificant.

Novick et al. (2001) *CP(2, ?, ?/40 + ?/160 + ?/320 + ?/640)* used a wearable compression system with a directional microphone. Compression ratio was not specified but it was subject and frequency dependent and was fitted according to a proprietary algorithm. A hearing-in-noise test in a twelve-talker babble (signal-to-noise ratio of +8 dB) showed no significant effect of release time on speech intelligibility.

Van Toor and Verschuure (2002) *CP(4, ?, 2/16-64 + 2-16/64-512 + 32-64/1024-2048)* compared three wearable, four-channel compression systems. They used a fast, an intermediate, and a slow system. Time constants changed over channels, the high frequency channels being the fastest. All systems used a short signal delay to reduce overshoot. Compression was curvilinear, i.e., the compression ratio depended on the input level. The systems were fitted according to the DSL [i/o] algorithm (Cornelisse et al., 1995) with both a flat response, and a fitting with high-frequency emphasis. They used three types of noises, stationary speech-shaped noise, fluctuating speech-shaped noise, and low-frequency car noise. No significant effect of time constants on speech reception threshold was found for both the stationary speech-shaped noise and the low-frequency car noise. Only for fluctuating noise did time constants have a significant effect. With the flat frequency response the slow system gave significantly less improvement than the intermediate or fast system. When the data was averaged over all noises (including stationary noise) and time constants, compression with high-frequency emphasis gave significantly better results than compression with a flat frequency response. A subjective performance test (APHAB, developed by Cox and Alexander (1995) and translated into Dutch) also yielded no consistent preferences for any set of time constants.

Humes et al. (2004) *CP(2, 1 + 2-4, ?/?)* compared a wearable compression hearing aid to a linear aid. They tested speech intelligibility in a multi-talker background after 1 month and 6 months acclimatization to the new aids. They also measured user satisfaction and sound quality. No significant differences were found between the two systems.

Olsen et al. (2004) *CP(3, 1 + 2/5 + 5/10, 5-60/11-60 + 100-135/300-400)* compared four fast-acting compression systems to linear amplification in fully modulated noise. Additionally they used a fifth system with a delay to reduce overshoot. The nominal compression ratio was 2 or 5. For channels with a hearing loss exceeding 60 dB the compression ratio was increased to 5 or 10. This occurred (mostly) for frequencies above 2000 Hz. The linear system resulted in significantly better speech intelligibility than the compression systems.

In summary, for multi-channel compression in fluctuating noise, results were mostly worse than for linear amplification. Only the experiments of Moore et al. (1999) showed improved speech intelligibility for multi-channel compression. This improved speech intelligibility was not found for stationary noise. Previous studies did not reveal a systematic effect of time constants, number of channels and noise type on speech intelligibility.

3.1.2 Conclusions based on previous research

The available results from previous research do not present a clear picture of the effects of compression in fluctuating noise. Although some research did show improved speech intelligibility for amplitude compression relative to linear amplification, most studies either showed a degradation or no difference. Compression results (relative to linear amplification) tended to be slightly more favourable in fluctuating noise than in stationary noise.

As was the case for stationary noise, for fluctuating noise possible interaction effects of the compression parameters are still unknown. The current study is intended to investigate the effect of compression in fluctuating noise, and possible interaction effects. We therefore conducted a full parametric investigation with respect to the combination of number of channels, compression ratio and the two time constants.

3.2 Methods

The present experiment was performed in conjunction with the experiment in stationary noise. Experimental methods and subjects in this experiment are the same as for stationary noise. Most of the details have been described in Chapter 2. Appropriate values of the compression parameters are based on previous research and are chosen identical to the values for compression in stationary noise (as described in Table 2.2).

All compression conditions were presented with two types of background noise: speech-shaped fluctuating noise (the results of which are presented here) and speech-shaped stationary noise (Chapter 2). Ten subjects selected at random performed the experiment with stationary noise first and with fluctuating noise second, the other ten subjects vice versa.

Speech materials consisted of Dutch sentence material for a Speech Reception Threshold in noise test (SRT test), developed by Versfeld et al. (2000). This material consisted of two sets (a male and a female speaker) of 39 lists of 13 sentences each. For each type of noise (stationary or fluctuating) each subject was presented with materials from one speaker only in all conditions.

The fluctuating noise used in this experiment consisted of the long-term average speech spectrum (LTASS) of each speaker with the modulations of a single talker. This noise was constructed as follows. For each speaker, fifty sentences were concatenated and passed through the same six-channel filter bank as used for compressing the stimuli. From the output of each channel, the speech signal envelope was calculated by means of a Hilbert transform. The (six-channel) speech-shaped noise was then multiplied by these speech envelopes. After multiplication the outputs of the six channels were summed. This resulted in a single-talker modulated noise with the frequency content of the original speech material. Since this six-channel modulated noise might to some extent be intelligible, we time-reversed this noise to avoid informational masking (Summers and Molis, 2004).

The speech stimuli were embedded in the fluctuating background noise. For each stimulus the starting point of the fluctuating noise was randomly chosen. The noise extended from 2 seconds before speech (onset), to 1 second after speech (offset). The dynamic range of both the original speech and the fluctuating noise was 26 ± 6 dB (the standard deviation was calculated from the dynamic range of all the separate sentences; wide band measurements between the 1^{st} and 99^{th} percentile, applying an integration time constant of 125 ms). Within each list, sentence order was randomized.

This was done to minimize possible recognition effects for a repeated list. The set of sentences per list was fixed because the lists were balanced with respect to intelligibility. For each subject a particular list was used only once or twice.

In order to estimate test-retest variance, measurements were made in duplicate. Due to time constraints the duplicate measurements were made for only 13 predetermined conditions. These duplicate conditions were evenly distributed over the experiment consisting of several (3 to 6) sessions (2^{nd}, 5^{th}, 7^{th}, 10^{th} measurement, etc.). In total, the same 50 compressive processing conditions as the previous chapter were used. The full parametric study consisted of three numbers of channels, four compression ratios, and five sets of time constants. The numbers of channels were 1, 2, and 6 channels (see Table 2.2). The compression ratios (CR = 1, 2, or 3) were divided in two separate ratios for low and high frequencies. The cross-over frequency was 1 kHz. Note that although the compression ratio in some channels might be equal, all channels were independently compressed. The time constants consisted of 5 combinations of attack (4–40 ms) and release times (4–400 ms), see Table 2.2.

3.3 Results

3.3.1 Learning effect

Using the duplicate measurements, a t-test for dependent samples (paired by condition) showed a significant improvement (0.6 dB) of the retest score relative to the test score ($p < 0.0005$). No significant interaction was found between this learning effect and subject ($F = 1.0$; $p = 0.5$), or compression condition (NC, CR, or T_a/T_r) ($F = 0.7$; $p = 0.7$). Test-retest measurements which occurred in the same session did not differ significantly from test-retest pairs in two different sessions ($F = 1.7$; $p = 0.2$). Regression analysis over all subjects showed that the learning effect depended only on the number of previous measurements, irrespective of the sessions involved (+0.023 dB per presentation number, $r^2 = 0.02$, $p < 0.05$). Figure 3.1 gives the difference between the retest and the test score for all subjects as a function of presentation number. Although the learning effect was small, it was significant and therefore a correction was applied. The maximum correction was +1.5 dB at the end of the experiment.

The learning effect in fluctuating noise corresponded well to the effect in stationary noise (see Fig. 2.1). For both background noises the learning effect depended only on the number of previously encountered lists/conditions and not on the number of previous sessions. The correction for fluctuating noise was slightly larger than for stationary noise (+0.023 and +0,019 dB per presentation number, respectively).

The learning effect did not depend on which noise type was presented first to the subjects. Figure 3.2 shows the learning effect for the sequence in which the noise type was

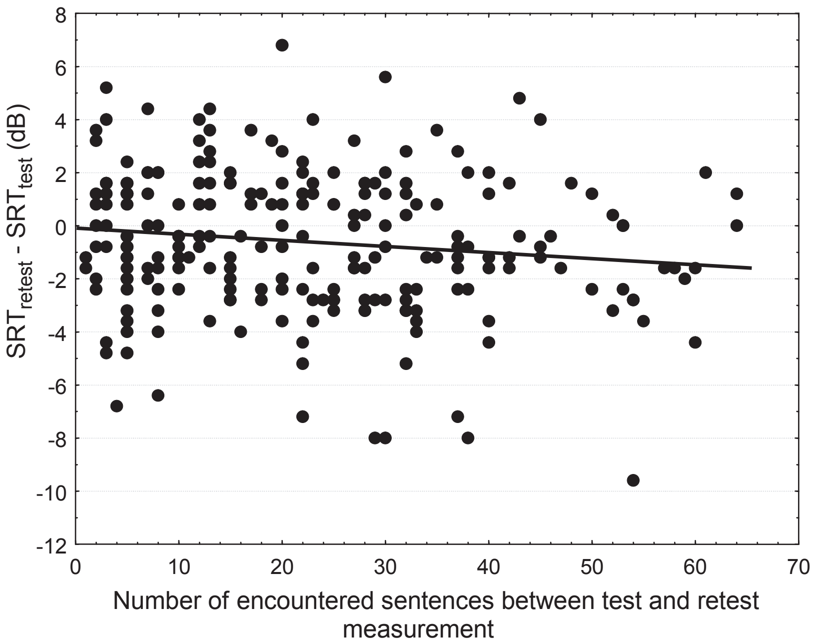


Figure 3.1: *Retest SRT minus test SRT for fluctuating noise, for all subjects. The test and the corresponding retest conditions are distributed at random over the entire experiment. The abscissa represents the difference in serial position between test and retest. A negative value of $SRT_{retest} - SRT_{test}$ represents a better retest. The regression line was significant: -0.023 dB per presentation number, $r^2 = 0.02$, $p < 0.05$.*

presented (first or second; represented by the two columns) and for each noise type (rows).

The average variance of the duplicate measurements for fluctuating noise was 3.6 dB^2 (after correction for the learning effect). This variance was independent of compression condition ($F = 1.4$, $p = 0.2$). Prior to further analysis all the duplicate pairs were averaged and the average values were included in the analysis as one data point. Subsequent analyses do not include the variance of the duplicate pairs in their error estimation.

3.3.2 Speech intelligibility in fluctuating noise compared to stationary noise

The average aided performance with linear amplification of individual subjects in fluctuating noise spanned a range between −10.5 dB and +7.9 dB. As expected, this range was larger (two fold) than the range for stationary noise (−2.9 to +6.1 dB). Across all subjects, the average SRT obtained for fluctuating noise (mean SRT $= -2.6 \pm 4.1$ dB) was 2.8 dB lower than that obtained for stationary noise (+0.2 ± 2.5 dB).

In Figure 3.3 the SRT for fluctuating noise is plotted against the SRT for stationary noise. Only SRTs obtained with linear amplification are shown. Regression analysis showed a highly significant correlation between the SRT for the two noise types across

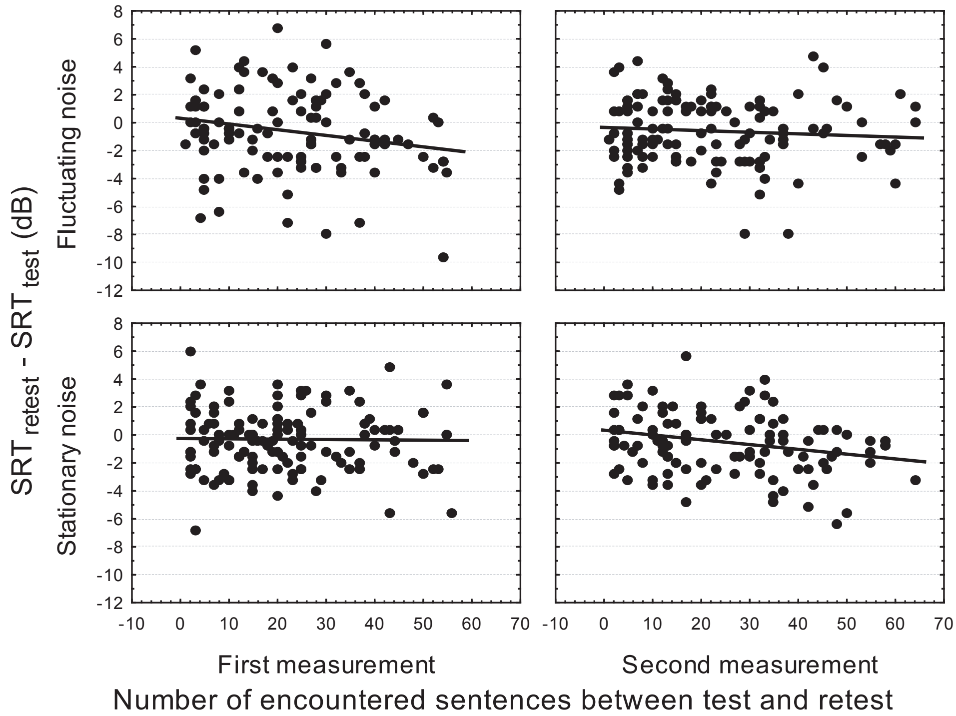


Figure 3.2: *Retest SRT minus test SRT for all subjects. The left two graphs show the data for the first presented noise type, the right graphs for the second. The upper row represents fluctuating noise, the lower row stationary noise. The test and the corresponding retest conditions are distributed at random over the entire experiment. The abscissa represents the difference in serial position between test and retest. A negative value of $SRT_{retest} - SRT_{test}$ represents a better retest. Two (out of four) regression lines were significant: the first presentation of fluctuating noise (top left) with an average improvement of 0.039 dB per presentation number ($r^2 = 0.04$, $p < 0.05$) and the second presentation of stationary noise (bottom right) with an average improvement of 0.034 dB per presentation number ($r^2 = 0.06$, $p < 0.01$).*

subjects ($r^2 = 0.67$; $p = 0.00001$). The difference between the two noise types increased at lower SRTs (i.e., less hearing impairment). In our subjects, the SRT for fluctuating noise deteriorated with increasing hearing loss; linear regression analysis showed a significant correlation between the SRT for fluctuating noise and pure tone average hearing loss at 0.5, 1, and 2 kHz ($r^2 = 0.3$, $p < 0.01$). For stationary noise, this trend was insignificant ($r^2 = 0.11$; $p = 0.15$, Fig. 2.2). As stated in the introduction, one of the rationales for using fluctuating noise is the common finding that speech intelligibility in normal-hearing listeners is better in fluctuating noise than in stationary noise, whereas sensorineurally hearing-impaired listeners have less benefit from the fluctuations. Figure 3.4 shows the difference between the SRT obtained for both kinds of noise (linear amplification) as a function of the subject's pure tone average hearing loss across 0.5, 1, and 2 kHz ($PTA_{0.5, 1, 2, kHz}$). The figure shows that the benefit of the fluctuations decreased with increasing hearing loss.

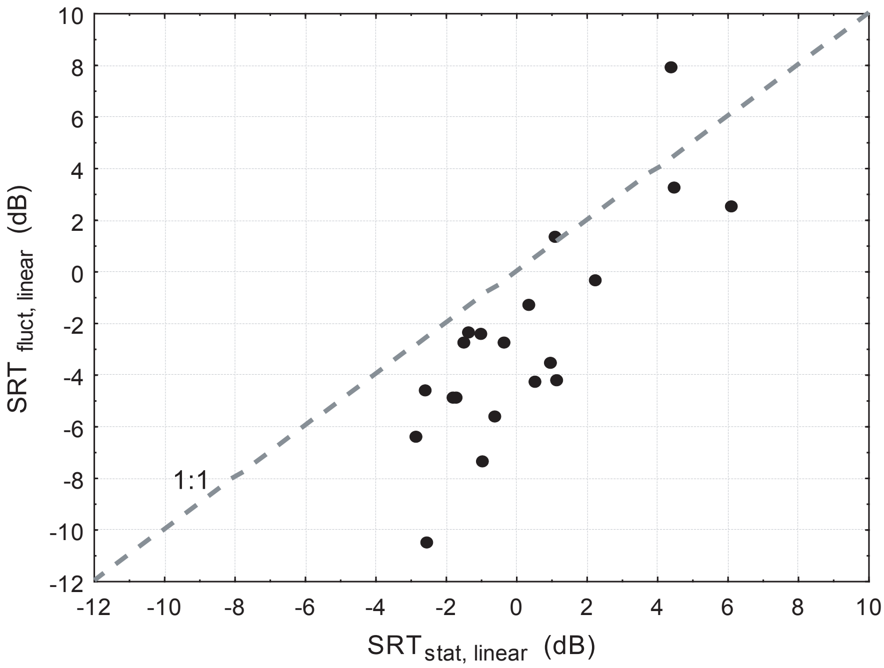


Figure 3.3: *SRTs obtained with linear amplification. The abscissa gives the speech reception thresholds for speech in stationary noise ($SRT_{stat,\ linear}$), the ordinate for speech in fluctuating noise ($SRT_{fluct,\ linear}$). The dashed line represents equal SRTs for fluctuating and stationary noise. Each dot shows data for one subject.*

Psychometric function for speech intelligibility in noise

Our psychometric function relates speech intelligibility in noise to the signal-to-noise ratio at which this intelligibility was obtained. Speech intelligibility is measured by the sentence correct score, which constitutes the ratio of the number of entirely correct responses and the total number of sentences that have been presented at a certain signal-to-noise ratio.

We did not measure the sentence correct score directly; instead we used an adaptive measurement procedure to obtain the speech reception threshold (section 2.2.3). The adaptive procedure is very efficient for estimating the signal-to-noise ratio at which 50% of the sentences were correctly repeated (i.e., the SRT), but is not optimized for estimating the slope of the psychometric function (Brand, 2000). However, our step size of 2 dB in combination with the large number of measurements yields a range in the excursions of the adaptive procedure around the 50% point sufficient for a reliable estimate of the psychometric slope.

The percentage of correct responses was calculated as follows. Data was pooled over different lists and subjects. This was possible because the signal-to-noise ratio for each sentence was expressed relative to the measured SRT for that list. Consistent with the adaptive SRT procedure, the data for the first three sentences of each list were

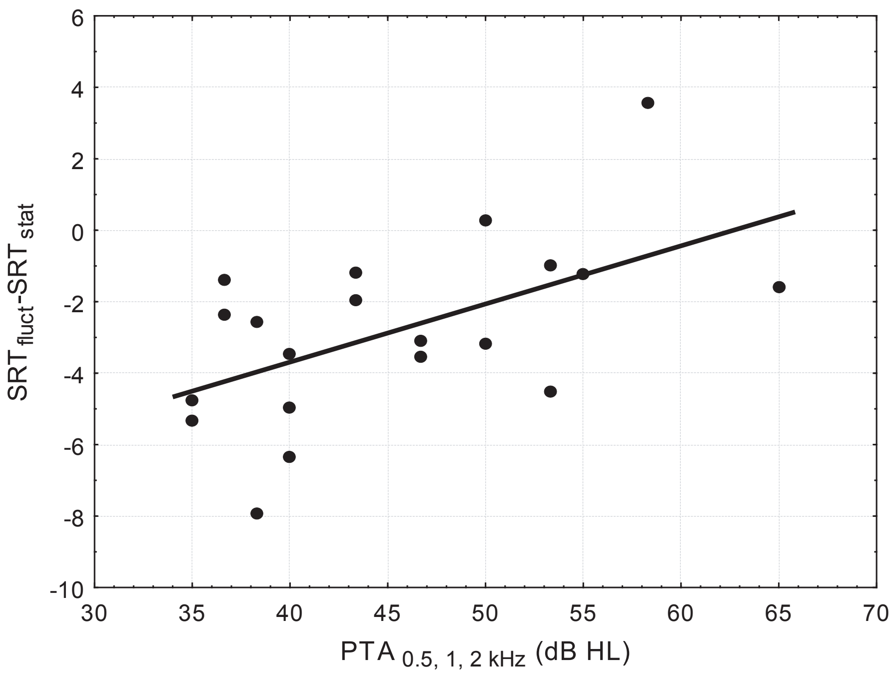


Figure 3.4: *Benefit derived from fluctuations in the background noise (release of masking) as a function of average pure tone hearing loss ($PTA_{0.5,\,1,\,2\,kHz}$). Lower values indicate a higher release of masking. Linear regression analysis yielded a significant correlation ($r^2 = 0.3$, $p = 0.01$). Linear amplification only. Each dot represents one subject.*

discarded. The percentage of completely correct responses was then calculated for each signal-to-noise ratio.

Linear amplification

Fig. 3.5 gives the psychometric function with linear amplification for both fluctuating and stationary noise. In order to obtain a reliable fit, the data of all subjects were pooled after which a single fitting was obtained. Due to the modulations in the fluctuating noise the signal-to-noise ratio differs between the words in a sentence. Fluctuating noise therefore tends to yield a shallower psychometric function than stationary noise. Indeed, the maximum slope of the psychometric function was 17.8 %/dB for fluctuating noise and 22.0 %/dB for stationary noise.

For adaptive SRT measurements, the slope of the psychometric function at the 50% point is assumed to be inversely proportional to the square root of the measurement error (Brand, 2000). Indeed, the ratio of the slopes ($slope_{stat}/slope_{fluct} = 1.24$) was about equal to the ratio of the square root of the measurement errors ($= 1.25$).

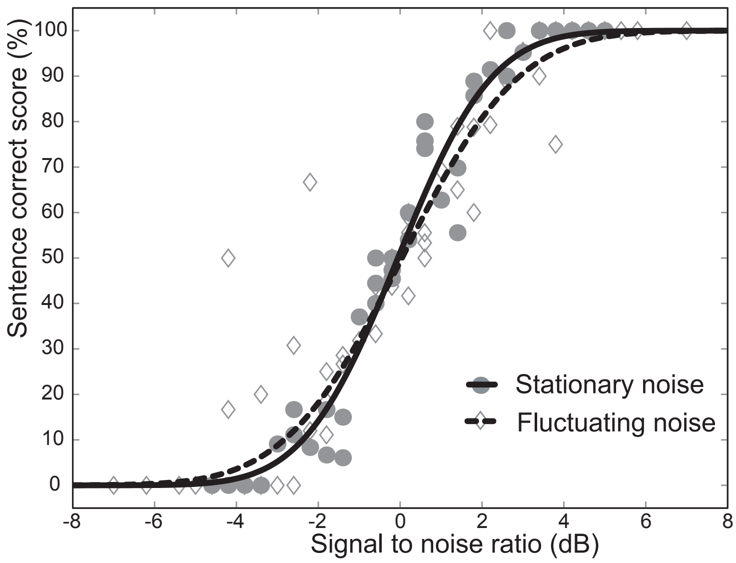


Figure 3.5: *Average discrimination curves for sentences with linear amplification. For each measurement the SRT was translated to 0 dB by shifting the data along the abscissa for each sentence according to the measured SRT for the corresponding list. A cumulative normal distribution was fitted to the data:* $S = \frac{1}{2}\,\mathrm{erf}\,(-(SNR-\mu)/\sigma\sqrt{2})$. *In which* S *is the sentence score and SNR is the signal-to-noise ratio. Both the mean* μ *and standard deviation* σ *were estimated according the least squares procedure* (μ_{fluct} = +0.04; $\sigma_{fluct} = 2.2$; $\mu_{stat} = -0.05$; $\sigma_{stat} = 1.8$).

Compression

The maximum slope of the psychometric function at the 50% point for compression conditions (excluding linear amplification) was 16.2 %/dB and 17.8 %/dB for fluctuating and stationary noise, respectively. For this calculation data was pooled across all compression conditions and all subjects. The slope for fluctuating noise decreased less under the influence of compression than the slope for stationary noise (1.6 %/dB, and 4.2 %/dB, respectively). For compression, the ratio of the slopes of the psychometric function ($\text{slope}_{\text{stat}}/\text{slope}_{\text{fluct}} = 1.1$) and the ratio of the square root of the measurement errors ($= 1.25$) were not equal anymore.

3.3.3 Concurrent analysis of single-, two-, and six-channel compression

All SRTs that were measured for compression conditions will be presented relative to the linear scores (like in Chapter 2). For this purpose, ΔSRT is defined as the SRT for a particular compression condition minus the SRT for the linear condition found for that subject. Thus, negative values of ΔSRT indicate a better speech reception threshold for compression than for linear amplification.

Single-channel compression does not allow for frequency dependent compression ratios, thus CR = 1/2 and CR = 2/3 were not included in this first analysis. A multivariate[1] repeated measures analysis of variance was used to analyse ΔSRT. Number of channels (3 levels) compression ratios (2 levels) and time constants (5 levels) were all treated as within subject variables. Table 3.2 shows the results of the analysis.

Table 3.2: *Multi-variate repeated measures analysis of variance. Single-channel compression is included and therefore frequency dependent compression ratios are omitted. (NC = 1, 2, 6; CR = 2/2, 3/3; T_a/T_r = 4/4, 4/40, 40/40, 4/400, 40/400).*

Effect	Degrees of freedom	Wilks' Lambda	F	p-value
NC	2	0.65	4.8	< 0.05
CR	1	0.96	0.7	0.4
T	4	0.62	2.5	0.09
NC * CR	2	0.98	0.1	0.9
NC * T	8	0.55	1.2	0.4
CR * T	4	0.84	0.8	0.6
NC * CR * T	8	0.54	1.3	0.3

In Figure 3.6 ΔSRT is given as a function of time constants, CR being the parameter and a separate panel for each number of channels. The ordering of time constants is primarily based on release time.

Interaction effects did not reach significance ($p \geq 0.3$). Only the main effect of number of channels was significant. Post-hoc analysis (HSD test) showed that six-channel compression yielded worse speech intelligibility than single-channel ($\Delta SRT_{NC=6} - \Delta SRT_{NC=1} = +0.4$ dB, $p < 0.05$) and two-channel compression ($\Delta SRT_{NC=6} - \Delta SRT_{NC=2} = +0.4$ dB, $p < 0.05$). Results for NC = 1 and NC = 2 were about equal ($\Delta SRT = -0.1$).

Fig. 3.6 shows that CR = 3/3 yielded (insignificantly) higher ΔSRTs than CR = 2/2. For time constants the best results (low ΔSRTs) occurred for $T_a/T_r = 40/40$. However, this factor was not significant ($p = 0.09$, Table 3.2). The low ΔSRT for $T_a/T_r = 40/40$ was largely caused by the good results for NC = 1, and 6.

None of the data points in Fig. 3.6 differed significantly from the results obtained with linear amplification. The best (but statistically insignificant) result was obtained for $CP(1, 2/2, 40/40)$: $\Delta SRT = -0.9 \pm 2.3$ dB. Note that for this condition the results with CR = 2/2 and CR = 3/3 are nearly equal, $\Delta SRT = -0.8 \pm 2.4$ dB for $CP(1, 3/3, 40/40)$.

3.3.4 Analysis with the inclusion of split-frequency compression ratios (1/2 and 2/3)

Frequency dependent compression ratios were omitted in the previous analysis be-

[1] Mauchley's test of sphericity yielded a significant result on T_a/T_r ($p < 0.01$). A multivariate analysis is thus required.

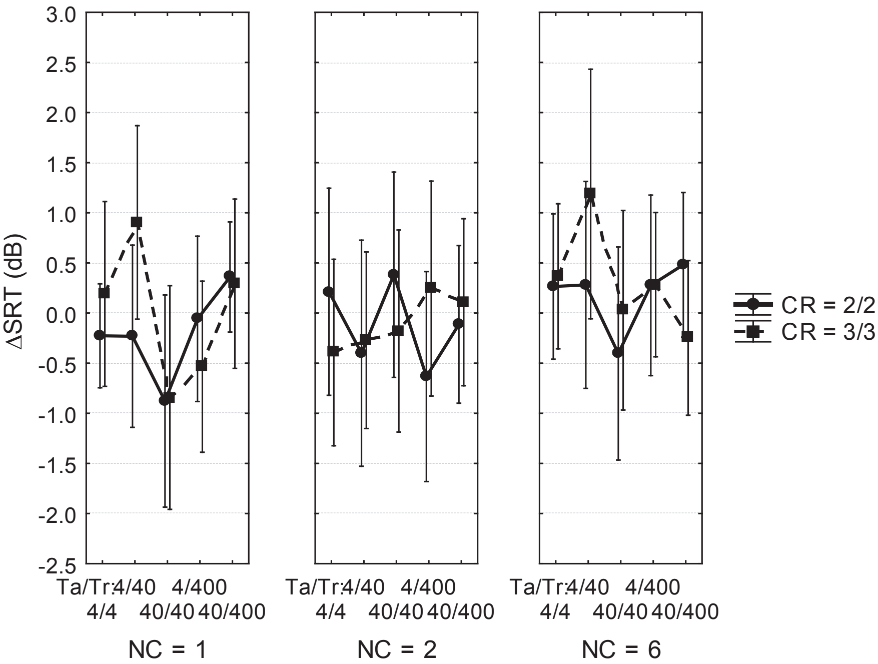


Figure 3.6: *Average results, ΔSRT is plotted as a function of time constants. ΔSRT is the SRT for compression minus the subjects' SRT for linear amplification, and lower values represent better speech intelligibility with compression. Each curve represents a compression ratio setting. The panels show results for single-, two-, and six-channel compression. Only CR = 2/2 and CR = 3/3 are included. Vertical bars denote 0.95 confidence intervals.*

cause we included single-channel compression. An additional multivariate[2] analysis was conducted on ΔSRT in which split-frequency compression ratios (1/2 and 2/3) were included. Number of channels (NC = 2, 6), compression ratio (1/2, 2/2, 2/3, and 3/3) and time constants (5 levels) were all treated as independent within subjects variables. The results of the analysis are given in Table 3.3.

Figure 3.7 shows ΔSRT as a function of time constants. The panels represent number of channels and the curves show data of the four compression ratios.

As for the previous analysis, the interaction effects did not reach significance. Again, only the main effect of number of channels was significant ($F = 5.5$; $p < 0.05$). As in the previous analysis, the average results for NC = 6 were worse than those for NC = 2 ($\Delta SRT_{NC=6} - \Delta SRT_{NC=2} = +0.3$ dB, $p < 0.05$, HSD test). This detrimental effect for NC = 6 was very consistent for all compression ratios and time constants. No significant result was found for compression ratio ($F = 2.0$; $p = 0.2$). A compression ratio of 2/3 gave the best results, CR = 3/3 the worst. The main effect of time constants was insignificant as well ($F = 0.2$; $p = 0.9$). The differences in scores were small (< 0.3 dB) and the best result was obtained for $T_a/T_r = 40/40$. The overall best result $\Delta SRT = -0.7$

[2]Mauchley's test of sphericity yielded a significant result ($p < 0.05$) for the highest interaction term (NC*CR*T). A multivariate analysis is thus appropriate.

Table 3.3: *Multi-variate repeated measures analysis of variance. Split-frequency compression conditions were included, and single-channel compression was omitted. NC = 2, 6. CR = 1/2, 2/2, 2/3, 3/3.* T_a/T_r *= 4/4, 4/40, 40/40, 4/400,40/400.*

Effect	Degrees of freedom	Wilks' Lambda	F	p-value
NC	1	0.77	5.5	< 0.05
CR	3	0.74	2.0	0.2
T	4	0.95	0.2	0.9
NC*CR	3	0.98	0.1	0.9
NC*T	4	0.90	0.4	0.8
CR*T	12	0.57	0.4	0.9
NC*CR*T	12	0.39	1.0	0.5

($\pm$2.5) dB was obtained for $CP(2, 2/3, 40/40)$.

3.3.5 Analysis with attack and release time

In the previous analyses, T_a and T_r were combined. An additional repeated measures analysis was conducted on ΔSRT, with both T_a and T_r as independent within variables ($T_a = 4$, 40 ms; $T_r = 40$, 400 ms). All compression ratios and number of channels were included. The analysis gave no significant main (T_a or T_r) or interaction ($T_a * T_r$) effects. Although insignificant, the results indicated that the longest attack time yielded slightly better ΔSRTs than the shortest one ($\Delta SRT_{T_a=40ms} - \Delta SRT_{T_a=4ms} = -0.1$dB). In contrast, the longest release time yielded slightly worse results than the shortest one ($\Delta SRT_{T_r=400ms} - \Delta SRT_{T_r=40ms} = +0.1$ dB).

3.4 Discussion

3.4.1 Results for fluctuating noise compared to results for stationary noise

Release of masking in fluctuating noise

No standard fluctuating noise was available for our speech materials. We therefore constructed a fluctuating noise by introducing the modulations of a speaker on stationary speech-shaped noise while reversing the modulations in time to avoid informational masking (section 3.2). In this section we will compare results of our fluctuating noise to the results from other research.

For normal-hearing listeners, speech intelligibility is generally better in fluctuating noise than in stationary noise with the same long-term energy content. This 'release of masking' is caused by the temporal gaps in the fluctuating noise. During these gaps

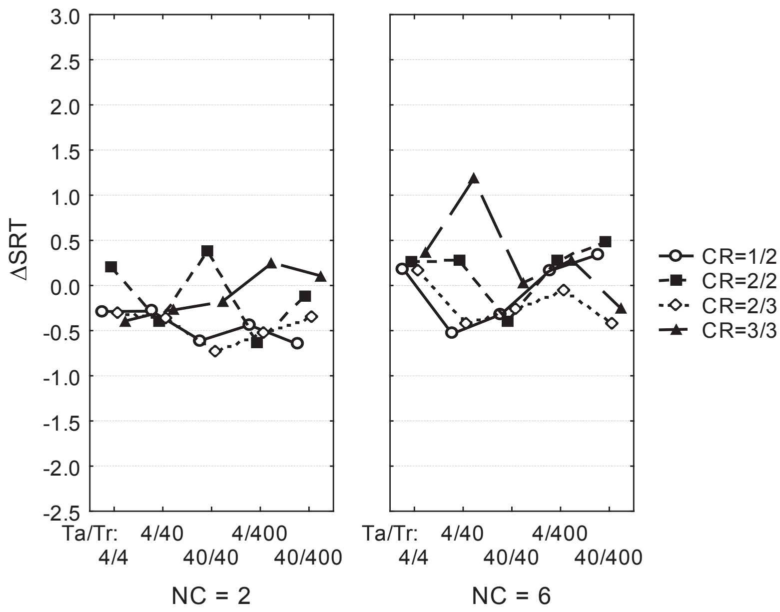


Figure 3.7: Δ*SRT as a function of time constants (T_a/T_r), compression ratio (CR) and number of channels (NC).* Δ*SRT is the SRT for compression minus the subjects' SRT for linear amplification, and lower values represent better speech intelligibility with compression. Split-frequency compression ratios (CR = 1/2 and 2/3) have been included, single-channel compression has been omitted. The ordering of T_a/T_r is primarily based on release time. Confidence intervals are omitted for clarity. The average 0.95 confidence interval was 1.5 dB.*

the noise level is relatively low and this can lead to a momentary high signal-to-noise ratio. The amount of forward masking depends on the duration of the gaps in the noise and on the temporal acuity of the listener, see section 1.3.3. Less forward masking leads to a larger release of masking (Festen and Plomp, 1990; Gustafsson and Arlinger, 1994; Bronkhorst and Plomp, 1992; Dubno et al., 2003). The reason that hearing-impaired listeners have typically a smaller release of masking than normal-hearing listeners may lay in the fact that sensorineurally hearing-impaired subjects have shown slower recovery (in dB/ms) from a masker noise, even when audibility effects were taken into account (Nelson and Freyman, 1987), see Fig. 1.6.

In our case, release of masking can be expressed in the difference between the SRT for fluctuating and for stationary noise. Since release of masking is generally less for hearing-impaired listeners, we expect that the SRT for fluctuating noise and for stationary noise approach each other for listeners with a larger hearing loss. The data from Fig. 3.3 agrees with our expectation: the difference between SRT_{fluct} and SRT_{stat} was largest for the lowest SRTs. Correlation analysis indicated that smaller hearing losses (i.e., lower SRTs) tended to show a larger benefit (up to 8 dB) from fluctuations in the background noise (see Fig. 3.4). Averaged over all subjects, the results for fluctuating noise were 2.8 dB better than for stationary noise.

We will attempt to compare this benefit from fluctuations to results from previously performed experiments. Eisenberg et al. (1995) found for single-channel sinusoidally modulated noise (with a modulation frequency of 31.5 Hz) a 1.9 dB better SRT than for stationary noise. On average, their subjects had less sensorineural hearing impairment than our subjects ($PTA_{0.5,1,2\,kHz}$ ranged from 25 to 37 dB, with an average of 30 dB; the average age was 62 yr).

Festen and Plomp (1990) also used single-channel modulated noise. They used speech modulations rather than sinusoids. Their subject population had a sensorineural hearing loss which was comparable to that of our population (their subjects' $PTA_{0.5,1,2\,kHz}$ ranged from 35 to 58 dB with an average of 47 dB; the average age was 57 yr). They found an average difference in SRT between fluctuating and stationary noise of 1.4 dB. Besides single-channel modulated noise, Festen and Plomp (1990) also used noise with two-channel speech modulations. For this background they found slightly worse speech intelligibility (0.5 dB) than for stationary noise. In contrast to the results of Festen and Plomp, our results indicate a relatively large benefit (on average 2.8 dB) for six-channel modulated noise. This difference is probably caused by differences in experimental set-up. First, we used frequency dependent amplification, whereas theirs was frequency independent. Our amplification assured a higher audibility of speech segments, which might have resulted in a larger release of masking. However, it should be noted that some studies have shown that the effect of audibility is limited, since audibility effects alone are not sufficient to explain the difference in release of masking between normal hearing subjects and hearing impaired subjects (Eisenberg et al., 1995; Bacon et al., 1998). Second, our background noise was time-reversed whereas the fluctuating noise of Festen and Plomp (1990) was not. For speech obtained from a speaker with the opposite-sex of the target speech, Festen and Plomp investigated the effect of time-reversal. Surprisingly, they found no difference in SRT between time-reversed speech and regular speech. Additionally, when the interfering time-reversed speech and the target speech were obtained from the same speaker, they found even worse speech intelligibility than for stationary noise with the same average long-term frequency content. Unfortunately they did not investigate this for regular speech or for many-channel speech modulated noise. Attempting to estimate the effect of time-reversal of background noise on speech intelligibility is difficult. Although the long-term frequency content is the same, regular and time-reversed fluctuating noise clearly differ in both temporal and informational content. This might lead to differences in speech masking. Although the experiment of Festen and Plomp seems to indicate that time-reversal of speech maskers does not increase the benefit of fluctuations in a background noise, many experiments suggest otherwise. It is generally assumed that the time-reversal of interfering speech causes an improved release of masking due to less informational masking. However, time-reversal might also lead to in increase in forward masking due to the asymmetrical shape of the temporal envelope of speech (Rhebergen et al., 2005). Summers and Molis (2004) found for moderately sensorineurally hearing impaired listeners (average $PTA_{0.5,1,2\,kHz} = 27$ dB) for a (single-talker) speech masker a (insignificant) benefit of 1.4 dB with respect to stationary noise. They used the same speaker for both target and masker. For a time-reversed masker the significant benefit was 3.3 dB, and was thus 1.9 dB larger than for regular speech. The average benefit of 2.8 dB that we measured for six-channel modulated time reversed speech, is somewhat less (0.5 dB) than found by Summers and Molis (2004) for time reversed real speech. This may be caused by the fact that their subjects were less hearing impaired, and that we used six-channel modulated noise instead of a single competing speaker. Our results lie between the results of Eisenberg et al. (1995), Festen and Plomp (1990) and Summers and Molis (2004).

Measurement error

In spite of the larger range in speech reception thresholds for fluctuating noise, fewer statistically significant results were obtained for this noise than for stationary noise. Nearly all p- and F-values (Tables 3.2 and 3.3) were smaller than the corresponding values obtained for stationary noise (Tables 2.5 and 2.6). This lack of significant differences could be caused by the larger measurement error for the speech reception thresholds in fluctuating noise. The average measurement error (both compression and linear amplification) was 1.9 dB, whereas it was 1.5 dB for stationary noise. Drullman and Bronkorst (2004) also found a larger measurement error for fluctuating noise than for stationary noise. They used the same speech materials (Versfeld et al. 2000), combined with fluctuating noise constructed from speech materials from Plomp and Mimpen (1979). For our data, the larger error might stem from two sources. First, for fluctuating noise the masking of specific words and syllables depended on the quasi-random occurrence of gaps in the noise. This typically increases the measurement error (Hagerman, 1997), since the masking of speech segments which attribute much to speech intelligibility can vary between sentences. Second, nearly all subjects mentioned that the listening task was much more demanding for fluctuating noise than for stationary noise due to the masking of arbitrary speech segments. Although the participants were allowed frequent breaks, this might have diminished subjects' attentiveness resulting in a larger measurement error.

The mean square of the effect, MS_{effect}, which can be seen as a measure of the variance explained by the variable under investigation, was for nearly all conditions smaller for fluctuating noise than the corresponding values for stationary noise[3]. This is supported by Figures 3.6 and 3.7 which show that, on average, compression parameters did not have a large influence on speech intelligibility in fluctuating noise. Thus the smaller number of significant differences originated in a larger measurement error (MS_{error}) for

[3]For this comparison, we used MS_{effect} of a univariate analysis. For multivariate analyses this also gives an acceptable estimate for MS_{effect}, with the exception of those effects for which Mauchley's test of sphericity was significant (T, without split-frequency compression ratios; NC * CR * T, with split-frequency compression ratios).

fluctuating noise, augmented by a smaller MS_{effect}.

For daily practice a larger measurement error in combination with a smaller effect on intelligibility implies that finding optimal parameter values based on speech intelligibility tests will prove to be more elaborate for fluctuating noise than for stationary noise. However, the (detrimental) effects of compression were smaller for fluctuating noise and this implies that it might be less critical to find optimal parameter values for this noise type.

3.4.2 Effect of compression on speech intelligibility

The effect of compression on speech intelligibility in fluctuating noise is somewhat disappointing. It was expected that fast-acting compression could increase speech intelligibility by lifting low level speech elements above the masking level. Although compression did not improve speech intelligibility in fluctuating noise, clear detrimental effects (as found in stationary noise for some compression parameters) were absent. For fluctuating noise, only the effect of number of channels was significant. Compression in six-channels yielded significantly worse speech reception thresholds than NC = 1 and NC = 2. This lower speech intelligibility might have originated in a reduction of spectral contrast caused by independent compression in six channels. Single- and two-channel compression resulted in SRTs near those obtained with linear amplification (ΔSRT = −0.1 dB for both NC = 1 and NC = 2).

For stationary noise we did not find a detrimental effect for a larger number of channels. There, NC = 1 yielded insignificantly worse results than NC = 2, and significantly worse results than NC = 6. These results were largely influenced by detrimental effects for NC = 1 with short T_a/T_r. It was speculated that fast-acting single-channel compression resulted in detrimental scores because of intermodulation distortion. In contrast to stationary noise, for fluctuating noise speech intelligibility for NC = 1 was not worse than for NC = 2 and NC = 6. Apparently the ability of fast-acting single-channel compression to lift the level of speech during gaps in the noise alleviates possible detrimental effects of spectral distortion.

Moore et al. (1999) found similar results for a single-talker background. For this noise type their experiment resulted in equal speech intelligibility for linear amplification and fast-acting compression with NC = 1 and NC = 2. For NC = 4 and NC = 8, speech intelligibility was significantly lower. Additionally, they used speech-shaped noise which had the (single-channel) modulations of a single-talker. In contrast to the single-talker background, for this fluctuating noise no significant difference was found between results of fast-acting compression (NC = 1, 2, 4, and 8). Van Buuren et al. (1999) also found comparable results for a single-talker background. For CR = 2 no significant differences were found. However, speech intelligibility for fast-acting compression with NC = 1 was slightly better than for linear amplification. Results for NC = 4, and NC = 6 were worse than for both NC = 1 and linear amplification. For a higher compression ratio (CR = 4), results for single-channel compression did not differ significantly from results obtained with linear amplification. However, for NC = 4 and NC = 16, results were significantly worse.

3.4.3 Best parameter values

The overall best result (ΔSRT = −0.9 dB) was achieved with single-channel compression, *CP(1, 2/2, 40/40)*. The second best score was only 0.04 dB worse, and was obtained with a CR of 3/3, *CP(1, 3/3, 40/40)*. A difference of ΔSRT = −0.9 dB may seem small. However, the psychometric function has shown that this corresponds to an increase in sentence score of 15%.

For two-channel compression the best result (ΔSRT = −0.7 dB) was obtained for *CP(2, 2/3, 40/40)*. For stationary noise this setting led to the best overall score (also ΔSRT = −0.7 dB). Since the effect of compression on speech intelligibility in fluctuating noise was rather limited, and since this setting resulted in good scores for both fluctuating and stationary noise, one might consider using these parameter values. For six-channel compression, the average results were worse than for single- and two-channel compression. The best result for six-channel compression was obtained for *CP(6, 1/2, 4/40)* and was ΔSRT = −0.5 dB.

3.5 Conclusions

Compression in fluctuating noise did not significantly improve or degrade speech intelligibility of moderately hearing-impaired subjects. Although multi-channel compression is known to reduce spectral and temporal contrasts, the effect on speech intelligibility in fluctuating noise was limited. This is probably caused by improved audibility of otherwise masked weak speech elements.

Only one significant result was obtained: six-channel compression led to significantly worse speech intelligibility than linear amplification. Contrary to stationary noise, fast-acting compression did not result in detrimental speech intelligibility, and interaction effects were not significant.

The fact that hardly any significant effect of compression on speech intelligibility was found implies that using speech intelligibility in fluctuating noise for finding the best setting for an individual can be very tedious. Parameter values which resulted in the best speech intelligibility for stationary noise yielded good results in fluctuating noise as well. For two-channel compression, *CP(2, 2/3, 40/40)* resulted in best speech intelligibility for both stationary and fluctuating noise. For six-channel compression *CP(6, 2/3, 40/400)* resulted in the best score for stationary noise and in a good score for fluctuating noise (ΔSRT = −0.4 dB).

Results for individual listeners 4

Rolph Houben and Guido F. Smoorenburg
Submitted for publication to Audiology & Neurotology

4.1 Introduction

The previous chapters addressed results averaged over twenty hearing-impaired subjects. The resulting group effects are important when comparing the efficacy of different processing schemes and for the formulation of generic fitting strategies. In clinical practice, hearing aids are often fitted based on these generic fitting rules; current examples are NAL-NL1, FIG6, DSL [i/o], and IHAFF (Dillon, 2001). These fitting rules, based on group averages, have the disadvantage that they may not cover inter-individual variation around the average value, which can be relatively large. For instance, Crain and Yund (1995) showed that the effect of compression parameters on speech intelligibility can vary substantially between individual listeners.

It is currently unknown which compression parameters (e.g., number of channels, compression ratio, time constants) yield best speech intelligibility for an individual. Moreover, it is unknown how to predict which individuals might benefit from compression over linear amplification, for speech at a comfortable listening level.

This chapter examines the effect of number of channels, compression ratio, and time constants on speech intelligibility of individual subjects. We used the experimental data of the previous chapters to re-analyse the results for individual listeners.

4.1.1 Previous research reporting individual results of compression

Although most research on compression has been concerned with group results, several previous papers describe individual results. Table 4.1 lists the papers of which the findings for individual subjects are described here.

For clarity, we have coded the compression characteristics. The nomenclature is identical to that of the previous chapters (see section 2.1.1 and 3.1.1). Again, three parameters are given, namely number of channels (NC), compression ratio (CR) and attack and release time ($T = T_a/T_r$). These three compression parameters (CP) are indicated by $CP(NC, CR, T_a/T_r)$. Linear amplification is indicated by $CR = 1$. A question mark indicates that the parameter values were not specified by the authors. Note that in contrast to the previous chapters, we do not limit this review to research for a specific noise type (e.g., stationary or fluctuating noise) or for a specified degree of hearing-loss (e.g., moderately hearing-impaired).

As early as 1978, Villchur propagated the presentation of results for individual listeners (Villchur, 1978). He ascribed the lack of clear results of compression experiments in part to the common practice of averaging results over subjects. According to him, one should determine which compression conditions give beneficial or degraded results for an individual, rather than averaging seemingly inconsistent results over several listeners. Despite Villchur's remark, most research on compression still presents only group results.

Table 4.1: *Overview of previous research that presents results for individual listeners. Q = quiet, S = stationary noise, and F = fluctuating noise. PTA and DR indicate that they used pure-tone audiometric threshold or dynamic-range based measures.*

Paper	Compression Parameters $CP(NC, CR, T_a/T_r)$	Number of subjects	$PTA_{0.5,1,2\,kHz}$ (dB HL) average or range	Background Noise	Tested Quality * = significant
Positive effect of compression *decreased* with hearing loss, or					
Negative effect of compression *increased* with hearing loss:					
Verschuure et al. 1993	*CP(1, 1 + 2 + 4 + 8, 10/22)*	18	15–>50	Q	syllable correct score
Yund and Buckles 1995a	*CP(8, 1 + 1-7, 4/4)*	16	20–53	S	PTA
Olsen et al. 2004	*CP(3, 1 + 2/5 + 5/10, 5-60/11-60 + 100-135/300-400)*	20	16–54	F	PTA* + SRT_{stat}* + release of masking
Verschuure et al. 1998	*CP(2, 1 + 2, 5/15)*	10	47–72	S	PTA*
Negative effect of compression *decreased* with hearing loss (positive for large loss):					
Moore et al. 1992	*CP(2, 1-3, 1/50-100)*	20	28–62	F	DR*
Stone et al. 1997	*CP(2, 1 + 2, 2/20)*	12	33–67	S	DR*
Effect of compression not dependent on hearing loss:					
Stone et al. 1997	*CP(2, 1 + 2, 2/20)*	12	33–67	F	DR
Nábělek 1983	*CP(3, 1 + ?, ?/?)*	10	25–73	S	DR + PTA
	CP(1, 1 + 2.5, 1/10+1/30+3/30+3/90+10/90+42/370)	9	25–73	S	speech scores + DR + PTA
	CP(1, 2.5 + 5 + 10, 1/30)	9	25–73	S	speech scores + DR + PTA
Neuman et al. 1994	*CP(1, 1 + 1.5 + 2 + 3 + 5 + 10, 5/200)*	10/10	32–85/42–70	S + F	DR_{speech}
Neuman et al. 1995	*CP(1, 1.5 + 2 + 3, 5/60 + 5/200 + 5/1000)*	20	37–85	S + F	DR_{speech} + PTA + age
Crain and Yund 1995	*CP(2 + 4 + 8 + 16 + 31, 1-9, 4/4)*	9	42–68	Q	-
Drullman and Smoorenburg 1997	*CP(6, 1 + 2 + 3 + 5, 0/0)*	16	90–115	Q	DR
Humes et al. 1997	*CP(1, 1 + 1∞, ?/?)*	41/53/16	33/44/62	F	unaided speech score
Dillon et al. 1998	*CP(1, 2, ?)*	140	8–75		PTA + age
Verschuure et al. 1998	*CP(2, 1 + 2, 5/15)*	10	12–42	F + S	PTA
	CP(2, 1 + 2, 5/15)	10	47–72	F	PTA
Moore et al. 1999	*CP(1 + 2 + 4 + 8, 1 + 1-2.9, 7/7)*	18	19–53	S + F	PTA
Franck et al. 1999	*CP(1, 2, 5/15)* and *CP(8, 2, 0.5/10)*	8	26–78	Q + S + F	PTA + DR + slope of auditory filter*
Goedegebure et al. 2001	*CP(1, 1 + 2 + 4, 5/5)*	14	13–60	Q + F	PTA + DR + UCL + loudness growth
Souza and Kitch 2001	*CP(1 + 2 + 3 + 4, 1 + 2 + 5, 3/25)*	16	43	Q	age
Van Toor and Verschuure 2002	*CP(4, ?, 2/16-2/64 + 2-16/64-512 + 32-64/1024-2048)*	38	18–70	S + F	PTA
Shanks et al. 2002	*CP(1, 1-2.7, 1/70)*	356	unknown	F	PTA (results due to increased audibility)

Nábělek (1983) *CP(3, 1 + ?, ?/?)* used ten subjects to compare compression to linear amplification. For a group of seven subjects with a small dynamic range (average $DR_{0.5,1,2\,kHz} = 13$ dB) he found better speech intelligibility for multi-channel compression than for linear amplification. For a group of three subjects with a larger dynamic range (average $DR_{0.5,1,2\,kHz} = 42$ dB) compression generally resulted in worse scores than linear amplification did. However, he found no significant correlation between benefit of compression and dynamic range ($DR_{0.5,1,2\,kHz}$ ranged from 7 to 49 dB). Note that stationary noise was added after compression (not simulating normal conditions where the combined signal of speech plus noise is compressed) and that the compression ratios were subject and frequency dependent. Nábělek also investigated the effect of single-channel compression with different time constants *CP(1, 1 + 2.5, 1/10 + 1/30 + 3/30 + 3/90 + 10/90 + 42/370)* and with different compression ratios *CP(1, 2.5 + 5 + 10, 1/30)* for nine subjects. (Again the noise was added after compression.) He performed a correlation analysis between the benefit of compression (scores with compression minus scores for linear amplification) and the speech intelligibility scores without compression (linear amplification), $PTA_{0.5,1,2\,kHz}$, the slope of PTA between $PTA_{1\,kHz}$ and $PTA_{4\,kHz}$, $DR_{1\,kHz}$, and $DR_{4\,kHz}$. He found no significant correlations.

Moore et al. (1992) *CP(2, 1-3, 1/50-100)* compared speech intelligibility of twenty moderately hearing-impaired listeners ($PTA_{0.5,1,2\,kHz} = 28$–62 dB HL) with compression to that with linear amplification. The experiments were performed in a twelve-talker babble. They found only weak correlations between the effect of compression and several audiometric parameters (for instance $DR_{0.5,1,2,4\,kHz}$ and various other combinations of DR for these frequencies; the slope of the loudness growth function at the same frequencies and at various combinations, and dynamic range of hearing for speech). The highest correlation ($r = 0.48$, $p < 0.05$) was found for $DR_{2,4\,kHz}$. This correlation showed that the largest benefit of compression was obtained for subjects with the smallest dynamic range. Their Figure 8 indicates that $DR_{2,4\,kHz} \approx 20$ dB corresponded to an improvement in speech reception threshold in noise of roughly 1.5 dB.

Verschuure et al. (1993) *CP(1, 1 + 2 + 4 + 8, 10/22)* measured speech intelligibility for 18 moderately to severely hearing-impaired listeners in quiet ($PTA_{0.5,1,2\,kHz}$ ranged from about 15 to more than 50 dB HL). Their system used a delay to reduce overshoot (see section 1.7.4). Subjects were divided into three groups based on their syllable correct score for linear amplification: above 85%, between 85 and 75%, and below 75%. The corresponding hearing losses (PTA) were not reported. Verschuure et al. found a trend in that the average benefit of compression was higher for the group of seven subjects who achieved a high syllable correct score ($>85\%$). For the other two groups compression tended to result in no difference or even in a deterioration of speech intelligibility.

Neuman et al. (1994) *CP(1, 1 + 1.5 + 2 + 3 + 5 + 10, 5/200)* did not find any significant correlations between the preferred compression ratio in several noise types and the dynamic range of hearing (DR_{speech}, defined as the uncomfortable level of broadband speech minus the speech reception threshold in quiet). Yet, for two low-level noises (ventilation and apartment noise), subjects with a small dynamic range of hearing (group of 10 subjects, $DR_{speech} < 30$ dB, $PTA_{0.5,1,2\,kHz} = 42$-70 dB HL) preferred compression over linear amplification more often than chance would predict. For cafeteria noise (their only noise with a presentation level that was always above the compression threshold) and for subjects with a dynamic range > 30 dB (10 subjects, $PTA_{0.5,1,2\,kHz} = 32$–85 dB HL), the preference of compression over linear amplification was at chance level for all three noise types. Only one subject out of 20 (DR_{speech} was unspecified) showed a significant preference for compression over linear amplification, and only for $CR = 1.5$. These findings led Neuman et al. to the conclusion that for slowly-acting compression, a compression ratio of about 1.5 to 2 is appropriate, especially for persons with a small dynamic range, and that it is not necessary to use an individual selection of compression ratios. Note that the available data suggests that for situations other than relative quiet, or for subjects with a $DR_{speech} > 30$ dB, linear amplification might be suitable as well.

In a subsequent experiment, Neuman et al. (1995) *CP(1, 1.5 + 2 + 3, 5/60 + 5/200 + 5/1000)* investigated the effect of release time on sound quality in both stationary and fluctuating noise. The dynamic range (DR_{speech}) of the twenty subjects spanned a range of 22 dB to 63 dB (and $PTA_{0.5,1,2\,kHz}$ ranged from 37 to 85 dB HL). The results showed several significant differences in the preference for release time. Some subjects consistently preferred long release times, while others preferred short times. For some, the preferred release time changed with compression ratio or noise type. However, again no significant correlations between the individual preferences and several audiometric characteristics (DR_{speech}, $PTA_{unspecified}$, shape of the audiogram, age) were observed.

Crain and Yund (1995) *CP(2 + 4 + 8 + 16 + 31, 1-9, 4/4)* conducted a regression analysis between consonant discrimination score in quiet and the number of compression channels. Their subjects had a $PTA_{0.5,1,2\,kHz}$ of 42 to 68 dB HL. Crain and Yund found a trend in that for about half of the nine subjects the scores improved with increasing number of channels. For the other half the scores decreased. However, these trends were statistically significant for two subjects only (one leading to an improvement of the scores with the number of channels, the other to a degradation).

Yund and Buckles (1995a) *CP(8, 1 + 1-7, 4/4)* compared results for compression to those obtained for linear amplification. The compression system was fitted according to the frequency-dependent dynamic range of their sixteen subjects ($PTA_{0.5,1,2\,kHz}$ ranged from 20 to 53 dB HL). Benefit of multi-channel compression in stationary noise decreased for increasing $PTA_{0.5,1,\,2,4\,kHz}$, and thus, implicitly with decreasing dynamic range. However, the correlation was insignificant ($r^2 = 0.16$, $p = 0.12$).

Drullman and Smoorenburg (1997) *CP(6, 1 + 2 + 3 + 5, 0/0)* measured the effect of compression on speech intelligibility of audiovisual stimuli in quiet for sixteen profoundly hearing-impaired subjects ($PTA_{0.5,1,2\,kHz}$ ranged from 90 to 115 dB HL, estimated from their Fig. 2). They found no significant correlations between dynamic range of hearing ($DR_{0.125,0.25,0.5\,kHz}$ and $DR_{1,2,4\,kHz}$) and benefit of compression for different compression ratios.

Humes et al. (1997) *CP(1, 1 + 1∞, ?/?)* used an amplification system in which the bass response was increased for low input levels (BILL). The experiments were conducted in both cafeteria noise and multi-talker babble. The subjects were divided into three groups with an average $PTA_{0.5,1,2}$ of 33, 44, and 62 dB HL, consisting of 41, 53, and 16 subjects, respectively (data was obtained from their Fig. 1). Group data showed only small differences between the effect of compression and linear amplification. No significant correlation was observed between effect of compression and the unaided scores.

Stone et al. (1997) used a two-channel system *CP(2, 1 + 2, 2/20)* of which the high frequency channel consisted of a fast-acting compressor, and the low frequency channel consisted of a linear amplifier. This system was preceded by a dual compression system that kept the overall signal level constant. The dual system consisted of a slowly-acting single-channel compressor ($T_a = 330$ ms, hold-off $= 560$ ms during which the gain was unchanged, $T_r = 1000$ ms) that was followed by a fast-acting single-channel compressor ($T_a = 3$ ms, $T_r = 80$ms). The $PTA_{0.5,1,2\,kHz}$ of the twelve subjects ranged between 33 and 67 dB HL, and the $DR_{2,4\,kHz}$ between 20 and 65 dB. For a single-talker background Stone et al. found no significant correlation between the effect of compression and $DR_{2,4\,kHz}$. However, for stationary noise the correlation was significant, although small ($r^2 = 0.15$, $p < 0.05$). The negative effect of compression tended to increase with dynamic range.

Dillon et al. (1998) *CP(1, 2, ?)* compared the preference of listeners for a low compression threshold (46 dB SPL) to that for a high threshold (56 dB SPL). Their 140 subjects had a $PTA_{0.5,1,2\,kHz}$ between 8 and 75 dB HL. Dillon et al. found no significant correlation between the preference scores and audiometric characteristics ($PTA_{0.5,1,2\,kHz}$, slope of hearing thresholds, age, years of hearing aid use).

Verschuure et al. (1998) *CP(2, 1 + 2, 5/15)* evaluated compression with overshoot reduction (identical to Verschuure et al., 1993) in four real-life background noises: restaurant noise, industrial noise, printing-office noise, and city background noise. They measured speech intelligibility of CVC words for two groups of subjects: moderately hearing impaired (10 subjects with $PTA_{0.5,1,2\,kHz}$ between 12 to 42 dB HL, estimated from their Fig. 6), and severely hearing impaired (10 subjects with $PTA_{0.5,1,2\,kHz}$ between 47 and 72 dB HL, estimated from their Fig. 8). The only significant correlation between the effect of compression and hearing threshold was found for the severely hearing-impaired subjects and industrial noise. For this group and noise type, there was a negative effect of compression that correlated with $PTA_{0.5,1,2\,kHz}$ ($r^2 = 0.56$). The industrial noise had a flat frequency spectrum and it was the most steady of the noises used (in Table 4.1 it is classified as stationary noise).

Moore et al. (1999) *CP(1 + 2 + 4 + 8, 1 + 1-2.9, 7/7)* compared the effect of the number of channels on speech intelligibility (SRT) in both stationary and fluctuating noise for 18 elderly subjects ($PTA_{0.5,1,2\,kHz} = 19$–53 dB HL). Their system used a delay to reduce overshoot. They found no significant correlation between effect of compression and threshold of hearing ($PTA_{0.5,1,2,4\,kHz}$, ranging from 25 to 58 dB HL; and $PTA_{2,3,4,5\,kHz}$, ranging from 42 to 76 dB) of the individual subjects. Moore et al. did not report correlation analyses between benefit of compression and dynamic range of hearing.

Franck et al. (1999) *CP(1, 2, 5/15)* and *CP(8, 2, 0.5/10)* applied compression on speech materials that had been spectrally enhanced. Both the single-channel and the multi-channel system used a delay to reduce overshoot. The experiment included eight young subjects with $PTA_{0.5,1,2\,kHz}$ ranging from 26 to 78 dB HL (estimated from their Fig. 1), and $DR_{0.5,1,2\,kHz}$ ranging from 22 to 51 dB. Franck et al. found no significant correlation between audiometric data (threshold at several frequencies, dynamic range at several frequencies, auditory filter width at 0.5 and 3 kHz) and the effect of compression on speech intelligibility of the spectrally enhanced materials. They also performed an analysis based on the slopes of the auditory filters (at 0.5 and 3 kHz), and found that the low-frequency skirt at 3 kHz was significantly related to the effect of single-channel compression. The benefit of compression was higher for subjects with a steep low-frequency filter slope at 3 kHz.

Goedegebure et al. (2001) *CP(1, 1 + 2 + 4, 5/5)* used a compression system in quiet and in restaurant noise. The system used overshoot reduction similar to the system of Verschuure et al. (1993). The fourteen subjects were moderately to severely hearing impaired, $PTA_{0.5,1,2\,kHz}$ ranged from 13 to 60 dB HL, estimated from their Fig. 5. Goedegebure et al. measured the effect of compression ratio on phoneme intelligibility. They investigated the relation between the effect of compression ratio and audiometric characteristics. In the analysis they used $PTA_{2,4\,kHz}$, low-high frequency contrast, slope of the thresholds, $DR_{2,4\,kHz}$, $UCL_{2,4\,kHz}$, and slope of the loudness growth function averaged over 2 and 4 kHz. All correlations were insignificant. Besides the above described compression system, they used a second system with the same compression parameters but in which the control signal was filtered to emphasize compression for the high frequencies. In quiet, listeners with a large dynamic range ($DR_{2,4\,kHz}$ between 30 to 50 dB) tended to have some benefit from compression, whereas listeners with a small dynamic range ($DR_{2,4\,kHz} < 25$ dB) only had worse phoneme intelligibility with compression. In the restaurant background noise, this trend was not present.

Souza and Kitch (2001) *CP(1 + 2 + 3 + 4, 1 + 2 + 5, 3/25)* used compression on a speech signal of which spectral information was removed (resulting in a flat spectrum) by randomly assigning a negative or positive sign to the waveform samples. This procedure leaves the envelope of the speech intact. The moderately hearing-impaired subjects were divided into two age groups. One group consisted of 7 young subjects (18 to 34 yr), the other of 9 aged subjects (71 to 95 yr). The thresholds of hearing for both age groups were roughly equal: the average $\mathrm{PTA}_{0.5,1,2\,\mathrm{kHz}}$ was 43 dB HL for both age groups (estimate obtained from their Fig. 1.). For all conditions (compression and linear amplification) the elderly had a significantly lower sentence recognition in quiet than the younger ones. All speech intelligibility scores decreased with increasing compression ratio, but the difference between the two age groups remained constant. This strongly suggested that the effect of compression ratio on speech intelligibility does not depend on age.

Van Toor and Verschuure (2002) *CP(4, ?, 2/16-2/64 + 2-16/64-512 + 32-64/1024-2048)* did not find any significant correlation between individually favoured time constants and the degree or steepness of the hearing loss of their subjects. The 38 subjects had a $\mathrm{PTA}_{0.5,1,2\,\mathrm{kHz}}$ between 18 and 70 dB HL (obtained from their Fig. 3). The experiments included both stationary and fluctuating noise.

Shanks et al. (2002) *CP(1, 1-2.7, 1/70)* compared the effect of single-channel compression on speech intelligibility to the effect of linear amplification. They analysed the obtained data separately for four groups of subjects with a different hearing loss. They used speech in fluctuating noise at three signal-to-noise ratios (−3, 0, and +3 dB) and at three levels (52, 62, and 74 dB SPL). Nearly all significant differences occurred for the lowest presentation level (52 dB SPL) and can be attributed to an increased audibility of speech due to more amplification for the compression condition than for the linear one. For the first group, consisting of 62 subjects with a large and sloping hearing loss ($\mathrm{PTA}_{0.5,1,2\,\mathrm{kHz}} > 40$ dB HL and $\mathrm{slope}_{0.5\text{–}4\,\mathrm{kHz}} > 10$ dB/octave), speech intelligibility was better for compression than for linear amplification.
For the second group, consisting of 54 subjects with a small and flat hearing loss ($\mathrm{PTA}_{0.5,1,2\,\mathrm{kHz}} < 40$ dB HL and $\mathrm{slope}_{0.5\text{–}4\,\mathrm{kHz}} < 10$ dB/octave), speech intelligibility was worse for compression than for linear amplification.
The third and fourth group consisted of 149 subjects with a small and sloping loss, and of 91 subjects with a large and flat loss, respectively. For these two groups, by and large no significant results were found.

Olsen et al. (2004) *CP(3, 1 + 2/5 + 5/10, 5-60/11-60 + 100-135/300-400)* investigated the benefit from four fast-acting compression systems in relation to audiometric factors. Two systems used $\mathrm{CR} = 2$ for the channels in which the hearing loss was less than 60 dB HL, and $\mathrm{CR} = 5$ for hearing loss exceeding 60 dB HL. The other two systems used $\mathrm{CR} = 5$, and $\mathrm{CR} = 10$, respectively. The experiments were conducted with twenty moderately hearing-impaired subjects ($\mathrm{PTA}_{0.5,1,2\,\mathrm{kHz}}$ ranged from 16 to 54 dB HL, obtained from their Fig. 1) in fully modulated noise (Hagerman, 2002). For linear amplification, they also measured speech intelligibility in slightly modulated noise: $SRT_{slight,\,linear}$ (Hagerman, 2002). In fully modulated noise, only about 30% of the listeners showed better speech intelligibility with fast-acting compression than with linear amplification. Correlation analysis showed that the negative effect of compression increased significantly for subjects with a worse $SRT_{slight,\,linear}$ ($r^2 = 0.4$, $p < 0.01$). Subjects with a good $SRT_{slight,\,linear}$ obtained benefit from compression (correlation suggested a benefit of 1 dB for subjects with a $SRT_{slight,\,linear}$ of about −8 dB). Olsen et al. also investigated the correlation between effect of compression and threshold data such as $PTA_{0.5,1,2\,kHz}$, $PTA_{2,3,4\,kHz}$, $PTA_{3,4,6\,kHz}$, slope of the thresholds: $PTA_{2,3,4\,kHz}$ minus threshold at 0.5 kHz. Of these measures, $PTA_{2,3,4\,kHz}$ and $PTA_{3,4,6\,kHz}$ correlated slightly ($r^2 < 0.26$) with the effect of compression. Correlation between the effect of compression and release of masking (i.e. the difference in score (SRT_{linear}) between fully and slightly modulated noise) showed no significant correlation. However, when only those subjects were selected that had improved speech intelligibility with compression (6 out of 20 subjects), the negative correlation between benefit of compression and release of masking became nearly significant ($r^2 = 0.18$, $p = 0.07$).

4.1.2 Conclusions based on previous research

Research on compression that reports significant results for individual subjects is relatively scarce, see Table 4.1.

Most research showed no significant correlation between the effect of compression and audiometric characteristics. Some papers describe a decrease of the positive effect of compression with hearing loss (Verschuure et al., 1993; Yund and Buckles, 1995a; Olsen et al. 2004) and some papers a decrease of the negative effect with hearing loss (Moore et al., 1992; Stone et al., 1997).

Neuman et al. (1994) found that subjects with a small dynamic range preferred compression ratios of 1.5 to 2 with slowly-acting compression. They concluded that individual selection of compression ratio was not necessary. Neuman et al. (1995) found that several subjects had a significant preference for a certain release time. However, they found no significant correlation between this preference and audiometric characteristics.

In conclusion, this extensive review shows that it is not clear which subjects would benefit most from compression. This however, does not negate the value of compression for hearing aids: the experiments were performed in situations in which the audibility of the stimuli was guaranteed for both compression and linear amplification. It is clear that compression can be very useful in situations where speech would not have been audible without compression (Dillon, 1996). Unfortunately, the present review shows that we can not predict which compression characteristics will ensure optimal speech intelligibility at comfortable input levels. In view of the inconclusive results from the literature quoted above we performed additional analyses on our data of which the average results were reported in the previous chapters.

4.2 Results

4.2.1 Learning effect

The previous analyses (Chapters 2 and 3) showed a small average learning effect for both stationary ($SRT_{retest} - SRT_{test} = -0.5$ dB, $p < 0.001$) and fluctuating noise ($SRT_{retest} - SRT_{test} = -0.6$ dB, $p < 0.0005$). In Chapters 2 and 3 a correction for this effect was applied based on regression analysis. An analysis for each subject individually showed that the learning effect was insignificant in all subjects. Moreover, the overall learning effect did not correlate significantly with hearing loss or age. Therefore, the correction from Chapters 2 and 3 was applied again, here.

4.2.2 Correlation of hearing loss with age

Hearing thresholds and dynamic range did not correlate significantly with age. We conducted an analysis for several combinations of pure tone average hearing loss (e.g., $PTA_{0.5,1,2\,kHz}$, $PTA_{1,2,4\,kHz}$, $PTA_{2,4\,kHz}$, $PTA_{2,4,8\,kHz}$, etc.) and dynamic range (e.g., $DR_{0.5,1,2\,kHz}$, $DR_{1,2,4\,kHz}$, $DR_{2,4\,kHz}$, $DR_{4\,kHz}$, $DR_{2,4,8\,kHz}$, $DR_{2,4,8\,kHz}$–$DR_{0.25,0.5,1\,kHz}$, etc.). The age of our subjects ranged from 41 to 83 year (Table 2.3). Speech intelligibility (50%-point) of consonant-vowel-consonant (CVC) words presented in quiet also showed no significant correlation with age. This finding was probably caused by the etiology of the subjects included. Only four subjects had a hearing loss that originated from aging, while the hearing loss of the other subjects was unknown or was ascribed to other factors (e.g., excessive noise exposure, heriditary hearing loss), see Table 2.2.1.

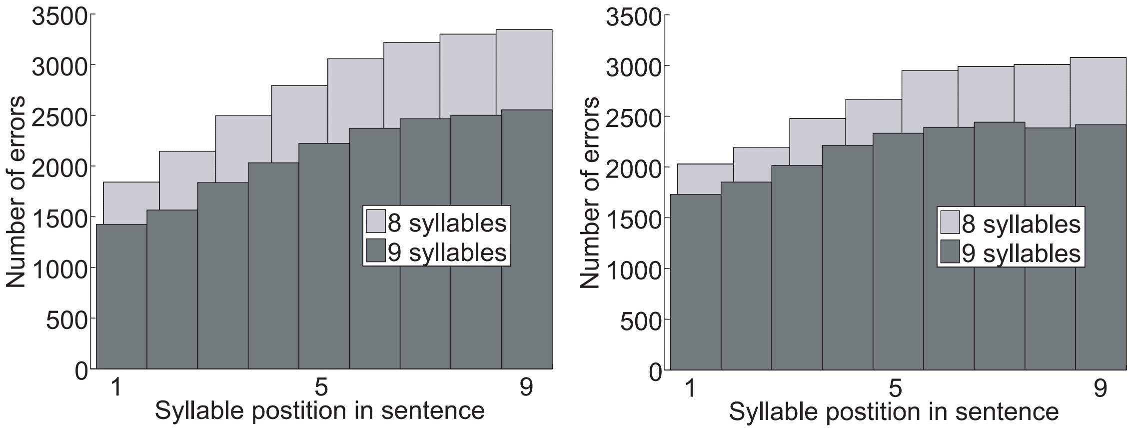
The speech reception threshold for both noise types ($SRT_{stat,\,linear}$ and $SRT_{fluct,\,linear}$) tended to increase with age. For stationary noise the correlation was insignificant ($r^2 = 0.1$, $p = 0.15$), and for fluctuating noise it was just significant ($r^2 = 0.2$, $p < 0.05$).

4.2.3 Response errors

The place in the sentence where most errors were made might reveal individual differences in listening strategy or it might reveal an effect of compression on speech intelligibility. Our experimental design allowed for the recording of the location of incorrectly reproduced syllables. Figure 4.1 shows the total number of errors for each location in a sentence, pooled over all experimental conditions and all subjects[1]. The

[1]Results for sentences consisting of 8 or 9 syllables are shown only. Together these two sentence lengths account for 98% of the measured errors. The remaining stimuli (consisting of sentences with 6, 7, or 10 syllables) all showed the same trend.

figure clearly shows that the number of errors increased towards the end of the sentences. This trend was very robust and was present for all but one subject (s10). Moreover, different conditions (including linear amplification) all resulted in a comparable increase in errors over the length of a sentence. This indicates that increase of errors towards the end of the sentences is probably not related to listening strategy or signal processing.



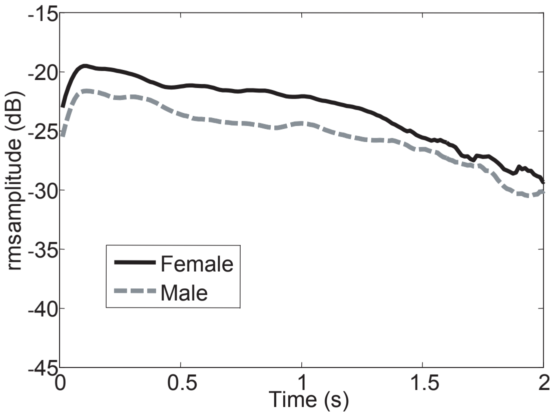
(a) *Stationary Noise* (b) *Fluctuating Noise*

Figure 4.1: *Error distribution within the stimuli. The bars give the number of errors (all subjects, all measurement conditions) which occurred at the corresponding syllable position within the sentence. Data is shown for sentences consisting of eight and nine syllables only.*

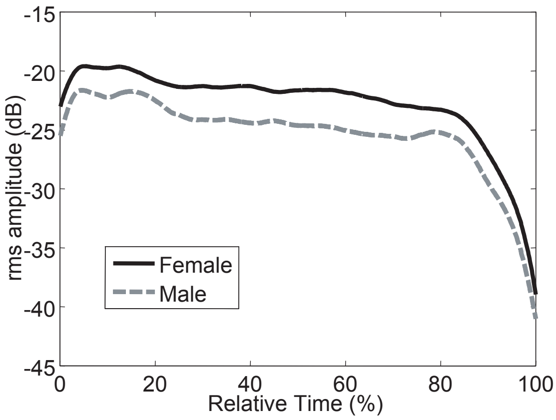
Figure 4.2(a) shows the average rms amplitude of the unprocessed speech materials as a function of time. The figure is analogous to Fig. 4 of Versfeld et al. (2000). It shows a large decrease in rms amplitude (dB) with time. The rms amplitude of the stationary noise (not shown in Fig. 4.2) was constant. For SRT measurements, this results in a decreasing signal-to-noise ratio during a sentence. This seems to explain the increase of errors towards the end of the stimuli[2]. In contrast to stationary noise, the amplitude of our fluctuating noise was not constant because the fluctuating noise was constructed from the concatenated envelopes of the original speech material (Chapter 3). However, the noise segments were always chosen at random, which ensured that the level variations were not synchronized with the speech.

In Figure 4.2(a) the average amplitude beyond 1.2 second is biased due to differences

[2]The adaptive measurement procedure only distinguished between entirely correct sentences and incorrect sentences. If a single mistake was made, the entire sentence was classified as incorrect. This implies that the last few syllables (which have the lowest signal-to-noise ratio) predominantly determined the score of the entire sentence. In other words, the sentence-based SRT shifted towards a word-based SRT. One might consider correcting this phenomenon by equalizing the signal level prior to adding noise. However, natural speech also shows a decline towards the end of the sentence. If an estimate of real-life speech intelligibility is required, this level decline might actually be desirable.



(a) *Original data*



(b) *Normalized sentence length*

Figure 4.2: *Mean rms amplitude of the unprocessed speech materials as a function of time. The left panel shows the original data, the right panel the data after the length of each sentence was normalized. All sentences were linearly scaled to obtain the same relative length. The lines represent the average for the female and the male speaker. For each speaker all 507 sentences were used. The rms was averaged over 125 ms.*

in sentence length (ranging from 1.2 to 3.0 s). Since we are specifically interested in the amplitude over the duration of the sentence itself, we have scaled the duration of the individual sentences. Figure 4.2(b) shows the normalized results. The beginning of each sentence is indicated by 0% and the end by 100%. The figure shows that the average rms level decreased steadily from the beginning until about 80% of the sentence length. (The total decline between 10-80% was about 3.7 dB and 3.1 dB for the female and male materials, respectively.) At about 80% (i.e., during the last syllable) the amplitude dropped sharply.

The increase in error over the length of the sentence followed the steady decline in speech amplitude. However, the sharp sharp drop in amplitude at 80% of the sentence length does not correspond to an increased error for the last syllable. This last syllable was probably relatively easy to predict, which might have compensated its lower amplitude.

4.2.4 Measurement error

Measurement error (i.e., the square root of the average variance of the repeated measurements, after correction for the learning effect) was 1.5 dB and 1.9 dB for stationary and fluctuating noise, respectively (Chapters 2 and 3).

The measurement error tended to increase slightly with age (0.2 dB per 10 yr), however this was insignificant ($r^2 = 0.1$; $p > 0.1$, for both noise types). The error showed no significant correlation with PTA (e.g., $PTA_{0.5,1,2\,kHz}$, $PTA_{1,2,4\,kHz}$, etc.) and dynamic range (e.g., $DR_{0.5,1,2\,kHz}$, $DR_{1,2,4\,kHz}$, etc.). However, further analysis showed that the measurement error was significantly larger for subjects with worse SRT scores. Figure 4.3 displays the measurement error averaged over all compression conditions for each subject as a function of the subject's SRT for stationary noise and linear amplification ($SRT_{stat,\,linear}$). For comparison, an additional average value for normal-hearing subjects is also shown (Versfeld et al., 2000). Correlation of measurement error with $SRT_{stat,\,linear}$ was significant for both stationary ($r^2 = 0.4$, $p < 0.005$) and fluctuating noise ($r^2 = 0.3$, $p < 0.01$). Correlation of measurement error for fluctuating noise, against $SRT_{fluct,\,linear}$ was weaker ($r^2 = 0.20$, $p = 0.05$) and is not shown in the figure.

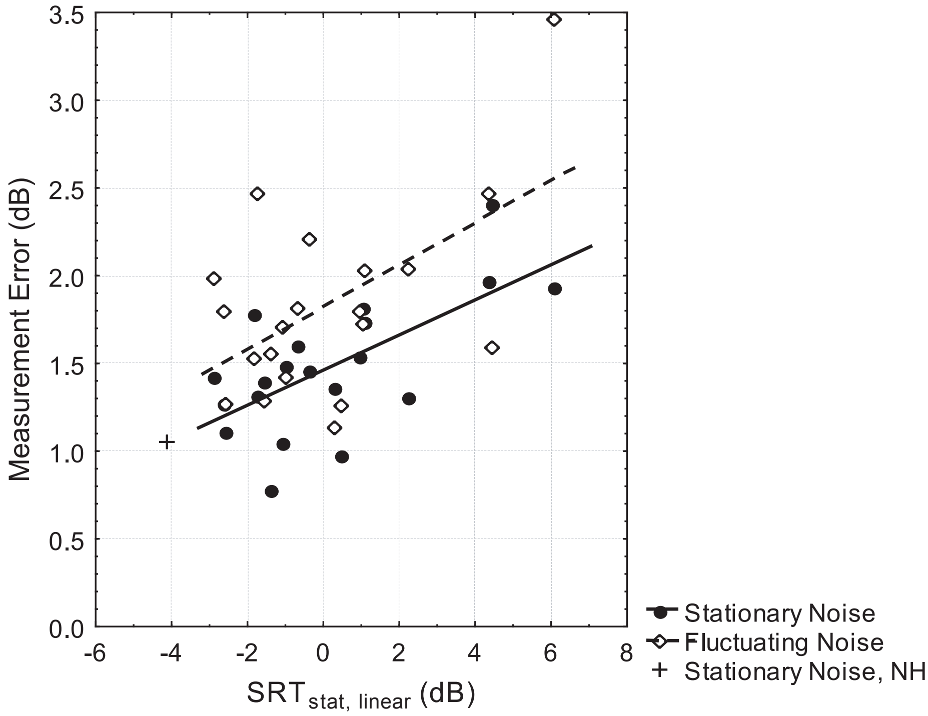


Figure 4.3: *Measurement error for stationary and fluctuating noise. The abscissa represents speech reception thresholds for stationary noise and linear amplification ($SRT_{stat,\,linear}$). Linear regression yielded significant correlations for both stationary ($r^2 = 0.4$, $p < 0.005$) and fluctuating noise ($r^2 < 0.3$, $p < 0.01$). The slopes of the regression lines are 0.10 dB/dB and 0.12 dB/dB for stationary and fluctuating noise, respectively. Additional data for normal hearing subjects from Versfeld et al. (2000) is represented by the + symbol. This data point is not included in our regression analysis and represents average data of 12 subjects for stationary noise and linear amplification.*

4.2.5 Analysis of Variance

Analogous to the previous chapters, the SRTs measured with compression will be presented relative to the subject's linear score. For this purpose, ΔSRT was defined as the SRT for a particular condition minus the SRT for the linear condition found for that subject (see section 2.3). Negative values of ΔSRT indicate a better speech reception threshold for compression than obtained for linear amplification.

A five-way analysis of variance (ANOVA) was conducted on the data of each of the twenty subjects. Prior to analysis, all the duplicate pairs were averaged and the average values were treated as a single data point. In order to conduct a full factorial analysis, the variance of the duplicate pairs (i.e., the square of the measurement error) was used as estimate of the error term of the ANOVA. An estimate was made for each subject separately[3].

Analogous to the analyses in Chapters 2 and 3, the ANOVAs were conducted twice: first with the inclusion of single-channel compression (and without split-frequency compression ratios) and second with split-frequency compression ratios (and without single-channel compression). The error variance for each subject was obtained from 13 duplicate measurements only, therefore we lowered the degrees of freedom of the error term accordingly[4], that is, from 30 (no split-frequency CRs) or 40 (with split-frequency CRs) to 12. Tables 4.2 and 4.3 contain the results for stationary noise, and Tables 4.4 and 4.5 the results for fluctuating noise. The tables show all p-values that were significant at the 5% level (which was the highest significance level that was found).

Stationary noise

Only two subjects (s7 and s18) had significant effects for interaction terms. For s7 these significant results were caused by two points with a very high ΔSRT, first *CP(6, 2/3, 4/40)* with ΔSRT = +9.4 dB, and second *CP(1, 3/3, 4/40)* with ΔSRT = +6.3 dB. Most likely these results were caused by a momentary lack of concentration. For s18, the significant result for CR * T was caused by consistently low values of ΔSRT for CR = 3/3 with $T_a/T_r = 4/40$. This effect was present for all number of channels (1, 2, and 6), but it was most prominent for NC = 1 ($\Delta SRT_{NC=1} = -1.8$ dB). This explains why the effect of CR * T was significant in the analysis with NC = 1 (Table 4.2) and not in the analysis without NC = 1 (Table 4.3).

Only six significant main effects were found in five subjects (s7, s10, s13, s14, and s18). For subject s7 the significance was again caused by the two points with a very high ΔSRT that were ascribed to lack of concentration. First, no significant effect was found for the number of channels. Second, for compression ratio, all five subjects had a significant effect. For s10, s13, and s14, a compression ratio of 3/3 resulted in significantly worse speech intelligibility than a compression ratio of 2/2. For s18, it was oppo-

[3]Although the measurement error did depend on subject, it was independent of compression parameters. For an individual it is therefore valid to use the error pooled over all compression conditions for that subject.

[4]Actually, the measured test-retest data is now used twice. The duplicate pairs were previously averaged and used as a single data point. Statistics suggest using the test (or retest) data only for this analysis, but in order to remain consistent to previous analyses we again used the average of the duplicate pairs.

Table 4.2: *Stationary noise; significant results of an analysis of variance for each subject separately. Single-channel compression was included and therefore split-frequency compression ratios (CR = 1/2, and 2/3) were omitted. The independent (between) factors were number of channels (NC = 1, 2, 6), compression ratio (CR = CR_{low}/CR_{high} = 2/2, 3/3) and time constants (T = T_a/T_r = 4/4, 4/40, 40/40, 4/400, 40/400).*

	Stationary noise						
	NC = 1, 2, 6			CR = 2/2, 3/3			
	p-value						
Subject	NC	CR	T	NC * CR	NC * T	CR * T	NC * CR * T
s1							
s2							
s3							
s4							
s5							
s6							
s7		< 0.05			< 0.05		
s8							
s9							
s10		< 0.05					
s11							
s12							
s13		< 0.05					
s14		< 0.05					
s15							
s16							
s17							
s18			< 0.05			< 0.05	
s19							
s20							

site: CR = 3/3 resulted in better speech intelligibility than CR = 2/2. Third, for time constants, only subject s18 (and s7) had a significant effect. For s18, both 40/400 and 40/40 resulted in better intelligibility than T_a/T_r=4/4 and 4/40.

Overall, only few significant effects were found: 7 out of 140 conditions (5%) for Table 4.2 and 5 out of 140 (4%) for Table 4.3. This small number of statistical results is at chance level (although one could argue about the significant results for CR, see Table 4.2). However, it is important to further evaluate these results for individual subjects, because in daily-practise hearing-aid fitting has to target individual hearing-impaired clients as well.

A comparison of the individual data to the group results of Chapter 2 shows that many subjects display similar (insignificant) trends as those seen for the statistically significant group results. First, number of channels; the group data showed significantly better speech intelligibility for NC = 6 than for NC = 1. The same trend was found for fifteen subjects (the five exceptions were s1, s4, s8, s14, s20), however it was insignifi-

Table 4.3: *Stationary noise; significant results of an analysis of variance for each subject separately. Split-frequency compression ratios (CR = 1/2, and 2/3) were included and single-channel compression was omitted. The independent (between) factors were number of channels (NC = 2, 6), compression ratio (CR = CR_{low}/CR_{high} = 1/2, 2/2, 2/3, 3/3) and time constants (T = T_a/T_r = 4/4, 4/40, 40/40, 4/400, 40/400).*

Stationary noise NC = 2, 6 CR = 1/2, 2/2, 2/3, 3/3							
	p-value						
Subject	NC	CR	T	NC * CR	NC * T	CR * T	NC * CR * T
s1							
s2							
s3							
s4							
s5							
s6							
s7			< 0.05	< 0.05	< 0.05	< 0.05	
s8							
s9							
s10							
s11							
s12							
s13							
s14							
s15							
s16							
s17							
s18		< 0.05					
s19							
s20							

cant for all. Second, compression ratio; in Chapter 2 a compression ratio of 3/3 resulted in the worst average speech intelligibility of all four compression ratios. All but two subjects (s12 and s18) displayed a similar (insignificant) trend. Third, T_a/T_r; for the group data, larger time constants resulted in better speech intelligibility. This effect was present (but insignificant) for eighteen subjects. The two remaining subjects (s2 and s7) achieved best results for $T_a/T_r = 4/4$. The group result for NC * T (Fig. 2.4) indicated that the worst speech intelligibility was obtained for fast-acting single-channel compression. This same effect was insignificantly present in the data of 10 subjects.

Fluctuating noise

The ANOVA results for fluctuating noise are shown in Tables 4.4 and 4.5. For fluctuating noise, no significant interaction effects were found. For the main effects, four subjects (s8, s13, s17, s20) had significant results. First, number of channels; only s13 showed a significant effect. In accordance with the group result, for this subject NC=2

Table 4.4: *Fluctuating noise; significant results of an analysis of variance for each subject separately. Single-channel compression was included and therefore split-frequency compression ratios were omitted. T represents the time constants ($T = T_a/T_r$). (NC = 1, 2, 6; CR = CR_{low}/CR_{high} = 2/2, 3/3; T_a/T_r = 4/4, 4/40, 40/40, 4/400, 40/400).*

Fluctuating noise NC = 2, 6 CR = 1/2, 2/2, 2/3, 3/3							
	p-value						
Subject	NC	CR	T	NC * CR	NC * T	CR * T	NC * CR * T
s1							
s2							
s3							
s4							
s5							
s6							
s7							
s8			< 0.05				
s9							
s10							
s11							
s12							
s13			< 0.05				
s14							
s15							
s16							
s17		< 0.05					
s18							
s19							
s20							

gave better speech intelligibility than than NC=6. Second, compression ratio; two subjects had significant effects: s20 for which CR=3/3 resulted in worse scores than CR=2/2, and s17 for which the results were reversed (CR=3/3 better than 2/2). Third, time constants; s8 had better results with slowly-acting compression ($T_a/T_r = 40/40$, 4/400, 40/400) than with fast-acting compression ($T_a/T_r = 4/4$, and 4/40). For subject s13, speech intelligibility for $T_a/T_r = 40/400$ was better than for $T_a/T_r = 40/40$.

The group data of Chapter 3 only yielded a significant result for the main effect of number of channels. On average, six-channel compression resulted in worse speech intelligibility than both single- and two-channel compression. For eleven of the twenty subjects this same trend was found in the individual results, but it was only significant for subject s13.

4.2.6 Significant results

The increase of measurement error for subjects with a worse SRT_{linear} suggests that the analysis for these subjects had a lower statistical power. This might have contributed

Table 4.5: *Fluctuating noise; significant results of an analysis of variance for each subject separately. Split-frequency compression ratios were included and single-channel compression was omitted. T represents the time constants ($T = T_a/T_r$). ($NC = 2, 6$; $CR = CR_{low}/CR_{high} = 1/2, 2/2, 2/3, 3/3$; $T_a/T_r = 4/4, 4/40, 40/40, 4/400, 40/400$).*

	Fluctuating noise						
	NC = 1, 2, 6			CR = 2/2, 3/3			
	p-value						
Subject	NC	CR	T	NC * CR	NC * T	CR * T	NC * CR * T
s1							
s2							
s3							
s4							
s5							
s6							
s7							
s8							
s9							
s10							
s11							
s12							
s13	< 0.05		< 0.05				
s14							
s15							
s16							
s17							
s18							
s19							
s20		< 0.05					

to the small number of statistically significant results. Figure 4.4 shows the number of significant post-hoc contrasts versus $SRT_{stat, linear}$. We used Tukey's Honestly Significantly Difference test on the highest level (NC * CR * T with adjusted degrees of freedom) for all compression conditions. The figure illustrates that compression led to few statistically significant results. However, no clear trend in the number of statistically significant results is visible. The results for subject s7 are biased due to two measurements with a high ΔSRT (see section 4.2.5).

4.2.7 Regression analysis

We analysed the effect of compression (ΔSRT) as a function of audiometric characteristics. In the analysis all data for a subject were pooled, except the parameter (NC, CR, or T) under investigation. No significant results were found for correlations of ΔSRT with pure-tone thresholds or dynamic range. However, for stationary noise significant results were obtained for the correlation between ΔSRT and $SRT_{stat, linear}$ and $CR = 1/2$ and $CR = 2/2$ (both, $r^2 = 0.3$; $p = 0.01$). Figure 4.5 shows these correlations.

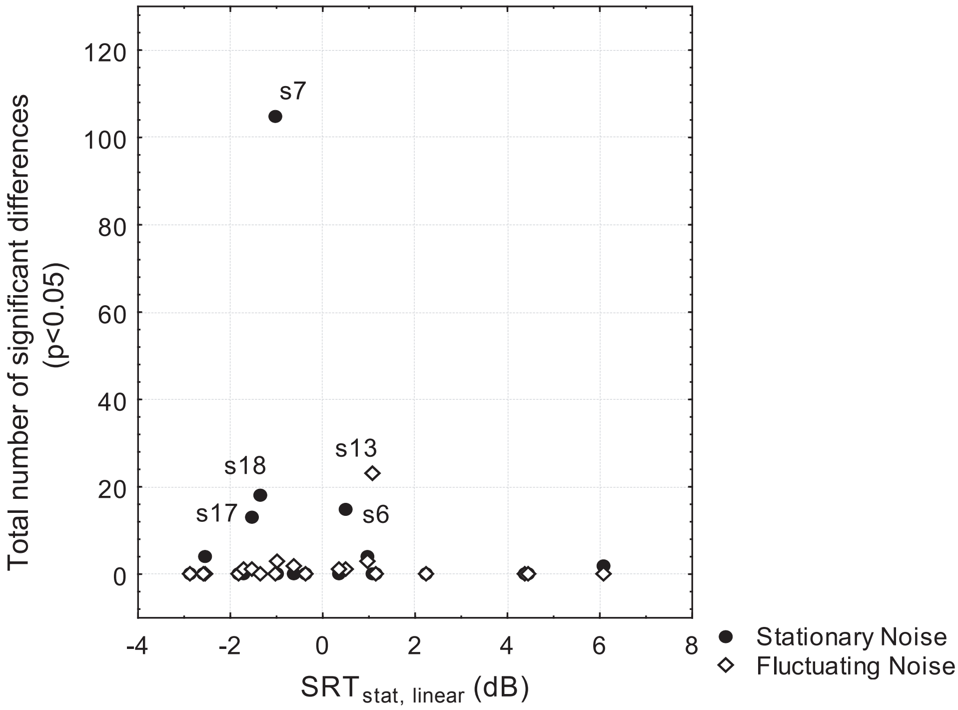


Figure 4.4: *Total number of statistical significant results based on Tukey's Honestly Significant Difference post-hoc test (with adjusted degrees of freedom). The data represents all 50 compression conditions, including conditions with NC = 1, and split-frequency compression ratios 1/2 and 2/3. Data for subjects with more than 10 significant results is marked with the subject number.*

No significant correlation was present for the higher compression ratios (CR = 2/3 and 3/3, $r^2 < 0.07$ and $p > 0.3$, respectively) as was the case with fluctuating noise for all compression ratios ($r^2 < 0.02$, $p > 0.5$).

4.2.8 Best result

In an attempt to find a predictor for the benefit of compression, we looked at the best (lowest) ΔSRT for each subject. Figure 4.6 shows ΔSRT_{best} for each subject as a function of the subject's score for linear amplification and stationary noise. Each point represents a single measurement (or the average of a duplicate pair). Only results for stationary noise are shown; error bars give the measurement error for each subject. The correlation between ΔSRT_{best} and $SRT_{stat, linear}$ was significant ($r^2 = 0.53$; $p < 0.0005$). For comparison, the figure also contains a regression line for the mean result of all 50 compression conditions ($r^2 = 0.16$; $p = 0.07$). Of course, a large part of the significant relation between ΔSRT and $SRT_{stat, linear}$ was caused by the increase of the measurement error for worse $SRT_{stat, linear}$. However, ΔSRT_{best} improved with 0.3 dB for every 1 dB increase in $SRT_{stat, linear}$, while the measurement error increased with 0.1 dB/dB only. Besides the best result, the average result of all compression conditions (ΔSRT_{all}) also

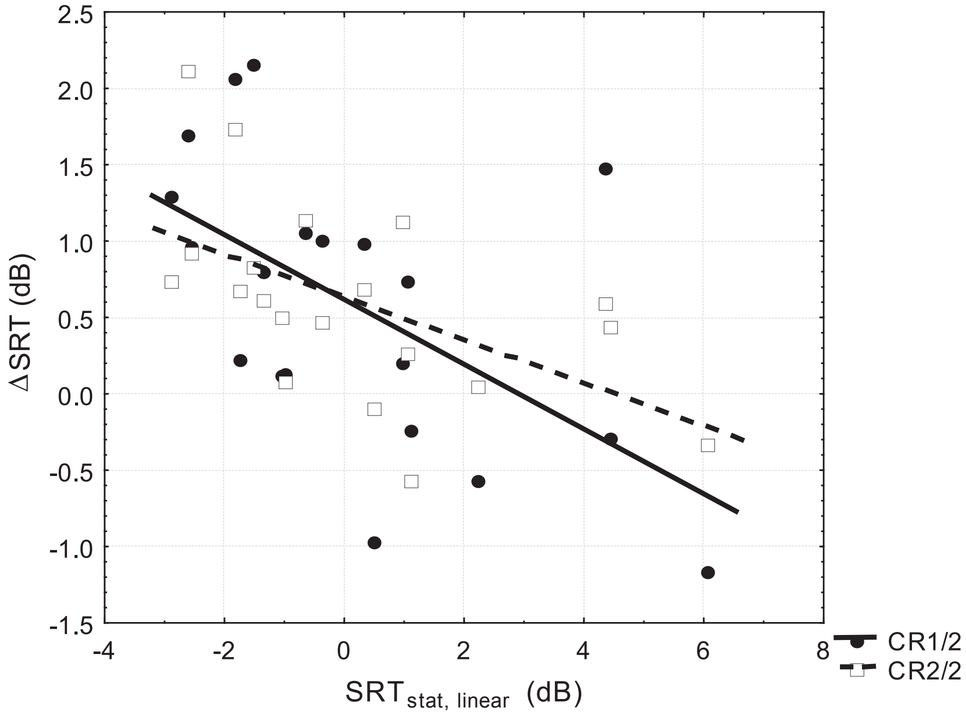


Figure 4.5: *Stationary noise; ΔSRT for CR = 1/2 and CR = 2/2 as a function of $SRT_{stat, linear}$. ΔSRT is defined as $SRT_{compression}$−SRT_{linear}. Lower values of ΔSRT indicate better results, and negative values represent a better SRT than obtained for linear amplification for that subject. Linear regression yielded a significant correlation ($r^2 = 0.3$; $p = 0.01$, for both CR = 1/2 and CR = 2/2).*

improved with $SRT_{stat, linear}$ (−0.12 dB/dB), see Fig. 4.6. The fact that both ΔSRT_{best} and ΔSRT_{all} decreased for increasing $SRT_{stat, linear}$ suggests that the improvement of ΔSRT with $SRT_{stat, linear}$ was not caused by the measurement error alone. For fluctuating noise (not shown in the figure) no significant correlations were found and the average result (ΔSRT_{all}) did not improve with increasing SRT_{linear}. These results suggest that the highest benefit of compression in stationary noise occurred for subjects with the worst $SRT_{stat,linear}$.

We also studied the relationship between ΔSRT_{best} and audiometric characteristics. Included were various combinations of PTAs, pure tone uncomfortably levels, and pure tone dynamic ranges. Additionally, we used supra-threshold measures such as CVC-scores in quiet and release of masking (expressed as SRT_{fluct}–SRT_{stat}). None of the correlations was significant.

In an attempt to find a predictor for the compression parameter setting (CP) that led to the best ΔSRT for an individual, we calculated the correlations between the setting (NC, CR, T) which resulted in ΔSRT_{best} and audiometric parameters. Al correlations were insignificant. Furthermore, we found no significant relationship between the setting for ΔSRT_{best} and setting of the subject's own hearing aid. Additional analyses in

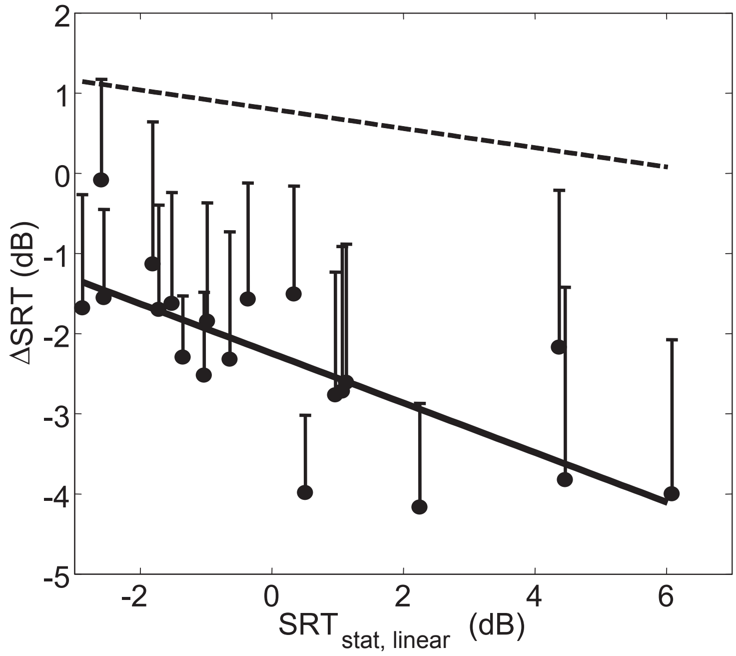


Figure 4.6: *ΔSRT_{best} for stationary noise versus SRT_{linear}. The data points show the lowest (best) ΔSRT for each subject for stationary noise only. Lower values of ΔSRT_{best} indicate better results. The error bars give the measurement error in each subject. Linear regression based on the data points only showed a significant correlation ($r^2 = 0.53$; $p < 0.0005$). The regression line showed an improvement in ΔSRT of -0.31 dB for every 1 dB increase in $SRT_{stat, linear}$. The striped line shows the insignificant regression line for the mean of all 50 compression conditions (-0.12 dB/dB).*

which the subjects were split into two groups, a group with a sloping hearing loss and a group with a flat hearing loss, again yielded no significant results.

4.3 Discussion

The ANOVAs for individual subjects yielded disappointingly few significant results. However, the main effects of Chapters 2 and 3 were clearly (but insignificantly) present in the data for many subjects. Interaction effects were not seen, except for the (insignificant) trend of a worse speech intelligibility for fast-acting single-channel compression in stationary noise (relative to slowly-acting single-channel compression or fast-acting multi-channel compression).

Correlation analyses between ΔSRT_{best} and pure-tone audiometric threshold, dynamic range, uncomfortable loudness levels, and CVC-scores in quiet, did not yield any significant result for this set of subjects. This lack of significant relationship between effect of compression and audiometric characteristics has also been described by several other studies (see Table 4.1). Correlation between ΔSRT and release of masking (i.e., SRT_{fluct}–SRT_{stat}) was also insignificant for both stationary and fluctuating noise. Olsen et al. (2004) also reported an insignificant correlation between benefit of compression in fluctuating noise and release of masking.

However, the correlation between ΔSRT_{best} and $SRT_{stat,\ linear}$ was significant. This was probably not caused by the age distribution of our set of subjects since $SRT_{fluct,\ linear}$ correlated only weakly with age, and the relationship between $SRT_{stat,\ linear}$ en age was insignificant. Moreover, the research of Souza and Kitch (2001) suggests that the effect of compression does not depend strongly on age.

The results for ΔSRT_{best} are most likely strongly influenced by measurement error. Still, the significant correlation was not entirely caused by measurement error, since the increase of measurement error with $SRT_{stat,\ linear}$ was three times as small as for ΔSRT_{best}. Moreover, the dependency of ΔSRT_{best} on $SRT_{stat,\ linear}$ was not present for fluctuating noise (for which measurement error increased equally). The correlation was also (insignificantly) present for the mean of all compression ratios (not for fluctuating noise). These other correlations suggest that the benefit of compression (ΔSRT_{best}) for stationary noise improved with increasing $SRT_{stat,\ linear}$.

For stationary noise, both $CR = 1/2$ and $CR = 2/2$ (pooled over all number of channels and time constants) had a significant correlation between ΔSRT and $SRT_{stat,\ linear}$ (Fig. 4.5). ΔSRT improved with increasing $SRT_{stat,\ linear}$. However, the measurement error increased with $SRT_{stat,\ linear}$ as well, resulting in few statistically significant differences (Fig. 4.4). The significant correlation for $CR = 1/2$ and $2/2$ with $SRT_{stat,\ linear}$ suggests that for the system with $CR = 1/2$ or $2/2$, subjects with a speech reception threshold between about +2 and +6 dB enjoyed more benefit than subjects with lower (better) SRTs. However, the lack of significant results for the main effect of compression ratio suggests that for these subjects $\Delta SRT_{CR=1/2}$ and $\Delta SRT_{CR=2/2}$ was not significantly better than results for higher compression ratios. These results do therefore not suggest to use low compression ratios.

The group data seems to represent the trends found for many individuals. The relatively large measurement error implies that the fine-tuning of a compression hearing aid based on the listener's speech intelligibility can be vary laborious. Rather than obtaining individual results for the fitting of compression in stationary noise, one might take the group results from Chapter 2. If individual results would be used, on might consider using main effects only, since the interaction effects seem less important.

Our results suggest that in order to predict the effect of compression in noise, one might look at speech intelligibility in noise, instead of pure-tone audiometric threshold or dynamic range.

4.4 Conclusions

This study compared the effect of compression parameters on speech intelligibility in stationary and fluctuating noise. The stimuli were presented entirely within the dynamic range of hearing of all subjects, and they were therefore audible without the need for compression. Thus, the effect of compression on speech intelligibility originated with the ability of compression to alter the dynamics of the stimuli, and not in the ability to increase audibility of the stimulus as a whole. This effect of compression could not be predicted with standard audiometric characteristics such as threshold of hearing or dynamic range. However, benefit of compression for stationary noise was shown to increase for subjects with a worse speech reception threshold ($SRT_{stat,\,linear}$). This suggests that a speech signal in noise can be more useful than pure-tone audiometric threshold or dynamic range data for determining a possible benefit of compression. Most current fitting rationales are threshold based, and prescribe a higher compression ratio for a larger hearing loss. While this might lead to substantial improvement of speech intelligibility at too low or too high input levels, it did not give improvement at the moderate levels used in this study. No significant correlation was found between dynamic range and the best compression ratio.

Overall, the effects of compression were rather small with respect to the within subject variability. For stationary noise, the benefit of compression tended to increase for subjects with a worse $SRT_{stat,\,linear}$. However, for both noise types, the measurement error was shown to increase with increasing SRT_{linear} as well. For a clinical setting, this indicates that the speech reception threshold is a laborious tool for assessing the best compression characteristics for an individual subject. Results indicate, however, that the SRT might be more suitable than pure-tone audiometric threshold or dynamic range. These results therefore agree with the results from the previous chapters: finding the best compression parameters based on optimizing speech intelligibility for an individual, is, at best, tedious.

5 Summary and clinical applicability

Many modern hearing aids apply compressive amplification to automatically increase the amplification of low-level sounds while keeping the high-level sounds at a comfortable loudness level. This amplification strategy can greatly improve user satisfaction. For instance, one of our subjects explained that with her new compression hearing aid she enjoyed her Sunday morning walks even more. Now she could hear birds sing without having to fumble with the volume knob or worry about too loud traffic sounds, simply because her hearing aid adjusted the volume automatically.

Compression can not only increase hearing aid comfort, but it can also influence speech intelligibility. There are many rationales for using compression to optimize speech intelligibility, and the number of different implementations of compression is even larger (Chapter 1). However, even the most advanced hearing aids suffer from the fact that eventually all processed sounds have to pass through the damaged ear. Although signal processing such as compression can counteract an impaired threshold and decreased dynamic range, it does not seem to be able to compensate for pathologically broad auditory filters (section 1.4.2). This inability is most evident for speech in a background noise, a condition in which hearing aids generally do not restore hearing to normal. Actually, for a noisy situation compression can even degrade speech intelligibility. Currently, the effect of various compression parameters on speech intelligibility in noise is unclear. Moreover, possible interactions between the various parameters are largely unknown.

In view of this hiatus, we conducted an extensive experiment in which we compared the effect of several compression parameters on speech intelligibility in a background noise. The experiment was a full parametric investigation, meaning that all combinations of the chosen parameter values were used. The parameters under investigation were number of channels (NC, 1, 2 or 6), compression ratio (CR_{low}/CR_{high} for frequencies below and above 1 kHz, respectively, ranging from 1/2 to 3/3), and attack/release time (T_a/T_r, ranging from 4 to 400 ms). The reference condition was amplification without compression (i.e., linear amplification). All subjects included had a moderate sensorineural hearing impairment.

Two types of background noise with a speech-shaped frequency spectrum were used: stationary and fluctuating noise. Stationary noise was chosen because it is well defined and can be regarded as a standard reference. Fluctuating noise was chosen because its resemblance to an interfering speaker is relevant for real-life communication, and because the benefit from temporal gaps in fluctuating noise is mostly less for hearing-impaired listeners than for normal-hearing listeners. In our experiments, the compression threshold was well below the average stimulus level to ensure compression of the entire signal. All stimuli were presented within the dynamic range of hearing of the individual subject and were therefore audible without the need for compression. This ensures that the effect of compression on speech intelligibility originated in the ability of compression to alter the *dynamics* of the stimuli and not in its ability to increase *audibility* of the stimulus as a whole (such as required for speech presented at too low a level or the previously mentioned low-level bird songs).

Chapter 1: Introduction

Chapter 1 starts with an overview of the workings of the peripheral ear. It describes sound transmission through the ear, and the consequences of damage to the ear. The chapter focuses on the cochlea, and on the current view that the cochlea is the primary origin of compressive amplification in the ear. It is explained that compressive amplification has a marked influence on auditory perception and can account for many of its nonlinear aspects. Based on literature, various estimates of the amount of compression are presented and from the data it was estimated that the bio-mechanical amplifier in the cochlea has a compression ratio of about 3 ± 2, depending on the level and frequency of the input sound wave. Hearing loss can lead to less cochlear compression, and many hearing aids apply compression to alleviate the hearing loss. Thus, the chapter concludes with a description of the properties and implementations of compression in hearing aids.

Chapter 2: Stationary noise

Chapter 2 explores the effects of compression on speech in a stationary background noise. For moderately hearing-impaired subjects in stationary noise, previous studies showed that compression with a small number of channels and with a low compression ratio did not degrade speech intelligibility relative to linear amplification. Some experiments even showed a slight improvement in speech intelligibility. Experiments with a large number of channels gave seemingly conflicting results. Hardly any previous studies have been designed to investigate interactions between compression parameters.

We measured the speech reception threshold (SRT) in noise, and expressed the SRT results for compression relative to the subject's result for linear amplification: ΔSRT. A lower value of ΔSRT means better speech intelligibility with compression than with linear amplification. We also introduced a notation to present the characteristics of compression under investigation by a set of compression parameters: $CP(NC, CR_{low}/CR_{high}, T_a/T_r)$, see section 2.1.1.

Our results indicate that most compression settings yielded lower scores than those obtained with linear amplification. Single-channel compression resulted in the worst speech intelligibility whereas the results for two-channel and six-channel compression were roughly equal to each other. There are some indications that the two-channel system might be optimized by investigating the best cross-over frequency (ours was 1 kHz). A compression ratio of 3/3 ($\mathrm{CR_{low}/CR_{high}}$) gave worse speech intelligibility than lower compression ratios. The combination of largest time constants ($\mathrm{T_a/T_r} = 40/400$ ms) gave best results. Our experiment was designed to investigate possible interaction effects between compression parameters. In other words, the study was designed to determine if the effect of a parameter (e.g., release time) is influenced by the value of another parameter (e.g., number of channels). We found a significant effect of NC * T meaning that the effect of time constants depended on the number of channels: fast-acting single-channel compression was detrimental for speech intelligibility. Moreover, the interaction NC * CR * T was significant, indicating that a specific combination of parameter values led to degraded (for instance *CP(1, 3/3, 4/4)*) or improved (*CP(2, 2/3, 40/40)*) results. Overall, the best result was obtained for two-channel compression *CP(2, 2/3, 40/40)* (ΔSRT = -0.7 dB). For six-channel compression the best speech intelligibility was achieved with *CP(6, 2/2 + 2/3, 40/400)* (ΔSRT = -0.4 dB) and the best result with single-channel compression was equal to that with linear amplification.

Chapter 3: Fluctuating noise

Chapter 3 investigates the effect of compression on speech intelligibility in a fluctuating background noise. The fluctuating noise consisted of spectrally speech-shaped noise with the time-reversed modulations of the same speaker as the target speech. The available results from previous research for moderately hearing-impaired subjects did not present a clear picture of the effects of compression in fluctuating noise. Although some studies showed improved speech intelligibility for amplitude compression (relative to linear amplification), most studies either showed a degradation or no difference. Compression results tended to be slightly more favourable for fluctuating noise than for stationary noise. Similar to stationary noise, for fluctuating noise possible interactions between the compression parameters were largely unknown.

The results from our experiments showed that compression had only a small influence on speech intelligibility. Although multi-channel compression is known to reduce spectral and temporal contrasts, its effect on speech intelligibility in fluctuating noise was limited. This is probably caused by the ability of compression to lift the level of the speech during gaps in the noise. For fluctuating noise, only one significant result was obtained: overall, six-channel compression led to significantly worse speech intelligibility than compression with NC = 1 or 2. Fast-acting single-channel compression gave less degraded results for fluctuating than for stationary noise and in contrast to results for stationary noise, interaction effects were not significant. Parameter values which resulted in the best speech intelligibility for stationary noise yielded good results for fluctuating noise as well. The best score (ΔSRT = -0.9 dB) was achieved with single-channel compression: *CP(1, 2/2, 40/40)*. A difference of -0.9 dB may seem small, but this corresponds to an improvement in sentence score of about 15%. The best two-channel result was found for *CP(2, 2/3, 40/40)* (ΔSRT = -0.7 dB). This combination also led to the overall best speech intelligibility for stationary noise (ΔSRT = -0.7 dB, Chapter 2). *CP(6, 2/3, 40/400)*, which led to the best speech intelligibility for six-channel compression with stationary noise (ΔSRT = -0.4 dB), gave a good result for fluctuating noise (ΔSRT $= -0.4$ dB) as well.

For fluctuating noise, the detrimental effects of compression were smaller than for stationary noise and the variance in the SRT results was larger. This suggests that stationary noise is more suited than fluctuating noise for measuring the effect of compression on speech intelligibility.

Chapter 4: Results individually examined

Chapter 4 evaluates the effect of compression on speech intelligibility for individual subjects. Most previous research did not find significant relationships between the effect of compression and audiometric characteristics. Past studies that did find a significant correlation with hearing loss (mostly tone-audiometric threshold or dynamic range of hearing) showed different effects: some studies found positive correlations, others reported negative correlations. It is therefore not clear which subjects can benefit most from compression. Moreover, we do not know which specific compression characteristics will yield optimal speech intelligibility for individual listeners. In view of these inconclusive results from the literature we conducted additional analyses on the large amount of data of which the pooled results were reported Chapters 2 and 3.

For our subjects, the effect of compression was rather small with respect to the within subject variability. Additionally the measurement error was larger for subjects with worse speech intelligibility in noise. The effect of compression was not significantly related to standard audiometric characteristics such as hearing threshold or dynamic range. However, for stationary noise, a significant correlation was found between the effect of compression (ΔSRT) and the SRT for linear amplification for the lowest two compression ratios (CR $= 1/2$ and CR $= 2/2$) only. Moreover, for this noise type, the best results obtained with compression improved significantly for subjects with a worse SRT. These correlations tentatively suggest that speech in stationary noise might be more useful for determining a possible benefit of compression than pure-tone audiometric threshold or dynamic range. In contrast, most current fitting rationales are threshold based.

Considerations for clinical applicability

The current parametric study was much more elaborate than will ever be possible in a clinical setting. However, one might apply our results to the practice of hearing aid fitting. This section will briefly discuss some issues that may arise when interpreting our results for use in a clinical setting.

Firstly, our results were obtained with Dutch speech material. Generally, the results of compression research are not regarded language specific. This assumption seems justified because the long-term average frequency content and the dynamic properties of many languages are very similar (Byrne et al., 1994)[1]. Moreover, Kam and Wong (1999) reported similar effects of compression for Cantonese (which is a tone language) as those obtained for English (which is an intonation language). It can therefore be expected that our results are also valid for other languages.

Secondly, the benefit of hearing aids can increase over time after the initial fitting. Unfortunately, our experimental design did not allow for acclimatization to each compression condition because of the large number (50) of different conditions. However, previous research indicates that the influence of acclimatization is rather limited. Turner et al. (1996) conducted a large review on previous research on acclimatization and concluded that there was a general tendency for hearing aid benefit to increase only slightly over time. In reply to the paper of Turner et al. (1996), Byrne (1996) stressed that measurements conducted without an acclimatization period may underestimate the potential benefit. Furthermore, direct comparisons between hearing aids might lead to the wrong conclusions because of a possible bias towards the characteristics of the subject's own hearing aid. Munro and Lutman (2003) found acclimatization for a high presentation level (about 90 dB SPL at the eardrum) only and not for lower levels. They suggested that the acclimatization effect depended on the sound level. Their subjects had not previously worn hearing aids and Munro and Lutman ascribed their finding to the fact that the subjects were not used to hearing sounds at this high a level. In our case, nearly all subjects used hearing aids prior to our experiments and were accustomed to the moderate sound levels of our study. Since acclimatization effects appear to be limited and since we found no significant correlation between the compression characteristics of the subjects' own hearing aid and their results in the experiment (Chapter 4), we assume that the results of our experiments remain valid. Still, it might prove beneficial to conduct a field trial which allows for sufficient acclimatization before using the results from our study.

Thirdly, this study investigated speech intelligibility at moderate input levels only. In real-life situations the range of encountered sound levels will be much larger than in our laboratory set-up with pre-recorded speech. However, our results can still be applied by fitting the hearing aid for these large dynamic ranges, while keeping in mind the influence of the parameter values on speech intelligibility. In other words, one should try to avoid compression with characteristics that lead to degraded speech intelligibility at normal speech levels (e.g., fast-acting single channel compression or high compression ratios).

[1]Note that whereas Dutch was not included in this study, closely related languages such as English, German, and Danish were, as well as more distant related languages such as Japanese, Russian, Arabic, etc.

How should we fit compression hearing aids?

Results indicate that for determining the best compression characteristics the SRT might be more suitable than pure-tone audiometric threshold or dynamic range. However, measurement error increased for subjects with a larger (worse) SRT and was generally larger for fluctuating noise. Moreover, the effects of compression seem limited for fluctuating noise, and since no clear individual predictors were found, one might consider using the average results. If one decides to use speech intelligibility to obtain a personalized fit, we suggest using stationary noise instead of fluctuating noise.

For our subjects, a compression ratio of 3 for the entire frequency spectrum gave lower speech intelligibility than lower compression ratios. Fast-acting single-channel compression led to degraded speech intelligibility for stationary noise. The influence of the number of channels was limited. For all channel configurations, short attack and release times resulted in lower speech intelligibility than longer time constants. These results suggest using multi-channel compression with compression ratios smaller than 3 and large time constants ($T_a/T_r > 4/40$).

Since there is more to life than speech intelligibility, it is important to look at other measures as well. A speech-based optimization is of not much use if the hearing-impaired is unsatisfied with the sound quality of the hearing aid and decides not to use it at all. The relatively small effects of compression on speech intelligibility warrants the use of compression for comfort enhancement in hearing aids. We suggest to use a first fit based on the average speech optimum for stationary noise, and then to individually optimize other factors such as comfort and ease of listening, audibility and clarity of sound, and music perception.

“They both savoured the strange warm glow of being much more ignorant than ordinary people, who were only ignorant of ordinary things”

Discworld scientists at work,

Terry Pratchett, ‘Equal Rites’,

HarperCollins, New York , 1987

Bibliography

ANSI (1987), *Specification of hearing aid characteristics*, American National Standards Institute, S3.22-1987.

ANSI (1996), *Specification of hearing aid characteristics*, American National Standards Institute, S3.22-1996.

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

Boike, K. T. and Souza, P. E. (2000), "Effect of compression ratio on speech recognition and speech-quality ratings with wide dynamic range compression amplification," Journal of Speech Language and Hearing Research 43, 456–468.

Boothroyd, A., Springer, N., Smith, L., and Schulman, J. (1988), "Amplitude compression and profound hearing loss," Journal of Speech and Hearing Research 31, 362–376.

Braida, L. D., Durlach, N. I., Lippmann, R. P., Hicks, B. L., Rabinowitz, W. M., and Reed, C. M. (1979), "Hearing aids–a review of past research on linear amplification, amplitude compression, and frequency lowering," Journal of the American Speech and Hearing Association Monographs 19, 1–114.

Brand, T. (2000), "Analysis and optimization of psychophysical procedures in audiology," Ph.D. thesis, University Oldenburg.

Bronkhorst, A. W. and Plomp, R. (1992), "Effect of multiple speech-like maskers on binaural speech recognition in normal and impaired hearing," Journal of the Acoustical Society of America 92, 3132–3139.

Busby, P. A., Tong, Y. C., and Clark, G. (1988), "Underlying structure of auditory-visual consonant perception by hearing-impaired children and the influences of syllabic compression," Journal of Speech and Hearing Research 31, 156–165.

Byrne, D. (1996), "Implications of acclimatization for hearing aid fitting practices and research," Ear and Hearing 17, 25S–26S.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R. M., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Kotby, N. M., Nasser, N. H. A., El Kholy, W. A. H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. I., Westerman, S., and Ludvigsen, C. (1994), "An international comparison of long-term average speech spectra," Journal of the Acoustical Society of America 96, 2108–2120.

Chi, T. S., Gao, Y. J., Guyton, M. C., Ru, P. W., and Shamma, S. (1999), "Spectro-temporal modulation transfer functions and speech intelligibility," Journal of the Acoustical Society of America 106, 2719–2732.

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," Journal of the Acoustical Society of America 97, 1854–1864.

Cox, R. M. and Alexander, G. C. (1995), "The abbreviated profile of hearing aid benefit," Ear and Hearing 16, 176–186.

Crain, T. R. and Yund, E. W. (1995), "The effect of multichannel compression on vowel and stop-consonant discrimination in normal-hearing and hearing-impaired subjects," Ear and Hearing 16, 529–543.

De Gennaro, S., Braida, L. D., and Durlach, N. I. (1986), "Multichannel syllabic compression for severely impaired listeners," Journal of Rehabilitation Research and Development 23, 17–24.

Dillon, H. (1996), "Compression?: Yes, but for low or high frequencies, for low or high intensities, and with what response times?" Ear and Hearing 17, 287–307.

Dillon, H. (2001), *Hearing Aids* (Thieme Medical Publishers, New York).

Dillon, H., Storey, L., Grant, F., Philips, A., Skelt, L., Mavrias, G., Woytowych, W., and Walsh, M. (1998), "Preferred compression threshold with 2:1 wide dynamic range compression in everyday environments," Australian Journal of Audiology 20, 33–44.

Dobie, R. (1997), "Drug treatments for sensorineural hearing loss and tinnitus." in *Neurotransmission and Hearing Loss*, edited by C. Berlin (Singular Publishing, San Diego).

Dreschler, W. A., Eberhardt, D., and Melk, P. W. (1984), "The use of single-channel compression for the improvement of speech intelligibility," Scandinavian Audiology 13, 231–236.

Drullman, R. and Bronkhorst, A. (2004), "Speech perception and talker segregation: effects of level, pitch, and tactile support with multiple simultaneous talkers," Journal of the Acoustical Society of America 116, 3090–3098.

Drullman, R. and Smoorenburg, G. F. (1997), "Audio-visual perception of compressed speech by profoundly hearing-impaired subjects," Audiology 36, 165–177.

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

Duijvestijn, J. A., Anteunis, L. J. C., Hendriks, J. J. T., and Manni, J. J. (1999), "Definition of hearing impairment and its effect on prevalence figures - A survey among senior citizens," Acta Oto-Laryngologica 119, 420–423.

Eguìluz, J. A., Ospeck, M., Choe, Y., Hudspeth, A., and Magnasco, M. (2000), "Essential nonlinearities in hearing," Physical Review Letters 84, 5232–5235.

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

Festen, J. and Van Dijkhuizen, J. (1999), "Modeling the speech-reception threshold for amplitude-compressed speech," in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific Publishers, Singapore).

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

Franck, B. A., Kreveld-Bos, C. S., Dreschler, W. A., and Verschuure, H. (1999), "Evaluation of spectral enhancement in hearing aids, combined with phonemic compression," Journal of the Acoustical Society of America 106, 1452–1464.

Geisler, C. D. (1998), *From Sound to Synapse* (Oxford University Press, Oxford).

Giguère, C. and Smoorenburg, G. (1997), "The generation of psychoacoustic combination tones in relation to two-tone suppression effects in a computational model," Journal of the Acoustical Society of America 102, 155–164.

Giguère, C. and Smoorenburg, G. (1999), "Computational modeling of outer hair cell damage: implications for hearing aid signal processing," in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific Publishers, Singapore).

Goedegebure, A., Hulshof, M., Maas, R. J. J., Dreschler, W. A., and Verschuure, H. (2001), "Effects of single-channel phonemic compression schemes on the understanding of speech by hearing-impaired listeners," Audiology 40, 10–25.

Goldstein, J. L. (1967), "Auditory nonlinearity," Journal of the Acoustical Society of America 41, 676–689.

Grimwood, W. K. (1949), "Volume compressors for sound recordings." Journal of the Society for Motion Picture Engineers 52, 49–76.

Gustafsson, H. A. and Arlinger, S. D. (1994), "Masking of speech by amplitude-modulated noise," Journal of the Acoustical Society of America 95, 518–529.

Hagerman, B. (1997), "Attempts to develop an efficient speech test in fully modulated noise," Scandinavian Audiology 26, 93–98.

Hagerman, B. (2002), "Speech recognition threshold in slightly and fully modulated noise for hearing-impaired subjects," International Journal of Audiology 41, 321–329.

Harte, J. M. and Elliott, S. J. (2005), "A comparison of various nonlinear models of cochlear compression," Journal of the Acoustical Society of America 117, 3777–3786.

Helmholtz, H. L. F. (1875), *On the Sensations of Tone as a Physiological Basis for the Theory of Music.* (Reprint, Dover 1954, New York).

Hickson, L. (1994), "Compression amplification in hearing aids," American Journal of Audiology 3, 51–65.

Hickson, L., Dodd, B., and Byrne, D. (1995), "Consonant perception with linear and compression amplification," Scandinavian Audiology 24, 175–184.

Houtgast, T. (1972), "Psychophysical evidence for lateral inhibition in hearing," Journal of the Acoustical Society of America 51, 1885–1894.

Humes, L. E., Christensen, L., Thomas, T., Bess, F. H., Hedley-Williams, A., and Bentler, R. A. (1999), "A comparison of the aided performance and benefit provided by a linear and a two-channel wide dynamic range compression hearing aid," Journal of Speech Language and Hearing Research 42, 65–79.

Humes, L. E., Christensen, L. A., Bess, F. H., and Hedley-Williams, A. (1997), "A comparison of the benefit provided by well-fit linear hearing aids and instruments with automatic reductions of low-frequency gain," Journal of Speech Language and Hearing Research 40, 666–685.

Humes, L. E. and Wilson, D. L. (2004), "A comparison of single-channel linear amplification and two-channel wide-dynamic-range-compression amplification by means of an independent-group design," American Journal of Audiology 13, 39–53.

IEC (1983), *Hearing Aids - Part 2: Hearing Aids with automatic gain control circuits*, International Electrotechnical Commission, IEC60118-2.

Jerger, S. and Jerger, J. (1981), *Auditory Disorders, A Manual for Clinical Evaluation* (Little, Brows and Company, Boston).

Kam, A. C. and Wong, L. L. (1999), "Comparison of performance with wide dynamic range compression and linear amplification," Journal of the American Academy of Audiology 10, 445–457.

Kates, J. M. (1993), "Optimal estimation of hearing-aid compression parameters," Journal of the Acoustical Society of America 94, 1–12.

Killion, M. C. (1993), "The K-Amp hearing aid: An attempt to present high fidelity for persons with impaired hearing," American Journal of Audiology 2, 52–74.

Killion, M. C., Staab, W. J., and Preves, D. (1990), "Classifying automatic signal processors," Hearing Instruments 41, 24–26.

King, A. B. and Martin, M. C. (1984), "Is AGC beneficial in hearing aids?" British Journal of Audiology 18, 31–38.

Kochkin, S. (1999), "'Baby Boomers' spur growth in potential market, but penetration rate declines," Hearing Journal 52, 33–48.

Kochkin, S. (2000), "MarkeTrak V: Why are my hearing aids in the drawer: The consumers' perspective," The Hearing Journal 53, 34–41.

Kochkin, S. (2002), "MarkeTrak VI: Consumers Rate Improvements Sought in Hearing Instruments," Hearing Review 9, 18–22.

Lippmann, R., Braida, L. D., and Durlach, N. I. (1981), "Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss," Journal of the Acoustical Society of America 69, 524–534.

Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (2003), "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing," Journal of the Acoustical Society of America 113, 951–960.

Lopez-Poveda, E. A., Plack, C. J., Meddis, R., and Blanco, J. L. (2005), "Cochlear compression in listeners with moderate sensorineural hearing loss," Hearing Research 205, 172–183.

Lyzenga, J., Festen, J., and Houtgast, T. (2002), "A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity," Journal of the Acoustical Society of America 112, 1145–1157.

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

Marriage, J. E. and Moore, B. C. J. (2003), "New speech tests reveal benefit of wide-dynamic-range, fast- acting compression for consonant discrimination in children with moderate-to-profound hearing loss," International Journal of Audiology 42, 418–425.

Møller, A. (2000), *Hearing. Its Physiology and Pathophysiology* (Academic Press, San Diego).

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

Moore, B. C. and Glasberg, B. R. (1986), "A comparison of two-channel and single-channel compression hearing aids," Audiology 25, 210–226.

Moore, B. C. and Glasberg, B. R. (1988), "A comparison of four methods of implementing automatic gain control (AGC) in hearing aids," British Journal of Audiology 22, 93–104.

Moore, B. C., Johnson, J. S., Clark, T. M., and Pluvinage, V. (1992), "Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss," Ear and Hearing 13, 349–370.

Moore, B. C., Peters, R. W., and Stone, M. A. (1999), "Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips," Journal of the Acoustical Society of America 105, 400–411.

Moore, B. C., Stainsby, T. H., Alcantara, J. I., and Kuhnel, V. (2004), "The effect on speech intelligibility of varying compression time constants in a digital hearing aid," International Journal of Audiology 43, 399–409.

Munro, K. J. and Lutman, M. E. (2003), "The effect of speech presentation level on measurement of auditory acclimatization to amplified speech," Journal of the Acoustical Society of America 114, 484–495.

Murugasu, E. (2005), "Recent advances in the treatment of sensorineural deafness," Annals Academy of Medicine Singapore 34, 313–321.

Nábělek, I. V. (1983), "Performance of hearing-impaired listeners under various types of amplitude compression," Journal of the Acoustical Society of America 74, 776–791.

Nábělek, I. V. (1984), "Discriminability of the quality of amplitude-compressed speech," Journal of Speech and Hearing Research 27, 571–577.

Nedzelnitsky, V. (1980), "Sound pressures in the basal turn of the cat cochlea," Journal of the Acoustical Society of America 68, 1676–1689.

Nelson, D. A. and Freyman, R. L. (1987), "Temporal resolution in sensorineural hearing-impaired listeners," Journal of the Acoustical Society of America 81, 709–720.

Nelson, D. A., Schroder, A. C., and Wojtczak, M. (2001), "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners," Journal of the Acoustical Society of America 110, 2045–2064.

Neuman, A. C., Bakke, M. H., Hellman, S., and Levitt, H. (1994), "Effect of compression ratio in a slow-acting compression hearing aid: paired-comparison judgments of quality," Journal of the Acoustical Society of America 96, 1471–1478.

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

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

Novick, M. L., Bentler, R. A., Dittberner, A., and Flamme, G. A. (2001), "Effects of release time and directionality on unilateral and bilateral hearing aid fittings in complex sound fields," Journal of the American Academy of Audiology 12, 534–544.

Olsen, H. L., Olofsson, A., and Hagerman, B. (2004), "The effect of presentation level and compression characteristics on sentence recognition in modulated noise," International Journal of Audiology 43, 283–294.

Oppenheim, A. V. and Schafer, R. W. (1989), *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs).

Oxenham, A. J. and Bacon, S. P. (2003), "Cochlear compression: Perceptual measures and implications for normal and impaired hearing," Ear and Hearing 24, 352–366.

Oxenham, A. J. and Plack, C. J. (1997), "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," Journal of the Acoustical Society of America 101, 3666–3675.

Palmer, A. and Shamma, S. (2004), "Physiological representations of speech," in *Speech Processing in the Auditory System*, edited by S. Greenberg, W. A. Ainsworth, A. N. Popper, and R. R. Fay (Springer, New York).

Peterson, M. E., Feeney, M. P., and Yantis, P. A. (1990), "The effect of automatic gain control in hearing-impaired listeners with different dynamic ranges," Ear and Hearing 11, 185–194.

Pickles, J. (1988), *An Introduction to the Physiology of Hearing* (Academic Press, London).

Plack, C. J. and Drga, V. (2003), "Psychophysical evidence for auditory compression at low characteristic frequencies," Journal of the Acoustical Society of America 113, 1574–1586.

Plomp, R. (1994), "Noise, amplification, and compression: considerations of three main issues in hearing aid design," Ear and Hearing 15, 2–12, published erratum appears in Ear Hear 1994,15, 6 (following table of contents).

Plomp, R. and Mimpen, A. M. (1979), "Improving the reliability of testing the speech reception threshold for sentences." Audiology 18, 43–52.

Preves, D. (1991), "Output limiting and speech enhancement," in *The Vanderbilt Hearing Aid Report II*, edited by G. A. Studebaker, F. H. Bess, and L. B. Beck (York Press, Parkton, MD).

Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (1998), "Basilar-membrane responses to clicks at the base of the chinchilla cochlea," Journal of the Acoustical Society of America 103, 1972–1989.

Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (2005), "Release from informational masking by time reversal of native and non-native interfering speech," Journal of the Acoustical Society of America 118, 1274–1277.

Rintelmann, W. F. (1972), "Effects of amplitude compression upon speech perception: a review of research," Scandinavian Audiology 1, 127–134.

Robinson, C. E. and Huntington, D. A. (1973), "The intelligibility of speech processed by delayed long-term averaged compression amplification," Journal of the Acoustical Society of America 54, 314.

Robles, L. and Ruggero, M. A. (2001), "Mechanics of the mammalian cochlea," Physiological Reviews 81, 1305–1352.

Roeser, R. J., Valente, M., and Hosford-Dunn, H. (2000), *Audiology - Diagnosis* (Thieme Medical Publishers, New York).

Rose, J., Brugge, J., Anderson, D., and Hind, J. (1967), "Phase-locked response to low-frequency tones in single auditory nerve fibres in the squirrel monkey," Journal of Neurophysiology 30, 769–793.

Rosengard, P. S., Oxenham, A. J., and Braida, L. D. (2005), "Comparing different estimates of cochlear compression in listeners with normal and impaired hearing," Journal of the Acoustical Society of America 117, 3028–3041.

Rouiller, E. M. (1997), "Functional organization of the auditory pathways," in *The Central Auditory System*, edited by G. Ehret and R. Romand (Oxford University Press, Oxford).

Ruggero, M. A. and Rich, N. C. (1991), "Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane," Journal of Neuroscience 11, 1057–1067.

Ruggero, M. A., Robles, L., and Rich, N. C. (1992), "Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory-nerve rate suppression," Journal of Neurophysiology 68, 1087–1099.

Sachs, M. B. and Kiang, N. Y. S. (1968), "Two-tone inhibition in auditory nerve fibers." Journal of the Acoustical Society of America 43, 1120–1128.

Shanks, J. E., Wilson, R. H., Larson, V., and Williams, D. (2002), "Speech recognition performance of patients with sensorineural hearing loss under unaided and aided conditions using linear and compression hearing aids," Ear and Hearing 23, 280–290.

Smith, S. (1997), *The Scientist and Engineer's Guide to Digital Signal Processing* (California Technical Publishing, San Diego).

Smoorenburg, G. F. (1972), "Combination tones and their origin," Journal of the Acoustical Society of America 52, 615–632.

Smoorenburg, G. F. (1974), "On the mechanisms of combination tone generations and lateral inhibition in hearing," in *Psychophysical Models and Physiological Facts in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, New York).

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

Souza, P. E. and Kitch, V. (2001), "The contribution of amplitude envelope cues to sentence identification in young and aged listeners," Ear and Hearing 22, 112–119.

Stone, M. A., Moore, B. C., Wojtczak, M., and Gudgin, E. (1997), "Effects of fast-acting high-frequency compression on the intelligibility of speech in steady and fluctuating background sounds," British Journal of Audiology 31, 257–273.

Stuart, A. and Phillips, D. P. (1996), "Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners," Ear and Hearing 17, 478–489.

Summers, V. and Molis, M. R. (2004), "Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level," Journal of Speech Language and Hearing Research 47, 245–256.

The Capital Region Otitis Media Project (2000), *Ear Infections in Children*, State of New York, Department of Health, New York, NY, Publication 4815.

The Mathworks, Inc. (2001), "MATLAB, R12.1," Natick, MA.

Turner, C. W., Humes, L., Bentler, R., and Cox, R. (1996), "A review of past research on changes in hearing aid benefit over time," Ear and Hearing 17, 14S–25S.

Tyler, R. S. and Kuk, F. K. (1989), “The effects of ’noise suppression’ hearing aids on consonant recognition in speech-babble and low-frequency noise [published erratum appears in Ear Hear 1989 Dec;10(6):399],” Ear and Hearing 10, 243–249.

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

Van Tasell, D. J. and Trine, T. D. (1996), “Effects of single-band syllabic amplitude compression on temporal speech information in nonsense syllables and in sentences,” Journal of Speech and Hearing Research 39, 912–922.

Van Toor, T. and Verschuure, H. (2002), “Effects of high-frequency emphasis and compression time constants on speech intelligibility in noise,” International Journal of Audiology 41, 379–394.

Verschuure, J., Benning, F. J., Van Cappellen, M., Dreschler, W. A., and Boeremans, P. P. (1998), “Speech intelligibility in noise with fast compression hearing aids,” Audiology 37, 127–150.

Verschuure, J., Dreschler, W. A., De Haan, E. H., Van Cappellen, M., Hammerschlag, R., Maré, M. J., Maas, A. J., and Hijmans, A. C. (1993), “Syllabic compression and speech intelligibility in hearing impaired listeners,” Scandinavian Audiology. Supplementum 38, 92–100.

Verschuure, J., Stikvoort, E., De Jong, R. M., Goedegebure, A., and Dreschler, W. A. (1996), “Compression and its effect on the speech signal,” Ear and Hearing 17, 162–175.

Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000), “Method for the selection of sentence materials for efficient measurement of the speech reception threshold,” Journal of the Acoustical Society of America 107, 1671–1684.

Versfeld, N. J. and Dreschler, W. A. (2002), “The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners,” Journal of the Acoustical Society of America 111, 401–408.

Villchur, E. (1973), “Signal processing to improve speech intelligibility in perceptive deafness,” Journal of the Acoustical Society of America 53, 1646–1657.

Villchur, E. (1978), “A critical survey of research on amplitude compression,” Scandinavian Audiology. Supplementum 6, 305–314.

Villchur, E. (1987), “Multichannel compression processing for profound deafness,” Journal of Rehabilitation Research and Development 24, 135–148.

Von Békésy, G. (1960), *Experiments in Hearing* (McGraw-Hill, New York).

Walker, G., Byrne, D., and Dillon, H. (1984), "The effects of multichannel compression/expansion amplification on the intelligibility of nonsense syllables in noise," Journal of the Acoustical Society of America 76, 746–757.

Yanick, P. (1976a), "Discrimination in the presence of competition with an AVC vs. a DRC hearing aid." Journal of the American Audiology Society 1, 169–172.

Yanick, P. (1976b), "Effects of signal processing on intelligibility of speech in noise for persons with sensorineural hearing loss," Journal of the American Audiology Society 1, 229–238.

Yund, E. W. and Buckles, K. M. (1995a), "Enhanced speech perception at low signal-to-noise ratios with multichannel compression hearing aids," Journal of the Acoustical Society of America 97, 1224–1240.

Yund, E. W. and Buckles, K. M. (1995b), "Multichannel compression hearing aids: effect of number of channels on speech discrimination in noise," Journal of the Acoustical Society of America 97, 1206–1223.

Yund, E. W., Simon, H. J., and Efron, R. (1987), "Speech discrimination with an 8-channel compression hearing aid and conventional aids in background of speech-band noise," Journal of Rehabilitation Research and Development 24, 161–180.

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., and Starr, A. (2005), "Perceptual consequences of disrupted auditory nerve activity," Journal of Neurophysiology 93, 3050–3063.

Nederlandstalige samenvatting

In moderne hoortoestellen gebruikt men geavanceerde signaalbewerking om het inkomende geluid te bewerken en te versterken. Het hoortoestel is meer dan een eenvoudige geluidversterker: het toestel reageert actief op het binnenkomende geluid. De signaalbewerking is complex. Het toestel wordt gestuurd door de eigenschappen van het ingaande geluid zoals geluidsterkte, fluctuaties in geluidsterkte en frequentie-inhoud van het signaal.

Desondanks zijn veel hoortoestelgebruikers toch niet volledig tevreden met hun toestel. Een veelgehoorde klacht is dat het moeilijk is om een gesprek te volgen in een rumoerige omgeving, zoals in een restaurant of op een feestje. Vooral mensen met een zogenaamd perceptief gehoorverlies hebben veel last van achtergrondlawaai. Een perceptief gehoorverlies wordt veroorzaakt door schade aan het binnenoor: het komt door "kapotte gehoorcellen". Zelfs als geavanceerde hoortoestellen worden gebruikt, leidt een perceptief gehoorverlies veelal tot een verminderd spraakverstaan in lawaai. Helaas komt deze vorm van gehoorverlies veelvuldig voor. Dit verlies heeft veel gangbare oorzaken zoals ouderdom, excessieve blootstelling aan hoge geluidniveaus, erfelijke factoren en voor het oor schadelijke medicatie.

Naast verminderd spraakverstaan heeft perceptief gehoorverlies nog een vervelend gevolg: geluiden met een laag niveau ("zachte geluiden") worden niet meer waargenomen terwijl geluiden met een hoog niveau ("harde geluiden") even hard en soms zelfs harder klinken. Er zit dus een kern van waarheid in de karikatuur van de (perceptief) slechthorende die zegt *"Kun je iets harder praten, ik versta je niet"* en vervolgens roept *"Hé! Je hoeft niet zo te schreeuwen, ik ben niet doof!"*. Om dit gevolg van perceptief gehoorverlies te verminderen, wordt vaak gebruik gemaakt van compressieve versterking. Deze compressieve versterking (of kort: compressie) kan automatisch het volume van het hoortoestel regelen. Zachte geluiden worden automatisch versterkt en harde geluiden blijven comfortabel doordat ze weinig of geen versterking krijgen. Zo vertelde één van onze proefpersonen dat ze door haar nieuwe compressieve hoortoestellen weer kan genieten van haar ochtendwandelingen. Ze kan nu de vogels horen fluiten zonder dat ze aan de volumeknop hoeft te draaien en zonder te vrezen voor pijnlijk harde geluiden van passerend verkeer.

Compressie in hoortoestellen kan niet alleen het luistercomfort verbeteren, het kan ook de spraakverstaanbaarheid beïnvloeden. Helaas is nog niet duidelijk wat de invloed is van belangrijke compressieparameters op het spraakverstaan in lawaai. Bovendien is het onbekend of het effect van een bepaalde parameter wordt beïnvloed door de instelling van een andere parameter (interactie). We hebben daarom een uitgebreid experiment opgezet waarin de effecten van vier compressieparameters in onderlinge samenhang werden onderzocht. Ons onderzoek was volledig parametrisch. Dit wil zeggen dat alle mogelijke combinaties van de compressieparameters zijn gebruikt. De onderzochte parameters zijn het aantal frequentiekanalen ($NC = 1, 2$, of 6), de compressieratio voor frequenties respectievelijk onder en boven 1 kHz ($CR_{low}/CR_{high} = 1/2$ tot $3/3$), en de in- en uitregeltijd ($T_a/T_r = 4$ tot 400 ms). Als referentie hebben we versterking zonder compressie, oftewel lineaire versterking, gebruikt.

In onze studies zijn twee soorten ruis gebruikt: stationaire en fluctuerende ruis. Beide typen ruis hadden de spectrale kenmerken (frequentie-inhoud) van spraak. Stationaire ruis klinkt als het constante geruis van een waterval. Deze ruis hebben we gekozen omdat deze eenduidig is gedefinieerd en omdat deze wordt gezien als een standaard. Onze fluctuerende ruis had de temporele veranderingen (modulaties) van achterstevoren afgespeelde spraak. Dit type achtergrondruis is gekozen omdat het klinkt als een onverstaanbare spreker en dus relevant is voor de dagelijkse praktijk van een slechthorende. Bovendien hebben slechthorenden in het algemeen minder profijt van de momenten van lage geluidsterkte in fluctuerende ruis dan normaalhorenden. Alle experimenten zijn uitgevoerd met een lage compressiedrempel zodat het hele signaal werd gecomprimeerd. Alle geïncludeerde proefpersonen hadden een matig perceptief gehoorverlies. De stimuli werden aangeboden op een geluidniveau dat binnen het resterend dynamisch bereik van de proefpersonen lag, zodat de spraak ook zonder compressie hoorbaar was. Dit heeft als voordeel dat het effect van compressie op het spraakverstaan werd veroorzaakt door de verandering in het spraaksignaal en niet door een betere hoorbaarheid van de spraak (zoals bijvoorbeeld nodig is voor te zachte spraak of voor de bovengenoemde zachte vogelgeluiden).

Hoofdstuk 1 geeft een overzicht van de werking van ons gehoor. Het beschrijft de geluidtransmissie door het oor en de gevolgen van schade aan het oor. De focus van het hoofdstuk ligt op de cochlea (het slakkenhuis). De huidige gedachte is dat de cochlea de primaire bron is van compressieve versterking in het oor zelf. Deze cochleaire versterking heeft een grote invloed op de auditieve perceptie. Het kan veel niet-lineaire aspecten van de geluidverwerking door het gehoor verklaren. Als de cochleaire compressie is verminderd door gehoorbeschadiging, dan is het wellicht zinvol om een hoortoestel van compressie te voorzien. Het hoofdstuk eindigt daarom met een korte beschrijving van eigenschappen en implementaties van compressie in hoortoestellen.

Hoofdstuk 2 bestudeert het effect van compressie op het spraakverstaan in een stationaire achtergrondruis. De studie richt zich op mensen met een matig perceptief gehoorverlies. Enkele voorgaande onderzoeken lieten een kleine verbetering in spraakverstaan zien met compressie ten opzichte van lineaire versterking. Hoewel sommige experimenten resulteerden in verslechterd spraakverstaan met compressie, bleek in het algemeen dat compressie het spraakverstaan niet verslechterde ten opzichte van lineaire versterking. Dit was met name het geval voor compressie met een beperkt aantal frequentiekanalen en een lage compressieratio. Experimenten met een groter aantal frequentiekanalen gaven onderling verschillende resultaten. Nagenoeg geen enkel voorgaand onderzoek was gericht op het vaststellen van mogelijke interacties tussen de compressieparameters.

In ons onderzoek werd de 'Speech Reception Threshold in ruis' (SRT) gebruikt als maat voor spraakverstaanbaarheid. De focus lag op het effect van compressieve versterking ten opzichte van lineaire versterking. Daarom hebben we ΔSRT geïntroduceerd: ΔSRT is het verschil tussen de spraakverstaanbaarheid met enerzijds compressie en anderzijds lineaire versterking. Een lagere ΔSRT komt overeen met een betere spraakverstaanbaarheid met compressie dan met lineaire versterking. De compressieparameters van een specifiek compressiesysteem noteren we als $CP(NC, CR_{low}/CR_{high}, T_a/T_r)$ (oftewel: het aantal kanalen, de compressieratio en de in- en uitregeltijd).

Ons experiment liet zien dat de meeste compressiecondities een slechtere spraakverstaanbaarheid opleverden dan lineaire versterking. De resultaten voor twee- en zeskanaals-compressie waren nagenoeg gelijk aan elkaar. Enkelkanaals-compressie gaf de slechtste resultaten. De resultaten suggeerden dat de tweekanaals-compressie wellicht zou kunnen worden verbeterd door optimalisatie van de frequentie waarbij de compressieratio werd gesplitst (1000 Hz). Een compressieratio van 3/3 gaf slechtere resultaten dan de lagere compressieratio's. De combinatie van de grootste tijdconstanten ($T_a/T_r = 40/400$ ms) resulteerde in betere spraakverstaanbaarheid dan snellere compressie.

Het experiment was specifiek opgezet om mogelijke interactie-effecten te bestuderen. Of anders gezegd, we wilden weten of het effect van een bepaalde parameter (bijvoorbeeld de uitregeltijd) zou worden beïnvloed door de instelling van een andere parameter (bijvoorbeeld het aantal frequentiekanalen). De resultaten lieten een statistisch significant effect van NC * T zien. Dit betekent dat het effect van de tijdconstanten af hing van het aantal frequentiekanalen: snelle enkelkanaals-compressie gaf een slechtere spraakverstaanbaarheid dan snelle meerkanaals-compressie. Bovendien was de interactie NC * CR * T significant. Dit geeft aan dat een specifieke combinatie van compressieparameters een verslechtering (bijvoorbeeld $CP(1, 3/3, 4/4)$) of een verbetering (zoals $CP(2, 2/3, 40/40)$) kan opleveren. De beste spraakverstaanbaarheid werd bereikt met tweekanaals-compressie ($CP(2, 2/3, 40/40)$, ΔSRT $= -0.7$ dB). Het beste resultaat met zeskanaals-compressie werd gemeten voor $CP(6, 2/2 + 2/3, 40/400)$

(ΔSRT $= -0.4$ dB). De beste spraakverstaanbaarheid voor enkelkanaals-compressie was gelijk aan de spraakverstaanbaarheid voor lineaire versterking. De verbeteringen van 0.7 en 0.4 dB komen overeen met verbeteringen in het spraakverstaan van respectievelijk 12 en 7%.

Hoofdstuk 3 richt zich op compressie van spraak in een fluctuerende achtergrondruis. De achtergrondruis had het frequentiespectrum van spraak, met de achterstevoren gedraaide temporele modulaties van de spreker. Voorgaand onderzoek voor matig perceptief gehoorverlies gaf geen duidelijk beeld van het effect van compressie in een fluctuerende ruis. Hoewel een enkel onderzoek een verbetering in spraakverstaanbaarheid liet zien onder invloed van compressie (ten opzichte van lineaire versterking), resulteerden de meeste onderzoeken in gelijke spraakverstaanbaarheid of zelfs in een verslechtering. Compressie in fluctuerende ruis gaf iets betere resultaten dan in stationaire ruis. Voor fluctuerende ruis is, net als voor stationaire ruis, weinig bekend over interacties tussen de compressieparameters.

Ons experiment liet zien dat compressie slechts een kleine invloed had op de spraakverstaanbaarheid. Ondanks de bekende spectrale en temporele contrastverlaging door meerkanaals-compressie bleek het effect van meerkanaals-compressie op de spraakverstaanbaarheid toch beperkt. Dit werd waarschijnlijk veroorzaakt doordat compressie de laag-energetische delen van spraak extra kan versterken tijdens de gaten in de ruis waardoor de temporele maskering vermindert. Het experiment leverde slechts één statistisch significant resultaat op: zeskanaals-compressie gaf slechtere spraakverstaanbaarheid dan enkel- en tweekanaals-compressie. Ten opzichte van stationaire ruis gaf enkelkanaals-compressie in fluctuerende ruis een minder slechte spraakverstaanbaarheid. In tegenstelling tot stationaire ruis waren interactie-effecten niet significant voor fluctuerende ruis. De compressie-instellingen die goede resultaten opleverden voor stationaire ruis, gaven ook goede resultaten voor fluctuerende ruis. De beste score (ΔSRT $= -0.9$ dB) werd gemeten voor enkelkanaals-compressie: $CP(1, 2/2, 40/40)$. Een verschil van -0.9 dB lijkt erg weinig, maar het komt overeen met een verbetering in zinscore van ongeveer 15%. De beste spraakverstaanbaarheid voor tweekanaals-compressie was gemeten voor $CP(2, 2/3, 40/40)$ (ΔSRT $= -0.7$ dB). Deze combinatie leverde ook de beste spraakverstaanbaarheid voor stationaire ruis (ΔSRT $= -0.7$ dB, hoofdstuk 2). De combinatie $CP(6, 2/3, 40/400)$ gaf voor zeskanaals-compressie de beste spraakverstaanbaarheid in stationaire ruis en resulteerde in een goed resultaat voor fluctuerende ruis (ΔSRT $= -0.4$ dB).

De resultaten suggereren dat stationaire ruis beter geschikt is dan fluctuerende ruis om het effect van compressie op de spraakverstaanbaarheid te meten. Voor stationaire ruis waren de (negatieve) effecten van compressie groter en de variantie in de SRT-resultaten kleiner.

Hoofdstuk 4 evalueert het effect van compressie op het spraakverstaan voor individuele luisteraars. In het merendeel van de voorgaande onderzoeken werd geen statistisch significante correlatie gevonden tussen het effect van compressie op de spraakverstaanbaarheid en de audiometrische kenmerken. Voorgaande onderzoeken die wel een significante correlatie gaven, lieten zowel positieve als negatieve correlaties zien. Het bleef daarom onduidelijk voor welke slechthorenden compressie het meeste nut heeft. Bovendien werd niet duidelijk welke specifieke compressie-eigenschappen het beste spraakverstaan opleveren voor individuele luisteraars. Omdat in voorgaand onderzoek geen duidelijk beeld werd geschetst, hebben we de grote hoeveelheid data van de voorgaande hoofdstukken opnieuw geanalyseerd.

Voor onze proefpersonen was het effect van compressie nogal klein ten opzichte van de variantie binnen proefpersonen. Bovendien was de meetfout groter voor proefpersonen met een slechter spraakverstaan in ruis. Het effect van compressie was niet significant gecorreleerd met standaard audiometrische kenmerken zoals de hoordrempel en het dynamisch bereik. Voor stationaire ruis lieten de resultaten echter een significante relatie zien tussen het effect van compressie (ΔSRT) en de SRT voor lineaire versterking voor de laagste twee compressie ratio's (CR $= 1/2$ en CR $= 2/2$). Voor stationaire ruis bleek het beste resultaat van compressie significant beter te zijn voor proefpersonen met een slechtere SRT. Deze correlaties suggereren dat een eventueel effect van compressie beter zou kunnen worden vastgesteld aan de hand van metingen van het spraakverstaan in stationaire ruis dan aan de hand van toondrempels of het dynamisch bereik. De meeste regels voor de aanpassing van hoortoestellen zijn echter gebaseerd op toondrempels.

Hoofdstuk 5 geeft een samenvatting en bespreekt kort de toepasbaarheid van de resultaten in de dagelijkse praktijk van hoortoestelaanpassing. Er wordt voorgesteld om een hoortoestel in eerste aanleg aan te passen op basis van spraakverstaanbaarheid in stationaire ruis, bijvoorbeeld aan de hand van de algemene resultaten van deze studie. Daarna kunnen andere belangrijke zaken individueel worden geoptimaliseerd zoals bijvoorbeeld hoortoestelcomfort, luistergemak, hoorbaarheid, helderheid van het geluid en muziekbeleving.

Dankwoord

In de afgelopen jaren heb ik in het AZU met veel plezier aan mijn proefschrift gewerkt. Graag wil ik enkele mensen bedanken voor hun ondersteuning. Zonder hen was de afgelegde weg een stuk zwaarder geweest.

Allereerst wil ik mijn promotor, prof.dr. Guido F. Smoorenburg, noemen. Guido, hartelijk bedankt voor al je hulp en voor de plezierige samenwerking. Je scherpe blik en kritische houding zorgden ervoor dat elke bespreking een leuke en leerzame ervaring was.

Organisatorisch was het was wel eens lastig dat ik geen co-promotor of dagelijkse begeleider had. Op zulke momenten sprong Sjaak Klis in de bres. Sjaak, ik heb je inzet, ondersteuning en gezelligheid erg gewaardeerd.

Tijdens een promotie breng je een groot gedeelte van je tijd door met je kamergenoten. Daarmee had ik het bijzonder getroffen op het AZU. Mijn eerste kamergenoot was Natalie Cappaert. Naast een liefde voor het vak deelde ik met haar een voorliefde voor dropjes, punkmuziek en het spelen van squash. Bovendien heb ik als fysicus veel geleerd van haar biologielessen. Natalie, ik vond het erg fijn om met je gewerkt te hebben. Maar als ik je vaker met squash had laten winnen, was je dan wel in Utrecht gebleven? Ik vind het heel leuk dat je mijn paranimf wilt zijn! Nadat Natalie Utrecht verruild had voor Amsterdam kwam Huib Versnel. Huib, als fysicus heb je een nuchtere blik op het vakgebied. Onder het genot van een 'tas koffie' hebben wij dan ook vele leuke en waardevolle discussies gevoerd. Dank!

Als AiO zit je allemaal in hetzelfde schuitje. Eenmaal van wal vertrokken op ontdekkingstocht in nieuwe gebieden is het erg fijn als je weet dat je enkele medevaarders hebt. Bij deze wil ik de AiO's bedanken die er mede voor gezorgd hebben dat ik niet afdwaalde of verdronk. Marjolein, vanaf het begin konden we het uitstekend samen vinden. Van *"maar het is toch maar 100 gulden"* tot het geweldige *"Volkswagen in een volksbuurt"!* Ciska, hartelijk dank voor je collegialiteit. Martijn, je eerst-praten-dan-denken-mentaliteit bracht een frisse sfeer op de afdeling, geweldig.

Hulpvaardige collega's zijn onmisbaar en daarom wil ik Pieter Lamoré bedanken. Piet, dankzij je grote audiologische kennis, je oog voor details en je hulpvaardigheid heb je mij heel goed geholpen, bedankt!

"Do Lipton employees take coffee breaks?" vroeg de acteur Steven Wright zich ooit af. Als fervent theedrinker was ik toch regelmatig in de koffiekamer te vinden. Ik wil de koffiekamerbezoekers bedanken voor hun gezelligheid: Bas, Frank, Frans, Ferry, Frits, Jan-Willem, Jeroen, John, Jurjaan, Kelly, Krista, Margreet, Tania, en alle anderen. Dankzij jullie heb ik een hele leuke tijd op het AZU gehad.

Onder het genot van een Guinness heb ik regelmatig gediscussieerd over allerhande akoestische onderwerpen met de leden van het ATO (Akoestisch Tafel Overleg). *"Guinness for strength"* luidt de slagzin. In welke mate het aan de Guinness lag of aan het gezelschap weet ik niet, maar deze bijeenkomsten waren zonder uitzondering uiterst inspirerend en boeiend. Jongens, bedankt!

Ook ben ik veel dank verschuldigd aan mijn proefpersonen. Ik heb veel respect voor deze mensen die urenlang naar zinnetjes hebben geluisterd in een kleine en vaak warme audiologische cabine. Maar zoals het materiaal zelf vermeldde: *"het werk is niet voor niets geweest"* [1] en *"gelukkig is het eindpunt bijna bereikt"* [1].

Ten slotte wil ik mijn familie en vrienden bedanken voor hun steun en welgemeende interesse. En Mark, heel erg bedankt dat je speciaal voor mij een pinguinpakje aantrekt!

[1] Geparafraseerd uit het voor dit proefschrift gebruikte testmateriaal.

Curriculum Vitae

10 September 1972	Born in Weert, the Netherlands
1984–1990	VWO ('pre-university education'), Philips van Horne Scholengemeenschap, Weert
1990–1996	Physics at Eindhoven University of Technology, Traineeship at IPO, Institute for Perception Research, Eindhoven, Traineeship at RTC Cork, department of Applied Physics, Cork, Ireland, Final research project at Philips Research Laboratories, Eindhoven
1997–2000	Peutz Consulting Engineers, Mook Department of Industrial Noise and Vibration Control
2000–2005	PhD student at Hearing Research Laboratories, University Medical Center Utrecht
2005–now	Academic Medical Center, Amsterdam, Department of Clinical and Experimental Audiology