

This article was downloaded by:[University of Glasgow]
On: 7 April 2008
Access Details: [subscription number 773513295]
Publisher: Informa Healthcare
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Audiology

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713721994>

Discrimination of release time constants in hearing-aid compressors

Gaëtan Gilbert ^a; Michael A. Akeroyd ^a; Stuart Gatehouse ^a

^a MRC Institute of Hearing Research, Glasgow Royal Infirmary, UK

First Published on: 25 February 2008

To cite this Article: Gilbert, Gaëtan, Akeroyd, Michael A. and Gatehouse, Stuart (2008) 'Discrimination of release time constants in hearing-aid compressors', International Journal of Audiology, 47:4, 189 - 198

To link to this article: DOI: 10.1080/14992020701829722

URL: <http://dx.doi.org/10.1080/14992020701829722>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Gaëtan Gilbert
Michael A. Akeroyd
Stuart Gatehouse

MRC Institute of Hearing Research,
Glasgow Royal Infirmary, UK

Key Words

Hearing aids
Compression
Release time
Discrimination

Abbreviations

CT: Compression thresholds (dB)
CR: Compression ratio
 G_{CAM65} : Insertion gain for a speech-
shaped noise at 65 dB SPL (dB)
x: Input level in a compression
channel (dB)
 G_{LD} : Level-dependent gain (dB)
 G_{ID} : Level-independent gain (dB)

Discrimination of release time constants in hearing-aid compressors

Abstract

We examined the ability of twenty-five hearing-impaired and eight normal-hearing listeners to discriminate between release time constants used for compression in hearing aids. The compressor was a standard three-channel system. The stimuli were normal and 'vocoded' sentences from a male and female database. In agreement with other studies looking at different outcomes, performance varied greatly across individuals. This variation was greater in hearing-impaired listeners, for whom the discriminability of a release time of 5 ms from one of 5000 ms (with the attack time fixed at 5 ms) ranged from chance to perfect. This variability was not significantly related to hearing impairment nor to individuals' compression ratios.

Sumario

Examinamos la habilidad de veinticinco hipoacúsicos y ocho normoyentes para discriminar entre la liberación de constantes de tiempo utilizadas para la compresión en auxiliares auditivos. El compresor fue un sistema estándar de tres canales. Los estímulos fueron frases con voces masculinas y femeninas de una base de datos, normales y distorsionadas con *vocoder*. En concordancia con otros estudios el desempeño varió ampliamente entre individuos. Estas variaciones fueron mayores entre los hipoacúsicos para quienes la discriminación de un tiempo de liberación de 5 ms contra uno de 5000 ms (con el tiempo de ataque fijo en 5 ms) varió del azar a la perfección. Esta variabilidad no se relacionó significativamente con la desventaja auditiva ni con los rangos individuales de compresión.

The advantage of using compression—such as wide dynamic-range compression or automatic volume control—in hearing aids is well established when the hearing loss is of cochlear origin, as it provides benefits across a wider range of situations than linear hearing aids (Braida et al, 1982). As compression is non-linear, it modifies the signal (Hickson & Thyer, 2003) to a degree dependent upon the time constants of the compression (Jenstad & Souza, 2005). However, the literature about the effectiveness of the time constants used for the compression is inconclusive. Gatehouse et al (2006) reviewed thirteen studies which contrasted alternative time-constant settings in terms of benefits: four studies found no effect of varying time constants (Bentler & Nelson, 1997; Novick et al, 2001; Moore et al, 2004; Muller et al 2004), three studies found that fast compression systems were superior to slow systems (Nábělek & Robinette, 1977; Jerlvall & Lindblad, 1978; Moore et al, 1993), three studies found that slow compression systems were superior to fast systems (King & Martin, 1984; Neuman et al, 1998; Hansen, 2002), and three studies found mixed results, giving different optima for different categories of listeners (Schweitzer & Causey, 1977; Neuman et al, 1995; van Toor & Verschuure, 2002).

In commercial hearing aids, the range of compression attack times varies from milliseconds to hundreds of milliseconds, and the range of compression release times varies from less than ten milliseconds to tens of seconds. Gatehouse et al (2003) showed that benefits of fast versus slow compression in hearing aids arise in different domains: fast compressors are generally better for speech intelligibility, while slow compressors are generally better

for listening comfort. The experimental studies cited above are characterized by a wide variation in the types of outcome measures across studies: objective speech intelligibility, self-rated speech intelligibility, sound quality ratings, or preference judgements. This inconsistency and heterogeneity of the outcome domains leads to a lack of guidance about an *optimal* setting of the time constants: while current hearing-aid prescription rationales define some of the fitting parameters, such as frequency-gain characteristic and the amount of compression, there is no consensus about how to set the time constants.

Another aspect of the literature about time constants in hearing aids is the variability among individuals. The inter-individual variability among hearing-impaired listeners has been observed in preference judgements of the quality of compressed speech (Neuman et al, 1995), in self-rated and objective speech intelligibility scores (van Toor & Verschuure, 2002; Gatehouse et al, 2006), and in questionnaires assessing the benefits that hearing-impaired listeners obtain from their hearing aids (van Toor & Verschuure, 2002; Gatehouse et al, 2006). The recurrent variability among hearing-impaired listeners has led some investigators (Neuman et al, 1995, Gatehouse et al, 2006) to suggest that the fitting of the time constants in a compression hearing aid might be individualized.

Nábělek (1984) studied the ability of listeners to discriminate between a single-channel compressed and uncompressed sentence or between various single-channel compressed sentences. In his first experiment, discrimination was measured for various combinations of attack times (3, 10, and 30 ms), compression

ratios (2.5:1 and 5:1), and various levels of a speech-shaped noise added after compression (signal-to-noise ratios of 5 and 10 dB). This experiment was conducted with normal-hearing listeners. Despite large individual differences, the obtained results suggested that the discrimination score was higher for larger compression ratios, longer release times, shorter attack times, and the more trained the listener was. In a second experiment, the discriminability of various compression speeds (attack and release time constants) was examined, again with normal-hearing listeners. For a compression ratio of 2.5:1 in a single-channel compressor, the average performance was 62%, 67%, and 72% for discriminating attack/release times of 1/30 ms from 10/90 ms, 1/10 ms from 1/30 ms, and 1/10 ms from 10/90 ms, respectively. A third experiment demonstrated variability across listeners with sensorineural hearing loss. Listeners were asked to distinguish between an uncompressed sentence and a single-channel, strongly compressed sentence (compression ratio = 10:1) for attack/release times of 10/90 ms and 1/10 ms. Performance across individuals ranged between chance and perfect.

The present study investigated the ability of normal and hearing-impaired listeners to discriminate between different release times of compression, regardless of any associated benefits or drawbacks, using a simulation of a three-channel, ten-band compressor. Note that we use the term *channels* to refer to the frequency ranges used for the compression stage, and the term *bands* to refer to the frequency ranges used for the frequency-response shaping stage. Nábělek (1984) noted that compression affected the loudness and timbre of a sound, as well as increasing any noise in the recordings and creating overshoots. In his experiment measuring the discriminability of time constants he concentrated on the resulting changes in sound quality (short-term loudness differences and spectral changes). We took a similar approach here, and so attempted to minimize the difference of long-term loudness, the enhancement of the recording noise with fast compressors, and the difference in overshoots (see Method for further details). We used a larger range of release time constants than Nábělek, making it possible to generate psychometric functions for compression discrimination. We also used a lower compression ratio: Nábělek's compression ratios (up to 10:1) would not be prescribed today, whereas the present study used a maximum compression ratio of 2.7:1. Our compression ratios were determined from an individualized fitting using CAMEQ (Moore et al, 1999).

In addition to normal sentences, we also used 'vocoded' sentences, which were processed to remove the spectral fine detail and temporal fine structure whilst preserving the envelope information (Shannon et al, 1995). Van Tasell and Trine (1996) showed that the envelope of a sentence was one feature adversely affected by compression. With the use of vocoded sentences, we assumed that listeners would put more attention on the remaining envelope information. As changing the release time for compression modifies the envelope, we expected better discrimination of release time with these stimuli. Another difference between normal and vocoded items is that the intelligibility of vocoded nonsense syllables is related to modification of the intensity difference between a consonant and a vowel (Balakrishnan et al, 1996; Freyman et al, 1991) while this is not observed with natural stimuli (Jenstad & Souza, 2005). We

used a five-channel vocoder, which would be expected to give near perfect sentence identification (Shannon et al, 2004).

Method

Listeners

Twenty-five hearing-impaired listeners participated. They were all suitable candidates for compression hearing aids, and were aged between 51 and 76 years old (mean 66.8 years; standard deviation 6.5 years). The hearing-impaired listeners were selected so that at least one of the three compression ratios prescribed by CAMEQ (Moore et al, 1999) was greater than or equal to 2:1. A minimum compression ratio of 1.4:1 was imposed in all three channels; this was done to allow the attack and release time constants to be measured according to ANSI S3.22 (1996). For nineteen of the listeners their poorer ear was chosen, but for the other six listeners, whose poorer ear required gains greater than our equipment supported, their better ear was chosen. The individual and the mean audiograms are shown in Figure 1 (thin solid lines, and a short-dashed bold line, respectively). Eighteen listeners were diagnosed with a purely sensorineural hearing loss on the basis of air-bone gaps less than or equal to 10 dB at frequencies of 0.5, 1.0, 2.0, and 4.0 kHz. The remaining seven hearing-impaired listeners had at least one air-bone gap above 10 dB at those frequencies. The contribution of the conductive loss was assessed by calculating the ratio of the air-bone gaps (in dB) to the pure-tone hearing thresholds (in dB) at frequencies of 0.5, 1.0, 2.0, and 4.0 kHz; the average values across the seven listeners were 0.24, 0.36, 0.15, and 0.09, respectively (with *maximum* values of 0.47, 0.54, 0.22, and 0.20); the average across all frequencies was 0.21. These values suggest that the contribution of the conductive loss was relatively small.

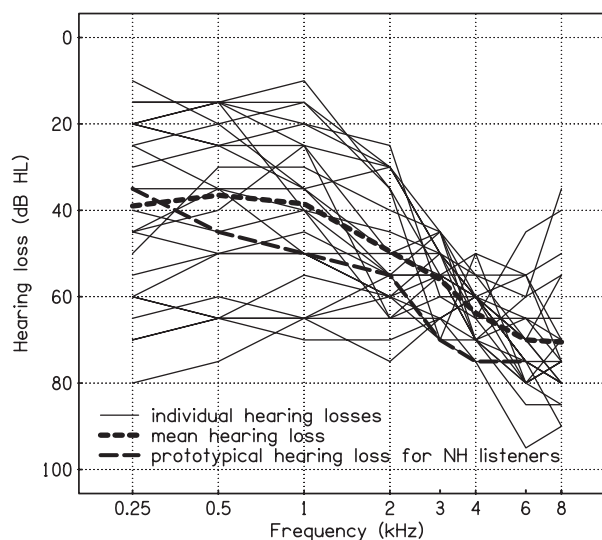


Figure 1. Pure-tone hearing thresholds for the twenty-five hearing-impaired listeners who took part in the study. The individual audiograms are shown by the thin solid lines, and the mean by the short-dashed bold line. The long-dashed bold line shows the prototypical moderate sloping hearing loss that was used to obtain compression parameters for the normal-hearing listeners.

Eight normal-hearing listeners also participated. They were aged between 23 and 35 years old (mean 28.1 years; standard deviation 4.3 years). The normal-hearing listeners had pure-tone thresholds below 20 dB HL for audiometric frequencies between 0.25 and 8.0 kHz. The stimuli were presented to their right ear. In order to simulate a hearing loss for these listeners, a prototypical moderate, sloping hearing loss was defined (shown by the long-dashed bold line in Figure 1). The compressed-amplified sentences were presented at a comfortable level to normal ears by attenuating them by half the prototypical hearing loss using a FIR digital filter.

Compression

The compressor was a Fast-Fourier-Transform (FFT) -based, three-channel, ten-band system, implemented as a series of programs in Matlab. The FFT frequencies included in each channel were 0–861 Hz, 948–2412 Hz, and 2498–7492 Hz; the FFT frequencies included in band are reported in Table 1. Each sentence was first sliced into 4-ms (176 points) frames, overlapping by 2 ms. Each frame was then multiplied by a Hanning window, padded with 168 zeros on both sides and then Fast-Fourier-transformed, giving 512 components spaced 86.1 Hz apart. The level of each component was then amplified by 8.9 dB to compensate for the loss of power due to the windowing and the zero padding; the phase was kept fixed.

Next, the level in each channel was determined by summing the levels across each FFT frequency within that channel. These levels were smoothed across frames using a single-pole lowpass filter whose characteristics were set to give the desired attack/release times: if the level in one frame was greater than the level in the previous frame, then the time constant of the filter was determined by the desired attack time; if it was lower, then the time constant was determined by the release time (these time constants depend on both the compression ratio and the desired attack-release times; they were taken from a look-up table determined in advance following the ANSI S3.22 standard). This smoothing was done for each channel independently.

Figure 2 shows a schematic Input/Output function and illustrates all the variables used below. The smoothed levels determined the *total* amount of gain applied, defined as the sum of a ‘level-dependent’ (compressive) gain G_{LD} , and a ‘level-independent’ (static) gain G_{ID} :

Table 1. Compression channels and frequency shaping bands characterizing the simulation of the compression hearing aid used in this study. Note that the FFT frequency resolution was 86.1 Hz.

Channel number	Band number	FFT bins (Hz)
1	1	0–172
	2	258–345
	3	431–603
	4	689–861
2	5	948–1206
	6	292–1723
	7	1809–2412
3	8	2498–3445
	9	3531–4823
	10	4909–7942

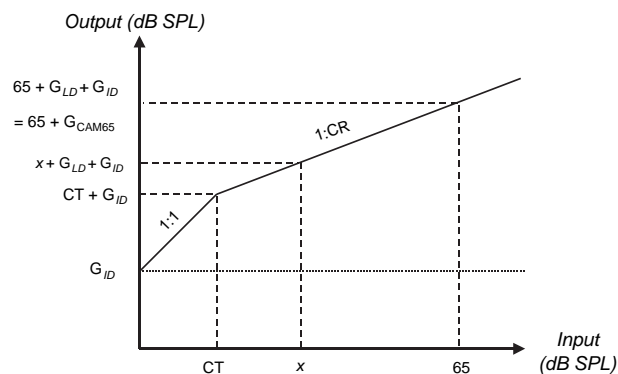


Figure 2. Schematic input/output function of a compressor showing the variables used in the Method section.

$$G_{LD} = (x - CT)(CR^{-1} - 1)$$

$$G_{ID} = (65 + G_{CAM65} - CT) - (65 - CT)CR^{-1}$$

where x is the input level in decibels, CT is the compression threshold (decibels), CR is the compression ratio, and G_{CAM65} is the speech-shaped-noise insertion gain (decibels). Note that the ‘level-independent’ gain is the same as the *total* gain for a signal at the channel compression threshold: i.e. for $x = CT$:

$$G_{LD} = (CT - CT)(CR^{-1} - 1)$$

Therefore, total gain = $G_{ID} + 0 = G_{ID}$

The value of x was the smoothed level for each channel (summed across each FFT bin in the channel) and frame: for bands 1, 2, 3, and 4 it was equal to the smoothed level in the low channel; for bands 5, 6, and 7 it was equal to the smoothed level in the middle channel; and for bands 8, 9, and 10 it was equal to the smoothed level in the high channel. Note that if x was less than the compression threshold, then the level-dependent gain was set to 0 (the level-independent gain was unaffected). The values of CT and CR were also defined on a per-channel basis and then copied into the relevant bands. The value of CT for bands 1, 2, 3, and 4 (low channel) was 40 dB; for bands 5, 6, and 7 (middle channel) it was 32 dB; and for bands 8, 9, and 10 (high channel) it was 25 dB. These values were approximately 25 dB below the long-term power of the sentences in each channel, so ensuring that almost no part of any sentence would fall outside the range of compression. The value of CR for bands 1, 2, 3, and 4 (low channel), for bands 5, 6, and 7 (middle channel), and for bands 8, 9, and 10 (high channel) were those prescribed for the three channels by CAMEQ. The value of G_{CAM65} was the only one to be defined on a band basis: this was the speech-shaped-noise insertion gain prescribed by CAMEQ at 65 dB SPL. Note that the gain prescribed by CAMEQ (G_{CAM65}) at 6000 Hz was mistakenly applied at 5000 Hz for all the normal-hearing listeners and for eight of the hearing-impaired listeners, resulting in the following frequency bins (in Hz) for bands 9 and 10, respectively: 3531–4393 Hz and 4479–7942 Hz.

The values of compression ratios and insertion gains were found for each listener individually using the CAMEQ procedure (Moore et al, 1999), which is part of the CAMFIT software

package. Its design goals are to deliver a specific loudness pattern that gives both equal loudness per critical band for the speech frequencies while also giving an overall loudness that is approximately normal for, in our case, an input level of 65 dB SPL. For the compressor we used—which was set-up and calibrated using sine waves—it would have been more usual to use the insertion gains prescribed for a sinusoid at 65 dB SPL instead of the insertion gains prescribed for a speech-shaped noise. We did not expect this, however, to have had a substantial effect on the experimental design, as the values of the insertion gains for the noise signal were almost the same (or slightly higher) than the values for sinusoidal signal, and so would only have led to an increase in audibility (the differences were, on average, -1.3 , 5.2 , and 9.8 dB for the three channels, and arise from a combination of the differences in the speech spectral shape and the width of the channels).

The final steps were to calculate the *total* gains at each of the 512 FFT frequencies by, for instance, giving each frequency bin between 0 and 172 Hz the gain determined for the first band (see Table 1). These were then smoothed across frequency to avoid the step changes from one band to another which would otherwise have given rise to excessive oscillations in the time domain. This was done by convolving the 512 gain values twice with a symmetric five-point exponential window. Note that the convolution was applied to the set of amplitudes for each of the FFT bins: thus the points in the smoothing window were 86.1 Hz apart, and so the maximum extent of the smoothing was ± 344 Hz. The magnitude at each of the 512 FFT frequencies was then amplified by this gain and then attenuated by 8.9 dB, in order to undo the previous correction for the loss of power due to Hanning windowing and zero padding. An inverse FFT of the resulting magnitude and phase spectrum gave the final 512-point (4-ms) waveform for this frame. All these waveforms were then combined using the overlap-and-add method (Allen, 1977). Note that the values of the five-point smoothing window were [0.12, 0.32, 0.87, 0.32, 0.12], which is a symmetric exponential function but normalized to give a total squared power of 1.0. The double convolution resulted in an effective convolution with [0.0098, 0.054, 0.22, 0.45, 0.70, 0.45, 0.22, 0.054, 0.0098].

Stimuli and apparatus

The 'normal' sentence stimuli were taken from the 270 'ASL' sentences (MacLeod & Summerfield, 1990) and the 336 'BKB' sentences (Bench & Bamford, 1979). Each sentence was about 1.5 seconds in duration and had a simple declarative structure (e.g. 'The ice cream was pink') with, usually, about five words. The ASL sentences were spoken by a single male speaker and the BKB sentences by a single female speaker. The RMS powers of the sentences varied slightly, as they were spoken naturally with minor variations in effort; the mean levels were 68.6 dB SPL and 69.2 dB SPL for the ASL and BKB sentences, respectively (the standard deviations were 1.2 and 1.4 dB). Each sentence was low-pass filtered at 8 kHz and the whole model ran at 44.1 kHz. The recordings were standard sets, often used in studies measuring speech intelligibility.

The 'vocoded' stimuli were derived from the same set of sentences using a five-channel noise vocoder. First, each sentence was filtered using a five-channel filterbank (the filters were sixth-order Butterworth bandpass filters with passbands of 50–354 Hz, 354–866 Hz, 866–1732 Hz, 1732–3464 Hz, and 3464–

8000 Hz). Second, the amplitude envelopes were found by applying a Hilbert transform, and then smoothed using a low-pass filter (first-order Butterworth, 32-Hz cutoff frequency). Third, these were multiplied by five independent and spectrally flat noises (0–22 kHz). In order to take into account the variation of the power with the bandwidth, each flat noise was scaled by the square root of the bandwidth before passing through the appropriate filter of the five-channel filterbank. The resulting spectrally limited noises were summed to create the final stimuli, and approximately time-aligned to account for the frequency-dependent group delay of the initial sixth-order filterbank.

Each sentence was compressed using each of four release times: 'very fast', 'fast', 'slow' and 'very slow', characterized by attack/release times of 5/5-ms, 5/50-ms, 5/500-ms, and 5/5000-ms respectively.

As in Nábělek's (1984) study, varying the speed of compression led to potential cues that were not the focus of the present study. Accordingly, we attempted to minimize the following three effects.

First, the overall loudness of the sentences varied, with shorter compression release times leading to louder sentences. To remove this effect, we attenuated or amplified each sentence until its loudness was equal to a reference loudness. This was accomplished by first determining the average spectrum of the sentence and using that to calculate the loudness (using the model developed by Moore et al, 1997). The average loudness across all the sentences of the set was taken as the reference loudness. Finally an iterative procedure was used to determine how much to attenuate or amplify each sentence by until its own loudness matched the reference loudness. It was on average, -2.5 dB, -1 dB, 1 dB, and 3 dB for sentences processed through compressors with attack/release time constants of 5/5-ms, 5/50-ms, 5/500-ms, and 5/5000-ms, respectively. Note that this gain was independent of frequency and level.

Second, although the sentences were recorded in a quiet, sound-treated environment, they contained a minor amount of recording noise. A fast release compression system might allow this recording noise to become audible in silent intervals. Accordingly, to prevent the listeners from using this as a cue, it was masked by a steady noise applied after compression. This noise was spectrally flat from 0 to 22 kHz and was continuous across the three intervals of each trial, starting 500 ms before the first sentence and ending 500 ms after the third sentence. For the hearing-impaired listeners, it was presented at 70 dB SPL (a spectrum level of 27 dB SPL), giving long-term-average signal-to-noise ratios of between $+20$ and $+30$ dB (the variation is due to each listener having an individualized gain function). For the normal-hearing listeners, it was set to an overall level of 40 dB SPL, yielding a long-term-average signal-to-noise ratio of around $+15$ dB (note that it was added *after* the converse half-gain rule was applied). The level of the masking noise added after compression was dependent upon the amount of recording noise contained in the 'ASL-normal' stimuli (see below) processed with the 5/5-ms compressor set for the prototypical hearing loss shown in Figure 1. The amount of recording noise was estimated by first calculating a 'near-instantaneous' power across the duration of each sentence (defined as the power in a series of 10-ms, 50% overlapping segments) and taking the minimum value per sentence. The level of the masking noise was

set to be approximately 3 dB above the highest point of the resulting distribution of those minimum values.

Third, in order to set the compressor into a pseudo-continuous state and to avoid any onset effects, every sentence had four ‘warm-up’ sentences put in front of it before the compression was applied, but which were removed after compression had been applied. Their total duration was 5.7 seconds, which was longer than the longest release time used.

The presentation of the stimuli and acquisition of the responses were controlled using a Matlab program running on a personal computer. The normal-hearing listeners were tested in an IAC acoustic booth using Sennheiser HD 580 circumaural headphones, after 16-bit D/A conversion by a RME DIGI 96/8 Pad audio interface and amplification by an Arcam A80 audio amplifier. The hearing-impaired listeners were tested monaurally in a quiet room using the same headphones as normal-hearing listeners, after 16-bit D/A conversion by an EDIROL Audio Capture UA-5 audio interface and amplification by a SAMSON C.que 8 headphone amplifier.

Experimental design

The task was to discriminate a reference sentence (5/5-ms compression) from a target sentence (5/50-ms, or 5/500-ms, or 5/5000-ms compression). The three release time comparisons are termed ‘5_vs_50’, ‘5_vs_500’, and ‘5_vs_5000’, respectively. The performance in these three conditions defined a psychometric function for the discriminability of release time. A three-interval, two-alternative, forced-choice task was used, for which the trial structure was either *reference-reference-target* or *reference-target-reference* and in which listeners were required to decide which of the last two intervals sounded different to the first one. A psychometric function was measured for each of the four combinations of stimuli: ‘ASL-normal’, ‘BKB-normal’, ‘ASL-vocoded’, and ‘BKB-vocoded’.

The experiment was conducted across three sessions. The first session was devoted to training. The listeners were first familiarized with vocoded speech by presenting, successively, 20-band, 10-band, and five-band vocoded sentences. Next the listeners were familiarized with the experimental method by presenting them with stimuli from the 5_vs_5000 condition, and then stimuli from the other two conditions were introduced. Listeners were given feedback during all the training sessions. The data were collected during the second and third visits. Each visit consisted of six blocks of 60 trials. Blocks using normal sentences alternated with blocks using five-band vocoded sentences. Within a block, 30 trials (10 trials for each of the three conditions) used ASL sentences and 30 trials used the BKB sentences. No feedback was given to the listeners during an experimental session. The sentence to be used in a trial was chosen at random from the 270 choices for the ASL set and from the 336 choices for the BKB set. Overall, there were 60 trials per listener for each point of every three-point psychometric function.

Results

Performance of hearing-impaired listeners

Each panel of Figure 3 shows discrimination scores as a function of the release time constant of the target. The different symbols represent the sentence type (circles = ASL, squares = BKB) and

the processing condition (filled symbols = normal sentences, open symbols = vocoded sentences). The dashed lines show the chance level (50%) and an arbitrary threshold level (75%, so approximately corresponding to a d' of 1). The bottom-right panel shows discrimination scores averaged across the twenty-five hearing-impaired listeners. The other panels show the results from representative individual listeners.

In general, performance improved with the release time constant of the target. A repeated-measures ANOVA with the set of sentences (BKB or ASL), the type of processing (normal or vocoded), and the release-time condition (5_vs_50, 5_vs_500, and 5_vs_5000) as factors confirmed this observation, giving a significant main effect of release time ($F(2,48) = 46.69$, $p < 0.001$). It also showed a marginally significant difference between the two sets of sentences ($F(1,24) = 4.64$, $p = 0.04$); the scores obtained with the BKB sentences were slightly higher than with the ASL sentences. The other main effects and interactions were not significant. Note that the Greenhouse-Geisser statistics were used when sphericity did not occur. We report the corrected p -values, but the uncorrected degrees of freedom.

An effect of the release time upon performance was not, however, observed for *all* individuals; some of the hearing-impaired listeners showed little or no ability to discriminate release time. This can be seen for the three individual-listener panels of Figure 3: one is for a ‘good’ listener (HI-1, upper-left), one for a ‘middle’ listener (HI-14, upper-right), and one for a ‘poor’ listener (HI-23, bottom-left). Listener HI-1 performed near threshold in the 5_vs_50 condition and at ceiling in the 5_vs_500 and 5_vs_5000 conditions, while listener HI-23 performed approximately at chance in all three conditions. Listener HI-14 showed similar results to the overall mean.

To study this further, we calculated for each listener the average score across all four stimulus types for each of the 5_vs_50, 5_vs_500, and 5_vs_5000 conditions. The results are shown in Figure 4, with the listeners ranked by their average score in the 5_vs_5000 condition (bottom panel). In that condition, listeners fell along a continuum, from chance to perfect. The top and middle panels show the corresponding results for the 5_vs_50 and 5_vs_500 conditions, respectively, using the same ranking of listeners: in general, a listener who performed relatively well in the 5_vs_5000 condition also performed well in the 5_vs_500 condition (the correlation between scores for the two was $r = +0.97$), and some of them also performed well in the 5_vs_50 condition ($r = +0.81$). In the 5_vs_50 condition (the hardest that we tested) three of the listeners performed at, or very close to, the 75% threshold, whilst in the 5_vs_5000 condition (the easiest) there were at least four listeners whose performance was approximately at chance.

Dependence on hearing loss and compression ratio

Figure 5 shows a scatter plot of the average performance in the 5_vs_5000 condition as a function of an individual’s hearing loss, calculated as the average of the pure-tone hearing threshold levels at 0.5, 1.0, 2.0, and 4.0 kHz for the test ear. The two measures were not significantly related: the correlation coefficient was 0.24. Also, the correlation coefficients between the compression ratios in each channel and the average performance were, respectively, 0.23, 0.20, and 0.11; all were statistically insignificant.

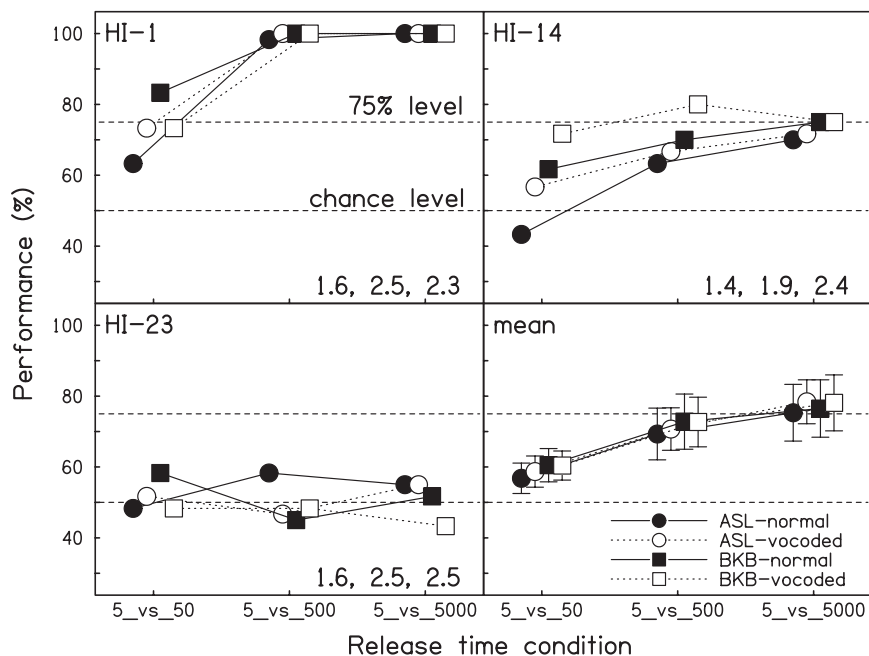


Figure 3. Each panel shows performance as a function of the release time constant of the target stimuli. The different symbols represent the sentence type (circles = ASL, squares = BKB), and the processing condition (filled symbols = normal, open symbols = vocoded). The chance level (50%) and an arbitrary threshold (75%) are represented with dotted lines. The upper-left, upper-right, and bottom-left panels show individual scores for a good, medium, and poor performer, respectively. The bottom-right panel shows discrimination scores averaged across the twenty-five hearing-impaired listeners (the error bars are 95% confidence intervals). The numbers in the individual-listener panels are compression ratios used for that individual in the low, mid, and high compression channels, respectively.

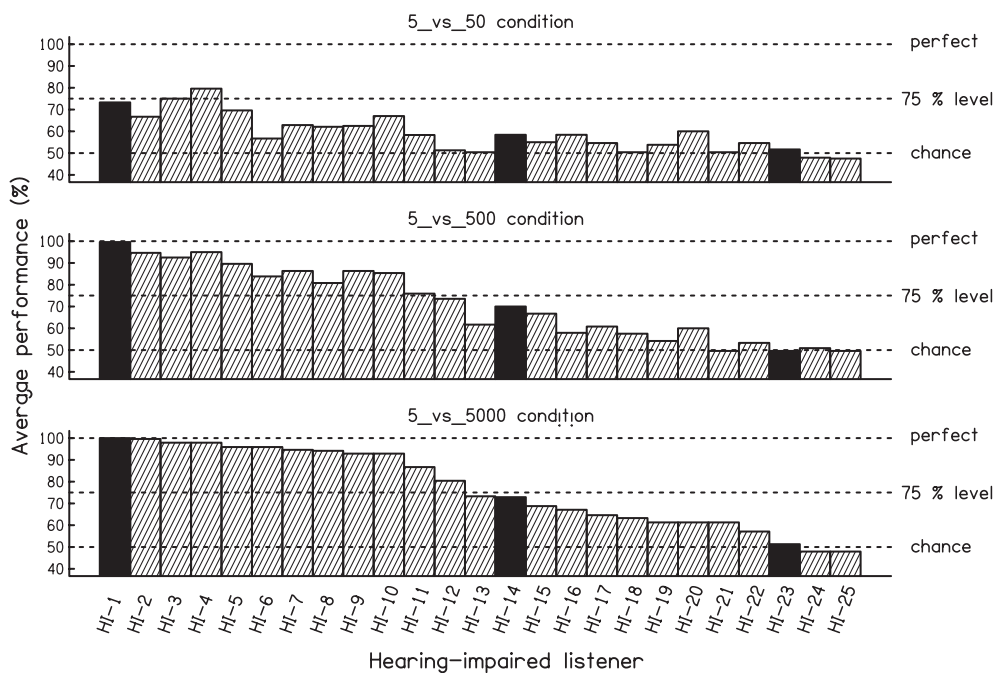


Figure 4. Average performance in the 5_vs_50 condition (top panel), 5_vs_500 condition (middle panel), and 5_vs_5000 condition (bottom panel) for the hearing-impaired listeners. The listeners are ranked according to their 5_vs_5000 results. The average performance was computed across the four experimental stimuli of ‘ASL-normal’, ‘BKB-normal’, ‘ASLvocoded’, and ‘BKB-vocoded’. The solid bars show the scores of the representative listeners that have been selected in the previous Figure (Figure 3).

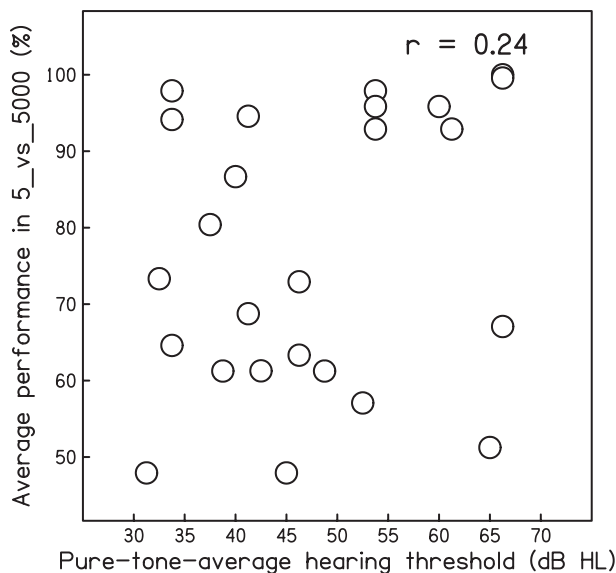


Figure 5. Scatter plot of the average performance in the 5_vs_5000 condition versus hearing level in the test ear. The average performance was computed across the four experimental stimuli of 'ASL-normal', 'BKB-normal', 'ASL-vocoded', and 'BKB-vocoded'.

Performance of normal-hearing listeners

Figure 6 shows the data from the eight normal-hearing listeners, using a similar format to Figure 3. The upper-left and lower-left panels are for three representative listeners, while the lower-right panel is for the across-listener mean. There, discrimination

increased with the release time of the target: an ANOVA showed a significant effect of release time ($F(2,14) = 133.24, p < 0.001$). Again, the scores obtained with the BKB sentences were slightly higher than with the ASL sentences ($F(1,7) = 7.16, p = 0.03$), but, unlike the scores for hearing-impaired listeners, there was also a significant interaction between the set of sentences and the release-time conditions ($F(2,14) = 6.88, p < 0.01$). Inspection of the results suggests that the better scores for BKB sentences only occurred in the 5_vs_500 condition.

The individual-listener panels of Figure 6 show data from a representative 'good' listener (NH-1), a 'medium' listener (NH-5), and a 'poor' listener (NH-8): as with the hearing-impaired listeners, there was a substantial variation across listeners. Figure 7 shows the results of a similar ranking method to before, but applied to the normally-hearing listeners. Although the number of listeners is smaller, the results suggest that the variability across normal-hearing listeners was less than with hearing-impaired listeners, in that none performed at, or near, chance in the 5_vs_5000 condition (bottom panel).

There was a small difference in the overall scores between the two groups: the mean scores for the normal and impaired listeners were 78% and 71% for the 5_vs_500 condition, and 90% and 77% for the 5_vs_5000 condition. This effect was, however, small and it should be noted that neither did the prototypical hearing loss (used to set the compressor for the normal-hearing listeners) exactly match the mean hearing loss of the hearing-impaired listeners nor was the number of listeners in the two groups equal. Nevertheless, the result is intriguing and would merit further study.

In the 5_vs_50 experimental conditions discrimination was higher for the best hearing-impaired listeners than for the best normal-hearing listeners. Although we cannot be certain, we

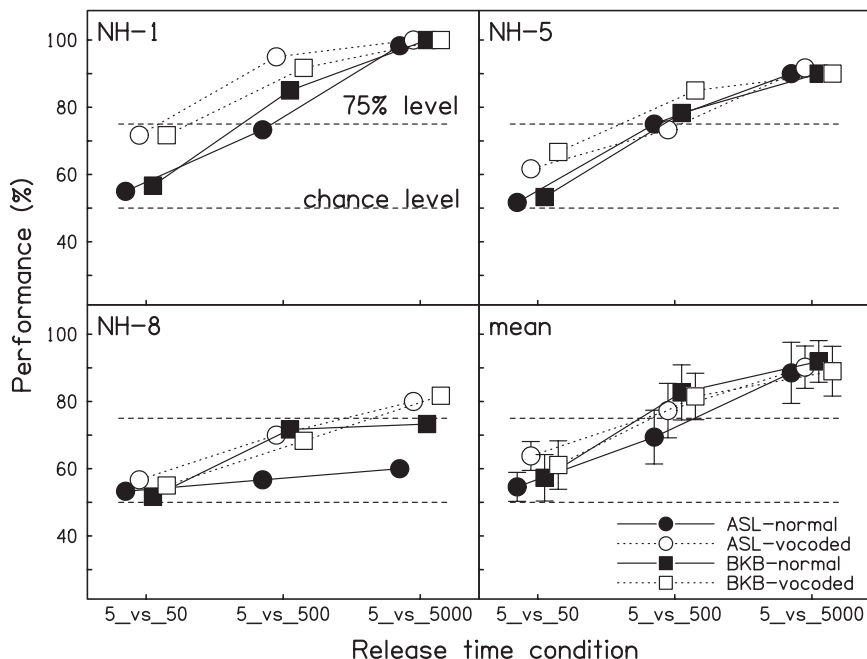


Figure 6. As Figure 3, but for the eight normal-hearing listeners.

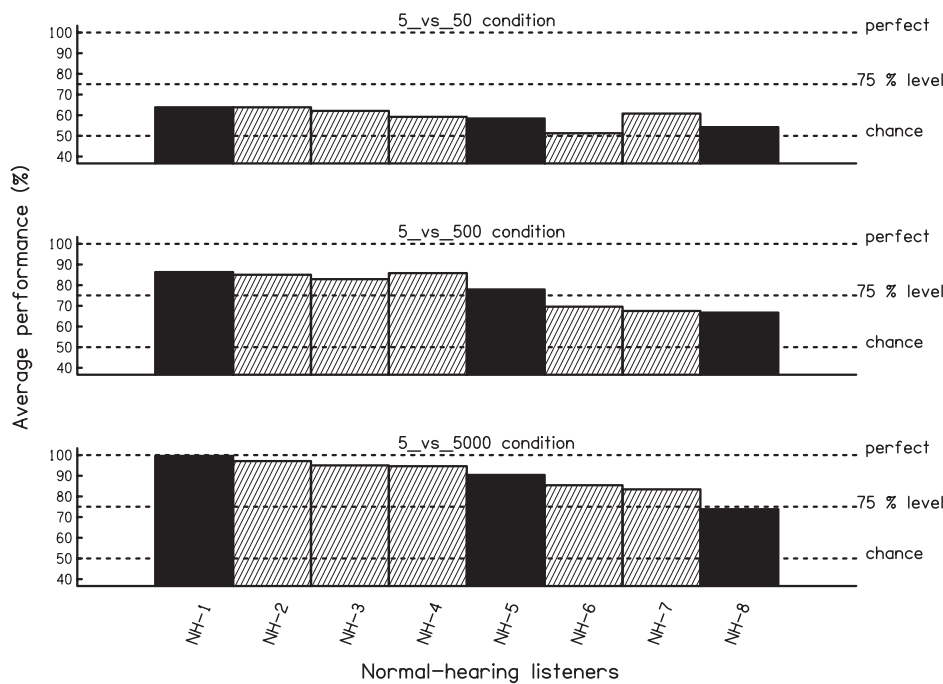


Figure 7. As Figure 4, but for the eight normal-hearing listeners. The solid bars show the scores of the representative listeners that have been selected in the previous Figure (Figure 6).

believe that this may be due to the reduction of the dynamic range of hearing-impaired listeners and a corresponding greater variation of the sensation of loudness (Moore, 1995).

Discussion

Inter-individual variability

This study investigated the ability of normal and hearing-impaired listeners to discriminate between the release time constants of compression. The results showed that this ability was highly variable across individuals; even for the easiest condition that we tested (5 ms vs 5000 ms) some hearing-impaired listeners performed at chance, whilst others performed perfectly. This variation parallels other studies that have observed large individual variation in the effects of compression, from studies of quality (Neuman et al, 1995), self-rated and objective speech intelligibility (van Toor & Verschuure, 2002; Gatehouse et al, 2006), to self-reported assessment of hearing-aid benefits (Gatehouse et al, 2006). It is possible that the variation may be related to differences in cognitive function. Lunner (2003) asked hearing-impaired listeners with low and high cognitive abilities to report the benefits they experienced in various situations. He compared two types of setting, a reference (standard) hearing-aid setting and another more complex setting for which signals were processed differently in noisy situations when speech sounds were detected. He found that, in noisy situations, the listeners with the highest cognitive abilities reported larger benefits from the more complex setting than the listeners with low cognitive abilities. Also, Gatehouse et al (2003) showed that listeners with high cognitive ability derived greater benefit from listening to speech in a background

noise using fast compression than listeners with low cognitive ability.

Our results showed that there were a substantial number of hearing-impaired listeners who could not reliably discriminate release time constants of either 500 ms or 5000 ms from the reference value of 5 ms ($N=13$ and 12 , respectively; see Figure 3). If these results can be generalized then they would suggest that there are some listeners to whom it would not matter what release time constant they are fitted with. However, we cannot rule out the possibility that those listeners can benefit from an undetectable change in release time in other domains. In contrast, other listeners *can*, sometimes remarkably well, discriminate such time constants. For them, there may be an optimal setting of the time constants, and it may become necessary to include a consideration of the time constant in the fitting of their hearing aids.

Relation to hearing sensitivity and compression

Performance in the 5_vs_5000 condition did not correlate significantly with individual hearing losses in the tested ears. The correlation coefficients with the individually prescribed compression ratios were also insignificant. This suggests that the ability to differentiate between two release time constants does not depend upon the severity of the hearing loss or upon compression ratios below or equal to 2.7:1 prescribed individually with a standard fitting method.

Stimulus effects

There were no significant differences between the psychometric function obtained with normal and vocoded sentences. This was surprising, as the use of vocoded sentences forces listeners to rely

on the remaining envelope information, which is one of the features most affected by compression. We had assumed that the lack of spectral or temporal fine-structure information in the vocoded sentences would lead to a perceptual emphasis on the remaining envelope information and thus an improvement in performance. The results showed that this did not happen. Instead, it would appear that the effectiveness of the envelope information in allowing discriminability of release time constants is independent of the presence or absence of the fine-structure information (note that the envelopes of vocoded sentences, after all, are similar to those of natural sentences).

In many of the conditions, the discrimination scores obtained with the BKB stimuli were generally slightly higher than with the ASL stimuli. The two sets differed in the nature of the speaker: the ASL set was spoken by a man, the BKB set by a woman. They also differed slightly in mean level, by 0.6 dB. However, we suggest a more-informative explanation is the difference in the recording noise (which should not be confused with the noise added after compression) between the ASL and the BKB sets. This difference will be affected by the compression speed. To quantify this difference, the level of the recording noise was estimated using the minimum 'near-instantaneous' power of the series of 10-ms, 50% overlapping segments that was used earlier (see Method). Across all the BKB sentences, its mean was found to be 38 dB SPL, but across the ASL sentences it was 27 dB SPL. This analysis suggests the difference in recording noises between the sentence sets was of the order of 10 dB. Analogous analyses, this time with compressed sentences, confirmed the effect of compression speed, in that the recording noise level was always higher with the 5/5-ms compressor than with the 5/5000-ms compressor. This effect was stronger in the high-frequency channel than in the low-frequency channel. For the ASL sentences, our measure of the recording noise level gave values approximately 20 dB below, 10 dB below, and 3 dB below the level of the noise added after compression in the low-, mid-, and high-frequency channels respectively, yet for the BKB sentences the corresponding values were 10 dB below, 5 dB above, and 15 dB above (all measured with the 5/5-ms compressor). This difference in the recording noise level for discriminating release times may explain the slightly better scores with the BKB sentences observed in this study.

Summary

We measured the ability of listeners to discriminate the release time of a 3-channel, 10-band compressor. The attack time was fixed at 5 ms; the compression ratios and gains were chosen for each individual using CAMEQ (Moore et al, 1999). The compression ratios varied between 1.4:1 and 2.1:1 in the low channel, 1.6:1 and 2.7:1 in the middle channel, and 1.9:1 and 2.5:1 in the high channel. The task was to discriminate between release times of 5 and 50 ms, 5 and 500 ms, and 5 and 5000 ms. The average discriminability scores for twenty-five hearing-impaired listeners were, respectively, 59%, 71%, and 77%; and for eight normal-hearing listeners, given a simulation of a prototypical moderate-sloping loss, were, respectively, 59%, 78%, and 90%. There was a substantial variation across listeners: in the group of hearing-impaired listeners, discrimination ranged

from chance to perfect even for the 'easiest' condition (5 vs. 5000 ms). This variation was not related to the hearing impairment or the compression ratios prescribed for each individual.

Acknowledgements

We wish to thank both Michael Stone (Department of Experimental Psychology, University of Cambridge) and Graham Naylor (Oticon A/S Research Centre, Eriksholm) for their helpful comments and advice on compression, Brian Moore (Associate Editor) and the two anonymous reviewers for their constructive comments on an earlier version of this manuscript, and all the listeners who generously gave their time. Stuart Gatehouse contributed substantially to all of this work before his death in February 2007. Part of this work was presented at the BSA Short Papers Meeting on Experimental Studies of Hearing and Deafness, September 2005 (University of Cardiff, UK).

References

- Allen, J.B. 1977. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Trans Acoust, Speech. Signal Processing*, 25, 235–238.
- American National Standards Institute. 1996. Specification of hearing aid characteristics. ANSI S3.22, New York.
- Balakrishnan, U., Freyman, R.L., Chiang, Y.C., Nerbonne, G.P. & Shea, K.J. 1996. Consonant recognition for spectrally degraded speech as a function of consonant-vowel intensity ratio. *J Acoust Soc Am*, 99, 3758–3769.
- Bench J. & Bamford J. 1979. *Speech-hearing Tests and the Spoken Language of Hearing-impaired Children*. London: Academic Press.
- Bentler, R. & Nelson, J. 1997. Assessing release time options in a two channel AGC hearing aid. *J Am Acad Audiol*, 6, 43–51.
- Braida, L.D., Durlach, N.I., de Gennaro, S.V. & Bustamante, D.K. 1982. "Review of recent research on multiband amplitude compression for the hearing impaired". In G.A. Studebaker & F.H. Bess (eds.) *The Vanderbilt Hearing Aid Report*. PA: Upper Darby, pp. 133–140.
- Freyman, R.L., Nerbonne, G.P. & Cote, H.A. 1991. Effect of consonant-vowel ratio on amplitude envelope cues for consonant recognition. *J Speech Hear Res*, 34, 415–426.
- Gatehouse, S., Naylor, G. & Elberling, C. 2003. Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol*, 42, 77–85.
- Gatehouse, S., Naylor, G. & Elberling, C. 2006. Linear and nonlinear hearing aid fittings: Patterns of benefit. *Int J Audiol*, 45, 130–152.
- Hansen, M. 2002. Effects of multichannel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear Hear*, 23, 369–380.
- Hickson, L. & Thyer, N. 2003. Acoustic analyses of speech through a hearing aid: Perceptual effects of changes with two-channel compression. *J Am Acad Audiol*, 14, 414–426.
- Jenstad, L.M. & Souza, P.M. 2005. Quantifying the effect of compression hearing-aid release time on speech acoustics and intelligibility. *J Speech Lang Hear Res*, 48, 651–667.
- Jerlval, L.B. & Lindblad, A.C. 1978. The influence of attack and release time on speech intelligibility. A study on the effects on normal-hearing and hearing-impaired listeners. *Scand Audiol Suppl*, 6, 341–353.
- King, A.B. & Martin, M.C. 1984. Is AGC beneficial in hearing aids? *Br J Audiol*, 18, 31–38.
- Lunner, T. 2003. Cognitive function in relation to hearing aid use. *Int J Audiol*, 42, 49–58.
- MacLeod, A. & Summerfield, Q. 1990. A procedure for measuring auditory and audio-visual speech reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *Br J Audiol*, 24, 29–43.

- Moore B.C.J. 1995. *Perceptual Consequences of Cochlear Damage*. New York, NY: Oxford University Press, pp. 62–87.
- Moore, B.C.J., Glasberg, B.R. & Baer, T. 1997. A model for the prediction of thresholds, loudness, and partial loudness. *J Audio Eng Soc*, 45, 224–240.
- Moore, B.C.J., Glasberg, B.R. & Stone, M.A. 1993. Effect on the speech reception threshold in noise of the recovery time of the compressor in the high-frequency channel of a two-channel aid. *Scand Audiol Suppl*, 38, 82–81.
- Moore, B.C.J., Glasberg, B.R. & Stone, M.A. 1999. Use of a loudness model for hearing aid fitting. III. A general method for deriving initial fittings for hearing aids with multi-channel compression. *Br J Audiol*, 33, 241–258.
- Moore, B.C.J., Stainsby, T.H., Alcantara, J.I. & Kuhnel, V. 2004. The effect on speech intelligibility of varying compression time constants in a digital hearing aid. *Int J Audiol*, 43, 399–409.
- Muller, T.F., Harris, F.P. & Ellison, J.C. 2004. Effect of release time on preferred gain and speech acoustics. *J Am Acad Audiol*, 15, 605–615.
- Nábělek, I.V. 1984. Discriminability of the quality of amplitude compressed speech. *J Speech Hear Res*, 27, 571–577.
- Nábělek, I.V. & Robinette, L.N. 1977. A comparison of hearing aids with amplitude compression. *Audiology*, 16, 73–85.
- Novick, M.L., Bentler, R.A., Dittberner, A. & Flamme, G.A. 2001. Effects of release time and directionality on unilateral and bilateral hearing aid fittings in complex sound fields. *J Am Acad Audiol*, 12, 534–544.
- Neuman, A.C., Bakke, M.H., Mackersie, C., Hellman, S. & Levitt, H. 1995. Effect of release time in compression hearing aids: Paired-comparison judgement of quality. *J Acoust Soc Am*, 98, 3182–3187.
- Neuman, A.C., Bakke, M.H., Mackersie, C., Hellman, S. & Levitt, H. 1998. The effect of compression ratio and release time on the categorical rating of sound quality. *J Acoust Soc Am*, 103, 2273–2281.
- Schweitzer, H.C. & Causey, G.D. 1977. The relative importance of recovery time in compression hearing aids. *Audiology*, 16, 61–72.
- Shannon, R., Zeng, F.-G., Kamath, V., Wygonski, J. & Ekelid, M. 1995. Speech recognition with primarily temporal cues. *Science*, 270, 303–304.
- Shannon, R.V., Fu, Q.-J. & Galvin, J. III. 2004. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryng Suppl*, 552, 1–5.
- van Tasell, D.J. & Trine, T.D. 1996. Effects of single-band syllabic amplitude compression on temporal speech information in nonsense syllables and in sentences. *J Speech Hear Res*, 39, 912–922.
- van Toor, T. & Verschuure, H. 2002. Effects of high-frequency emphasis and compression time constants on speech intelligibility in noise. *Int J Audiol*, 41, 379–394.