Canadian Audiologist

The Official Publication of the Canadian Academy of Audiology

Multi-Channel Compression: Concepts and (Early but Timeless) Results

Published May 10th, 2016

Inga Holube, PhD
Mead C. Killion
Volkmar Hamacher

SUMMARY

Most digital hearing instruments offer dynamic compression in several frequency bands (multi-channel dynamic compression). Some of the underlying concepts differ greatly with respect to the number of compression channels, the applied compression ratios and the values of the time constants involved. As these parameters bear great influence on the spectral and temporal structure of speech output signals, a study on the subjective preference of different parameter settings was conducted in 2000, using state-of-the-art technology of that day. Speech-in-quiet comparisons were performed using a single channel compression amplifier, a 4-channel compression amplifier, and a 16-channel compression amplifier. In each case, release times of 15 ms, 380 ms, and 1400 ms were coupled with fast attack times. Five normal-hearing and five hearing-impaired subjects participated in paired comparison tests. The normal-hearing subjects heard compression ratios of 1:1, 2:1 and 8:1. The hearing-impaired subjects had compression ratios lower than, equal to, and greater than that determined by DSL(i/o). Each subject judged 351 paired comparisons of the 27 different conditions. Using the technology available at that time, the results were consistent with few exceptions: Fewer channels of compression, longer release times and lower compression ratios were preferred by both subject groups. While it is clear that improved technology and compression strategies have developed significantly over the last 15 years, this paper may still serve as a useful reminder of the limitations of simple multi-channel compression with short times constants. Simple in this case means the compression in each channel operates independently of the others.

Note: The present paper is based on the first author's presentations at IHCON, Lake Tahoe, CA also presented at the congress of hearing aid acousticians EUHA, Germany, in the year 2000 (Holube et al., 2000a,b).

INTRODUCTION

When fitting sensorineural hearing impairment using linear hearing instruments, soft sounds are typically perceived as too soft while loud sounds are perceived as too loud. Wide-Dynamic-Range Compression (WDRC) solves this problem by using gain that depends upon the input level: Maximum gain for soft sounds and minimum gain for loud sounds. Multi-channel compression (MCC) performs this task in several frequency bands independently. Often cited advantages of MCC are frequency dependent loudness restoration, and therefore listening comfort without the use of a volume control, audibility of soft sounds in all frequency regions, and noise reduction in the case of different frequency spectrum of (loud) noise and (softer) speech. Today, MCC is a standard feature in every hearing instrument. However, the products and the respective fitting

approaches utilize different compression ratios, multiple knee-points, advanced interdependencies across channels and typically different time constants for attack and release in a different number of compression channels. In contrast, the present paper reviews a study which was conducted 15 years ago to investigate the impact of basic multi-channel compression on judged sound quality, as measured by subjective preference, using independent compression in each channel. It focuses on two parameters, compression time constants and the number of independent compression channels, and how they affect speech. Of particular interest was the effect of interactions between compression time constants and compression ratios as measured by the subjective sound quality impression of the subjects.

BACKGROUND

This study was not the first one to focus on the effect of compression. It is reasonable to conclude that the use of WDRC had an important rebirth after the results of Villchur (1973) were published. Although the theoretical advantages of multichannel compression in hearing aids were described by Villchur, his favorable experimental results were obtained with only two channels. When Villchur was a visiting scientist at MIT, a quick experiment in the late 1970s demonstrated that it was possible to have too much of what might have been a good thing: With a computer-based 16-channel compression system later described by Lippmann et al. (1981), each channel was set to an 8:1 compression ratio. While such high compression ratios are normally associated with compression limiting for loud sounds, and would never be realized in modern solutions, the result of the MIT experiment was that with a high vocal pitch, the vowels from the series "heed hid had hod hawed hood who'd" all sounded much the same. The explanation was immediately clear: Excessive compression in each of the 16 one-third-octave channels flattened the peaks in the vowel spectrum sufficiently so that they all sound nearly the same. This cannot happen with a single channel compression amplifier (Killion, 2015).

Several studies examined the effects of compression after Villchur's and Lippmann's results (e.g., Neuman et al., 1998; Moore et al., 1999; van Buuren et al., 1999; Boike and Souza, 2000; Hansen, 2002; Moore et al., 2011; Croghan et al., 2014). Each found decreasing sound quality with increasing compression ratios and decreasing release times, but none of these focused on the number of compression channels. In general, linear processing or long release times were preferred in subjective ratings. In most cases, speech recognition was only marginally influenced by the compression ratio or the number of compression channels. Unfortunately, systematic variations of the number of channels, compression ratio and release time in one system are rare and precise measurement procedures (like paired comparisons) are needed to resolve small differences between the different settings. Therefore, the experiment described here was undertaken to systematically explore the impact of different parameter settings for multiple channel compression systems, representing typical technology from around year 2000, in the basic case where each channel operated independently of the levels in the other channels (contrary to some of current technology) and a classical "simple" attack and release time mechanism is used.

Another reason for formally reporting these earlier data was given indirectly in the data reported by Killion (2004), obtained on more than 60 normal-hearing audiologists (depending on the experiment) and 27 hearing-aid wearers. In 2003, the output of a premium digital hearing aid from each of the six major manufactures was recorded with the aid of the KEMAR manikin. A string quartet from the Chicago Symphony Orchestra and a jazz trio were used as sound sources. The insertion response was deduced from the music recordings by subtracting the spectrum of the open-ear recording from the aided recording. Interestingly enough, the hearing aid with the lowest fidelity ratings from both normal-hearing and hearing-impaired subjects had "syllabic compression" with a release time less than 50 ms (audible as a defect to most listeners) and a

ragged frequency response. Both subject groups gave an average fidelity rating of less than 30% to that aid. One might argue that the low fidelity ratings resulted because that aid had been factory adjusted for speech, not music. But that aid did just as poorly on speech: It had the lowest intelligibility in noise of all the tested hearing aids. The recent trend appears to be toward the use of longer and adaptive time constants and reduced compression ratios, but informal listening test on some modern digital aids suggest that, while some have excellent sound quality and intelligibility in noise, not all of them do. One possible explanation for the latter may lie in the results reported here.

EFFECT OF MULTI-CHANNEL COMPRESSION

The following figures visualize the effect of different release time settings and of different numbers of compression channels on the temporal and spectral structure of speech. For demonstration purposes, the compression ratio was somewhat exaggerated by setting it to 8:1, and no conclusion based on listening tests should be made based on such settings. Fig. 1 shows the influence of the release time on the temporal structure of speech for a section from the fable “The north wind and the sun” (in German: “als sein Wanderer, der in einen warmen Mantel gehüllt war”, in English: “when a traveler … wrapped in a warm cloak”). A short release time reduces the contrast or difference between the peaks and valleys of the temporal envelope. For the graphs in Figure 1, a speech signal was compressed using release times of 1400 ms, 380 ms, and 15 ms as defined by ANSI. The attack time was 3 ms in all three panels showing compression. This figure shows that as the release time is decreased, the temporal contrast is decreased because the soft parts (valleys) of the signal are more amplified than the loud parts (peaks).

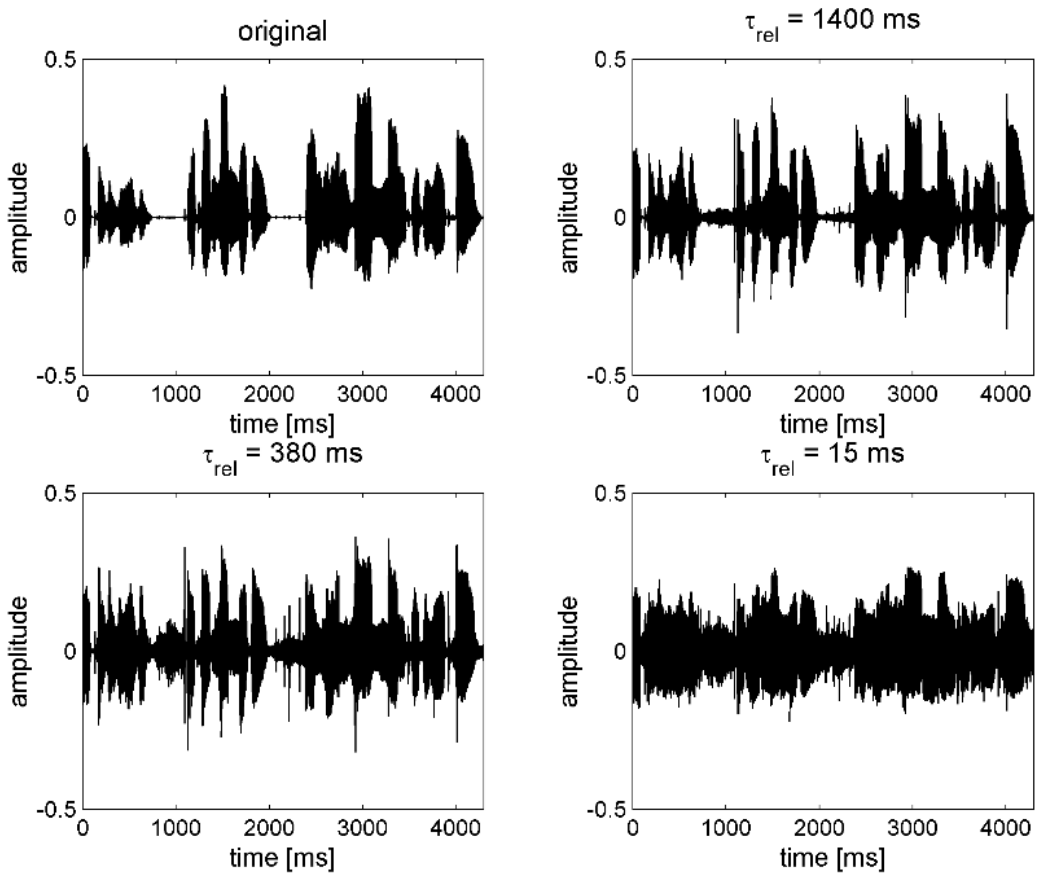


Figure 1. Influence of multi-channel compression release time on the temporal structure of speech for a 16-channel compression system with an exaggerated compression ratio of 8:1. The upper left panel shows the original signal.

Interestingly enough, the number of compression channels can also affect the frequency spectrum of speech. The spectrum of a speech segment processed through a 1-, 4-, and 16-channel dynamic compression system with a short release time (15 ms) and a compression ratio of 8:1 is shown in Fig. 2. In this case, as the number of channels increases, the spectral contrasts of speech are reduced, because the valleys in the spectrum receive greater amplification than the spectral peaks.

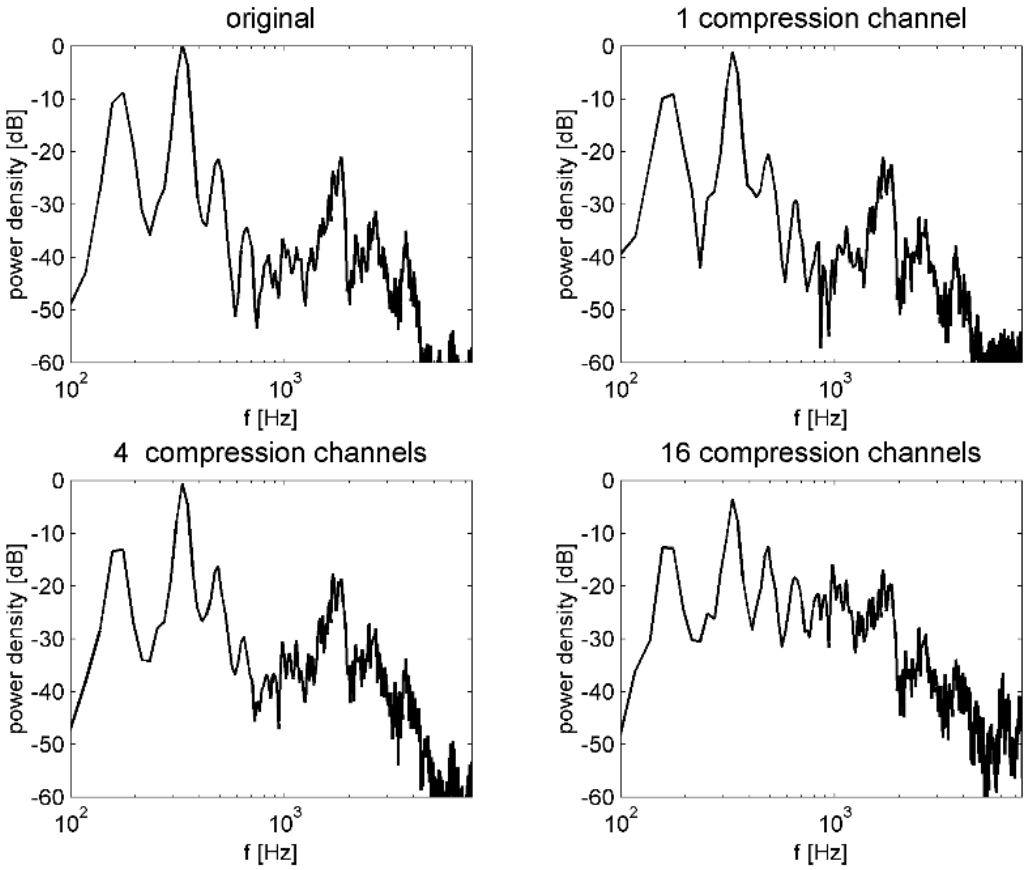


Figure 2. Influence of number of compression channels on the spectrum of a speech segment. The upper left panel shows the original signal.

Alternate Fig. 3 and 4 tell the same story somewhat dramatically in spectrogram form, where in each figure the top spectrogram shows the input and the bottom spectrogram shows the output. The middle spectrogram shows the gain applied to the input. Someone skilled at "reading" spectrograms and familiar with the German language might recognize the sentence "Einst stritten sich Nordwind und Sonne, wer von ihnen beiden wohl der stärkere wäre, als ein Wanderer, der in einen warmen Mantel gehüllt war" (in English: "The north wind and the sun were disputing which was the stronger, when a traveler … wrapped in a warm cloak"). With independent compression in 16 channels and a release time of 15 ms, the output spectrogram at the bottom of Fig. 3 is almost unreadable (and difficult to understand when listening) because so much information has been compressed out. With four channels of compression and a release time of 380 ms, the output spectrogram in Fig. 4 is easily to read because it retains nearly all the information in the original.

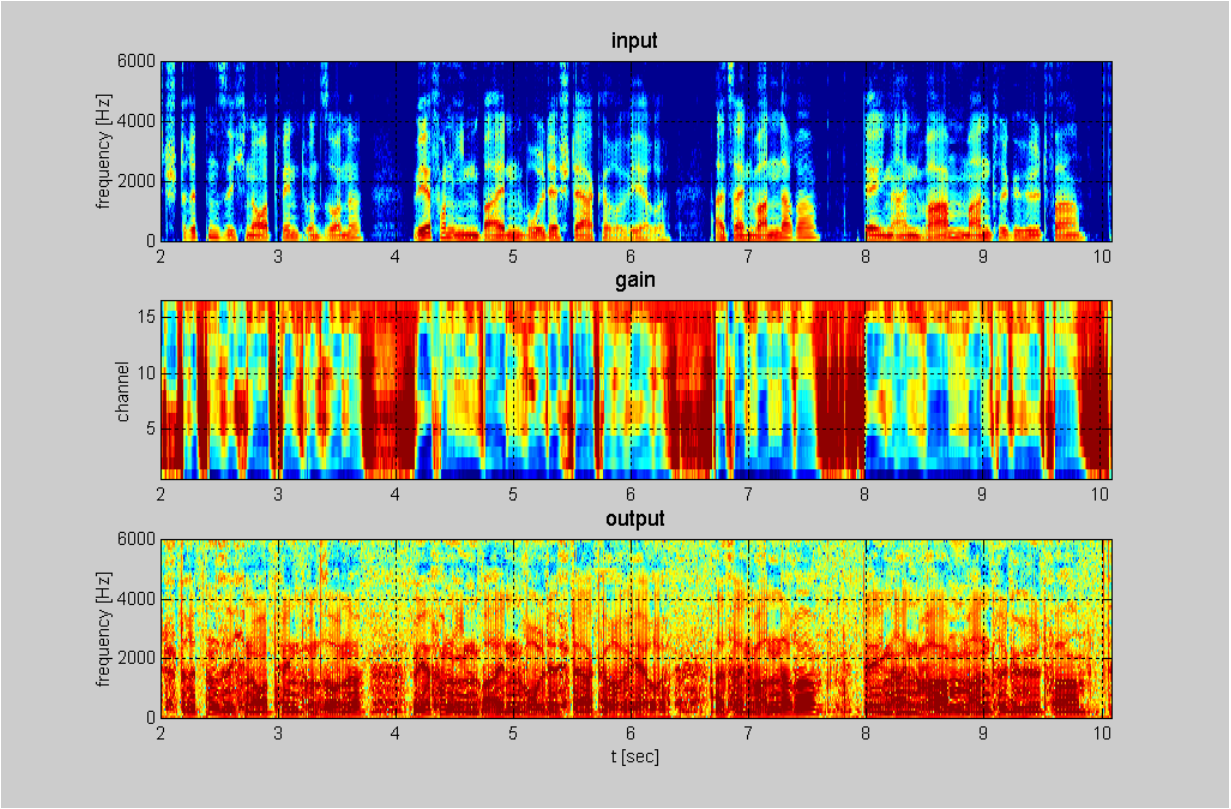


Figure 3. Independent, single knee-point compression in 16 channels with a release time of 15 ms and a compression ratio of 8:1.

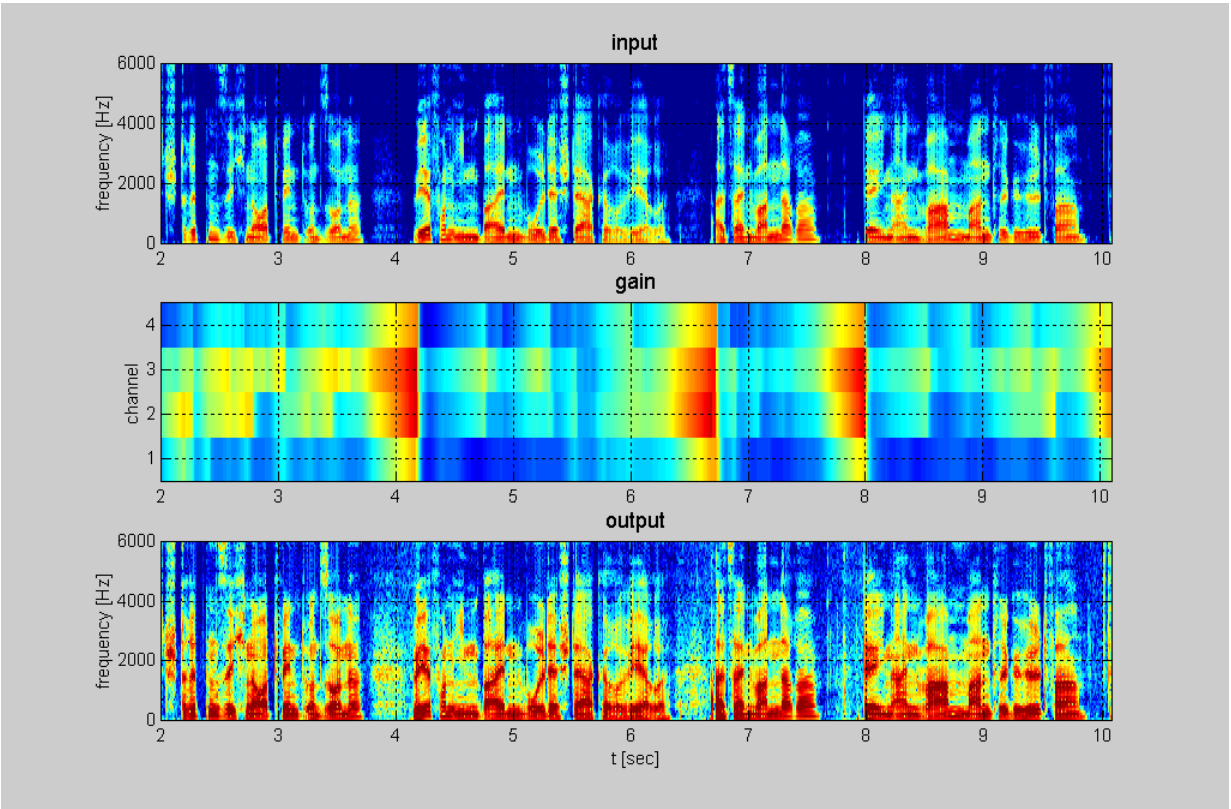


Figure 4. Independent, single knee-point compression in four channels with a release times of 380 ms and a compression ratio of 8:1.

METHODS

A flexible experimental test system with steep filter characteristics and multi-channel compression was used in the experiments described here (see Fig. 5). The input signal was filtered in 16 channels and linearly amplified in each channel according to the hearing loss with the gain settings

G1 to G16. In addition, compression resulted in an additional non-linear gain contribution in the 1, 4, or 16 channels. The 4- and 1-channel compression was derived from the levels in the 16-channels. To explain: The 16 channels were summed to the targeted number of channels, and the respective levels in each summed channel were determined and the corresponding gains were applied in dependence on the input-output function. Three different settings of the compression ratio were used for every group of listeners (normal-hearing and hearing-impaired). The single compression kneepoint was set to 40 dB SPL. The attack time was always set to five periods of the mid-frequency in every channel. The release time was either 15, 380, or 1400 ms according to ANSI S3.22.

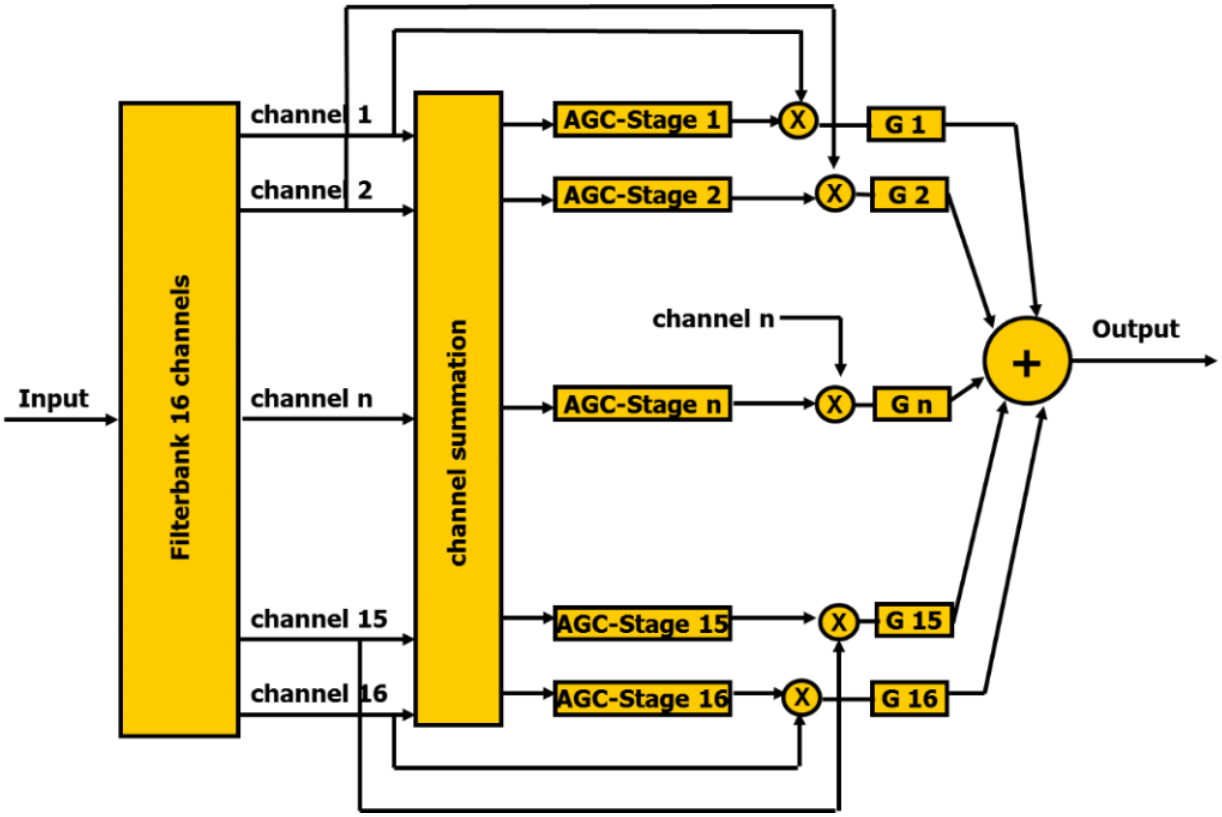


Figure 5. Multi-channel compression system with 16 channels.

Fig. 5 shows a block diagram of the experimental test system which was used. The experiment was steered by a control software running on a PC. Signals were routed via a sound card to external boards for digital signal processing, DSP1 and DSP2, which calculated the multi-channel compression algorithm. The output signals were then routed to one ear of a headphone and the listeners gave their response on a touch screen connected to the PC as shown in Fig. 6.

The task of the listeners was to judge the subjective overall impression of speech in quiet in a paired comparison tournament, i.e., they compared two parameter settings of the compression algorithm and selected that setting with better overall impression. 27 different settings of the compression algorithm, three numbers of channels, three release times and three compression ratios, resulted in 351 paired comparisons per listener. The speech samples (story told by a male speaker) were presented at a comfortable level. The average spectrum of speech was preserved independent of the number of compression channels by respective frequency dependent calibration to speech simulating noise. By use of broadband noise, rather than speech, the average spectrum was normalized even though individual vowels of actual speech were sometimes flattened, as described above. Before the paired comparisons, all parameter combinations were adjusted by each listener to equal loudness.

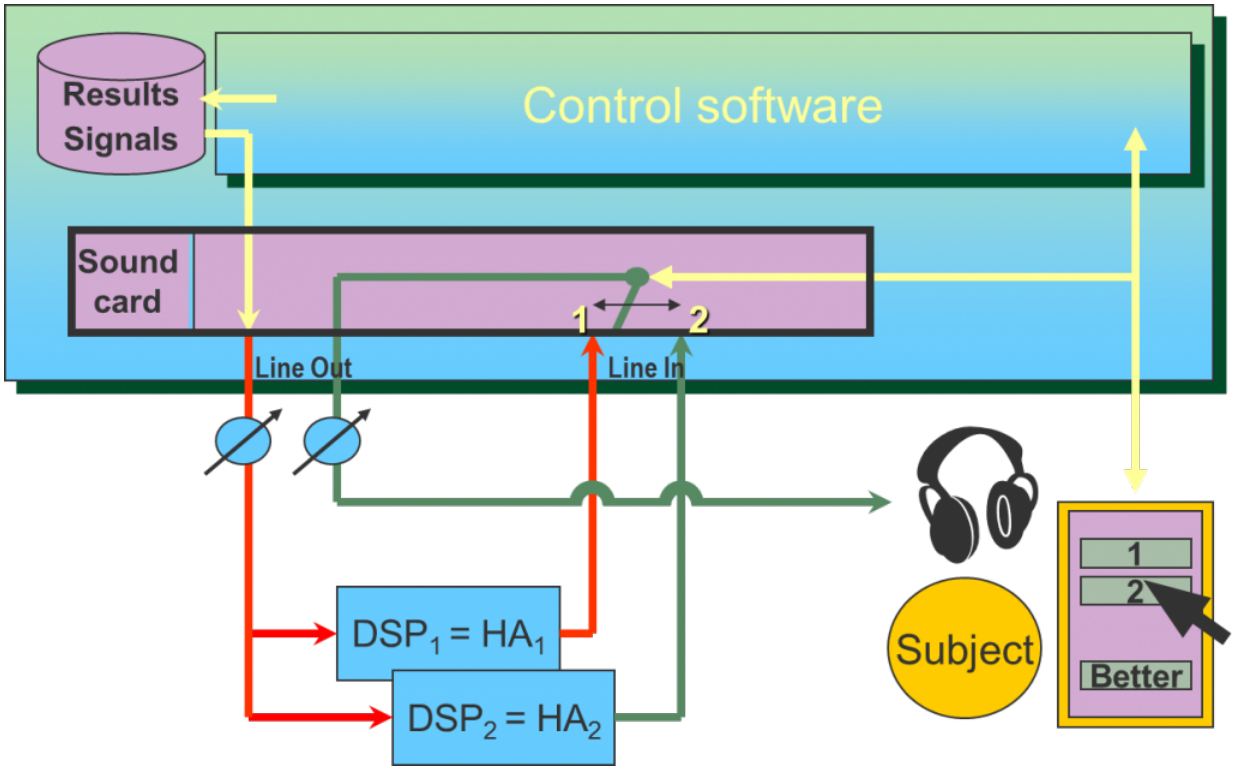


Figure 6. Experimental setup.

SUBJECTS

The experiments were conducted in the year 2000 at the Hörzentrum Oldenburg, Germany. Five normal-hearing listeners and five listeners with a moderate broadband hearing loss participated. Fig. 7 gives the hearing loss of the hearing-impaired listeners. The compression ratios (CR) for the normal-hearing listeners were 1:1 (linear), 2:1 and 8:1. The compression ratios for the hearing-impaired listeners were set frequency dependent for each individual listener according to the calculation in the DSL(i/o) fitting algorithm (denoted as “b”). Two additional frequency dependent settings of the compression ratios were used – one above (denoted as “c”) and one below (denoted as “a”) the DSL(i/o) recommendation, Fig. 8 shows the average settings used.

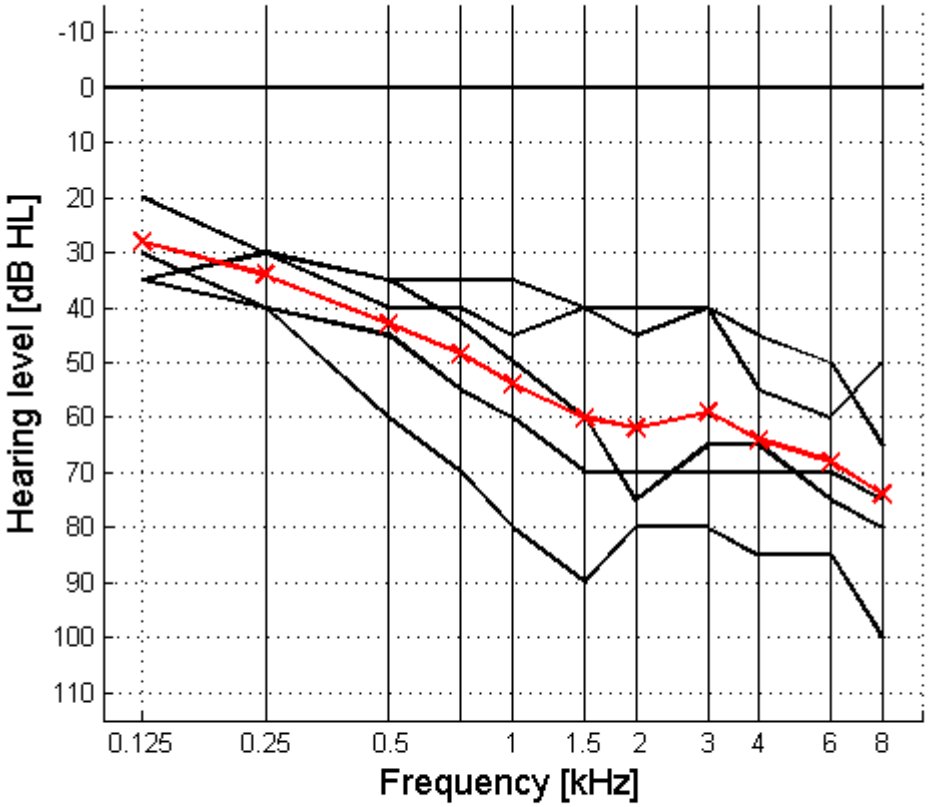


Figure 7. Hearing loss of the five hearing-impaired listeners. The average hearing loss is shown in red.

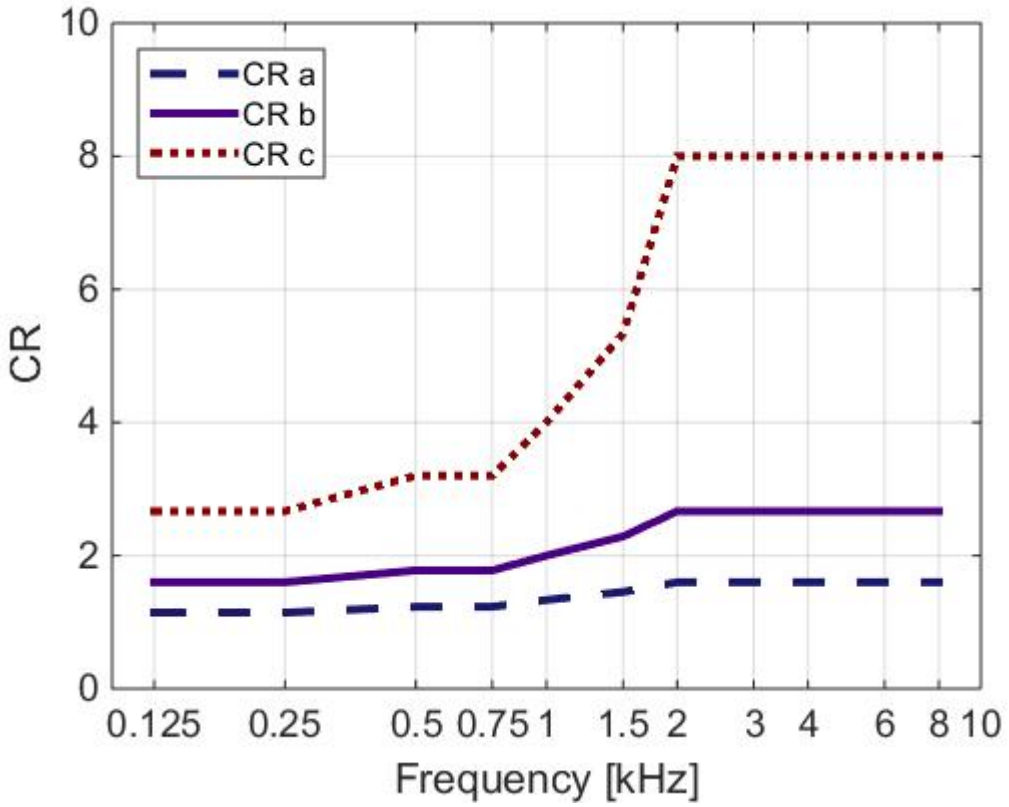


Figure 8. Median CR settings for the hearing-impaired listeners. The solid purple line (b) shows the setting according to the hearing loss calculated with DSL(i/o). The other two lines give a lower and a higher setting.

RESULTS

Fig. 9 shows the number of wins as a measure for the subjective overall impression for each parameter combination in the paired comparison set with normal-hearing listeners. The subjective preference increases with decreasing number of channels, decreasing CR and increasing release time. Similar results can be observed for hearing-impaired listeners, as shown in Fig. 10.

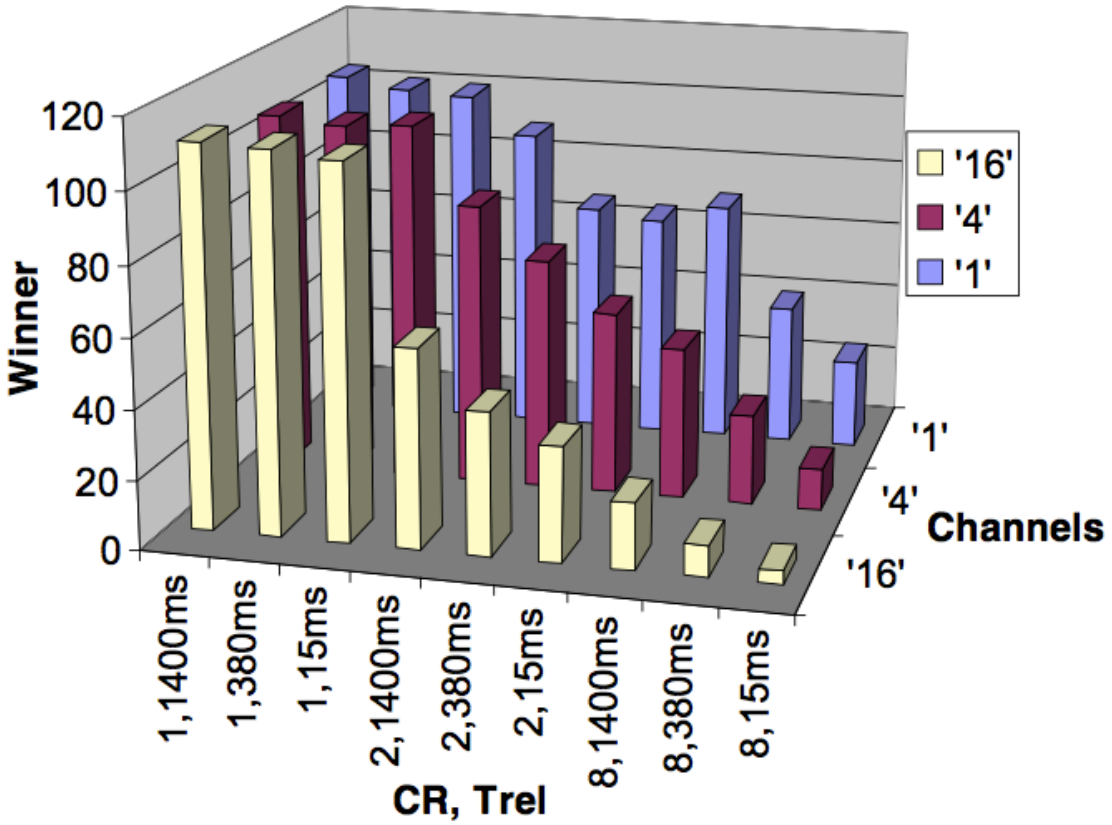


Figure 9. Number of wins for normal-hearing listeners for each of the 27 parameter combinations. The figure shows three rows for the different number of channels (1, 4, or 16) and nine combinations for CR and release time. The first number of the labels on the x-axis give the CR (1:1, 2:1, or 8:1) and the second number gives the release time (15, 380, or 1400 ms).

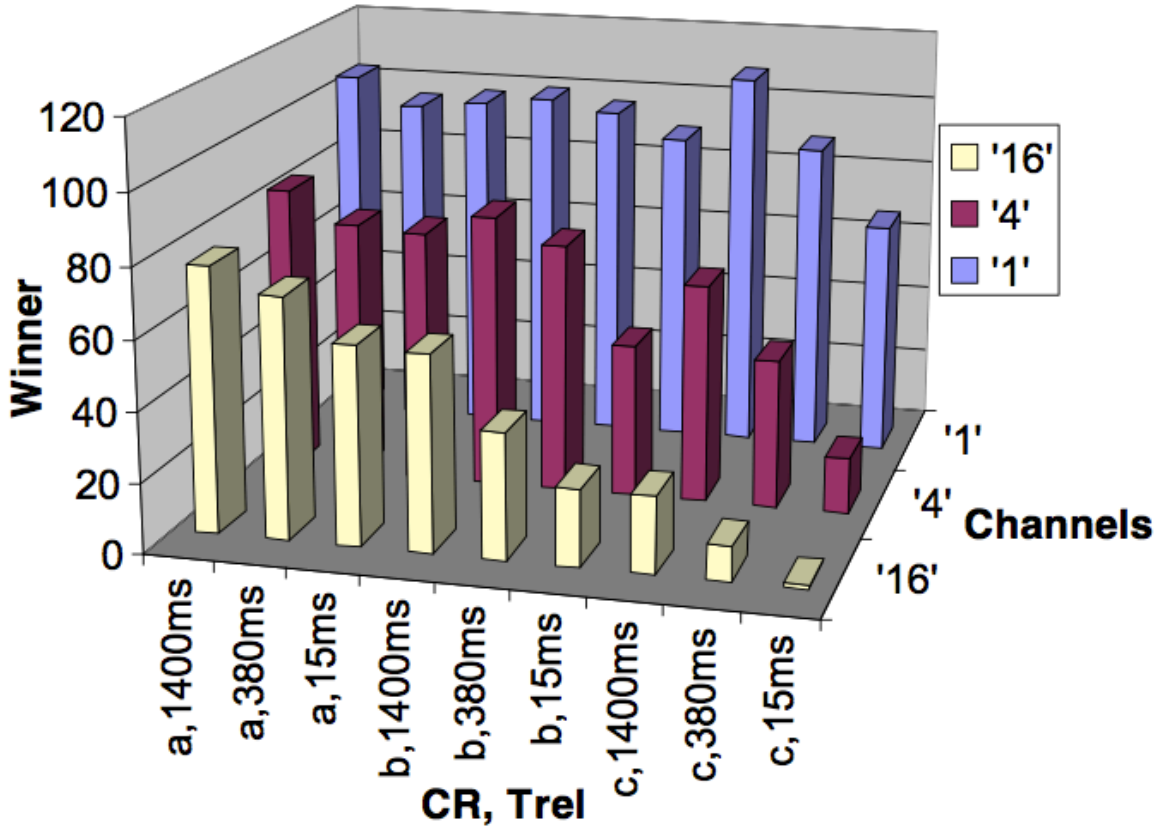


Figure 10. Number of wins for hearing-impaired listeners for each of the 27 parameter

combinations. The figure shows three rows for the different number of channels (1, 4, or 16) and nine combinations for CR and release time. The first number of the labels on the x-axis give the CR (“a” for low, “b” for medium, or “c” for high CR) and the second number gives the release time (15, 380, or 1400 ms).

DISCUSSION

This study, conducted in 2000 and using typical technology of the time (and with channels that operated independently of each other), showed a subjective preference for long release times, low compression ratios and a small number of channels in paired comparison tests for normal-hearing as well as hearing-impaired listeners. It was expected that normal-hearing listeners would prefer as less compression as possible, but it was not clear which parameter (compression ratio, release time or number of channels) has the most influence on the subjective impression and whether hearing-impaired subjects would prefer that setting which corresponds to their hearing loss respective dynamic range. The results showed that the more compression the less preference for normal-hearing as well as for hearing-impaired listeners independent on whether this compression was achieved by a higher number of compression channels, a higher compression ratio or shorter time constants. These results are supported by the analysis of the temporal and spectral contrast of speech signals as shown in Fig. 1 and 2 and are in line with literature results described briefly at the beginning of this paper. As the temporal and spectral contrasts are reduced by varying the release time and the number of channels, the subjective preference for the corresponding parameter combinations is decreased.

Now, the question remains why one should use compression in hearing aids at all and not linear amplification, which seems to be subjectively preferred, and also which compression is the best compression. It has to be taken into account that the experiment was conducted using technology representing state-of-the-art in 2000 for five hearing-impaired listeners with a broadband hearing loss only and only one signal was used to examine the subjective preference for different parameter settings: speech in quiet at a comfortable level. Also, for different number of compression channels the spectrum was matched to the same speech simulating noise and loudness differences between different settings were compensated. These actions were being done to reduce the differences in the compression settings to sound differences beside frequency shaping and loudness impression.

Thus the results obtained here for speech "pre-compressed" to a comfortable level can not be generalized to everyday life or to modern compression technology. Manifold signals with different levels, spectra, and time structure as well as certain amounts of noise are processed by hearing aids. Due to the limited dynamic range of hearing impaired listeners (which might even be highly dependent on the frequency region especially for steeply sloping hearing losses), those signals need different amounts of gain for restoration of audibility and loudness while avoiding discomfort. Also, dynamic compression results in less gain for high level noise, e.g., car noise in the low frequency region. But since the dynamic range of hearing-impaired listeners is rarely less than 30 dB (the dynamic range of speech) this can be achieved mostly with long release times and a small number of compression channels. The minimum number of necessary compression channels seems to be given by loudness restoration, audibility for soft sounds and noise reduction which does not mean “the more channels the better” (On the other hand, an increased number of channels allows for higher resolution in other processing areas like feedback cancellation, noise reduction and adaptive directionality.) A disadvantage of long release times is a decrease in detectability of soft sounds that occur soon after loud sounds. This problem can be solved by using short time constants for short loud sounds only, i.e. a combination of long and short release times dependent on the level changes in the input signal. Such technology has been aroud for many years and indeed, one of the popular single-channel compression (Killion, 1991) used a level-dependent high-frequency

boost coupled with adaptive compression (20 ms recovery from short transients moving to 600 ms recovery for long transients). This avoided "spectral flattening" and excessive compression while maintaining audibility for soft sounds. For many years after the first introduction of digital hearing aids, both fidelity listening tests and intelligibility-in-noise tests suggested superior sound quality using Killion's analog configuration. By the time of the listening-test and speech-in-noise study reported by Killion (2004), multi-channel digital hearing aids with excellent sound quality and intelligibility in noise had been introduced. These improvements have continued during the last decade.

Although, the effect of compression on speech quality shown in this contribution is 15 years old, its general finding that short release times in combination with a high number of compression channels reduce subjective quality is still valid and was recently confirmed for music (Croghan et al., 2014). Nevertheless, the findings of this analysis are limited to classical compression algorithm with “simple” attack and release time mechanisms.

CONCLUSIONS

- This analysis showed that short time constants and a large number of independent compression channels using a state-of-the-art technology circa 2000 reduce the temporal and spectral contrast of speech.
- In subjective paired comparisons, long time constants, low compression ratios and a small number of compression channels were preferred in the investigated technology configuration by normal-hearing and hearing-impaired listeners.

ACKNOWLEDGEMENTS

The study was conducted during the employment of Inga Holube and Volkmar Hamacher at Siemens Audiological Engineering Group, Erlangen, Germany. Thanks to Sivantos GmbH and Hörzentrum Oldenburg, Germany, for their permission to publish the study and several (former) colleagues for their support and data collection.

REFERENCES

Boike, K. T., Souza, P. E. (2000). “Effect of compression ratio on speech recognition and speech-quality ratings with wide dynamic range compression amplification,” J. of Speech, Language, and Hearing Research 43(2), 456-468.

Croghan, N. B. H., Arehart, K. H., Kates, J. M. (2014). “Music preferences with hearing aids: effects of signal properties, compression settings, and listener characteristics,” Ear and Hearing 35(5), e170-e184.

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

Holube, I., Wesselkamp, M., Hamacher, V., Gabriel, B. (2000a). “Multi-channel dynamic compression: Concepts and results,” International Hearing Aid Research Conference IHCON, Tahoe City, California.

Holube I. (2000b) “Multi-channel dynamic compression: Concepts and results,” 45th International Congress of Hearing Aid Acousticians, Germany.

Killion, M. C. (1991). "High Fidelity and Hearing aids,” Audio 75(1), 42-44.

Killion, M. (2004). “Myths that Discourage Improvements in Hearing Aid Design,” The Hearing Review 11(1), 32-40, 70.

Killion, M. (2015). Personal communication.

Lippmann, R. P., Braida, L. D., Durlach, N. I. (1981). “Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss,” J. Acoust. Soc. Am. 69(2), 524-534

Moore, B. C. J., Peters, R. W., Stone, M. A. (1999). “Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips,” J. Acoust. Soc. Am 105(1), 400-411.

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

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(5), 2273-2281.

Van Buuren, R. A., Festen, J. M., Houtgast, T. (1999). “Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality,” J. Acoust. Soc. Am. 105(5), 2903-2913.

Villchur, E. (1973). “Signal processing to improve speech intelligibility in perceptive deafness,” J. Acoust. Soc. Am. 53(6), 1646-1657