

Effects of Instantaneous Multiband Dynamic Compression on Speech Intelligibility

Tobias Herzke

Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany
Email: tobias.herzke@uni-oldenburg.de

Volker Hohmann

Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany
Email: volker.hohmann@uni-oldenburg.de

Received 1 May 2004; Revised 26 January 2005

The recruitment phenomenon, that is, the reduced dynamic range between threshold and uncomfortable level, is attributed to the loss of instantaneous dynamic compression on the basilar membrane. Despite this, hearing aids commonly use slow-acting dynamic compression for its compensation, because this was found to be the most successful strategy in terms of speech quality and intelligibility rehabilitation. Former attempts to use fast-acting compression gave ambiguous results, raising the question as to whether auditory-based recruitment compensation by instantaneous compression is in principle applicable in hearing aids. This study thus investigates instantaneous multiband dynamic compression based on an auditory filterbank. Instantaneous envelope compression is performed in each frequency band of a gammatone filterbank, which provides a combination of time and frequency resolution comparable to the normal healthy cochlea. The gain characteristics used for dynamic compression are deduced from categorical loudness scaling. In speech intelligibility tests, the instantaneous dynamic compression scheme was compared against a linear amplification scheme, which used the same filterbank for frequency analysis, but employed constant gain factors that restored the sound level for medium perceived loudness in each frequency band. In subjective comparisons, five of nine subjects preferred the linear amplification scheme and would not accept the instantaneous dynamic compression in hearing aids. Four of nine subjects did not perceive any quality differences. A sentence intelligibility test in noise (Oldenburg sentence test) showed little to no negative effects of the instantaneous dynamic compression, compared to linear amplification. A word intelligibility test in quiet (one-syllable rhyme test) showed that the subjects benefit from the larger amplification at low levels provided by instantaneous dynamic compression. Further analysis showed that the increase in intelligibility resulting from a gain provided by instantaneous compression is as high as from a gain provided by linear amplification. No negative effects of the distortions introduced by the instantaneous compression scheme in terms of speech recognition are observed.

Keywords and phrases: hearing aids, automatic gain control, fast-acting dynamic compression, speech intelligibility.

1. INTRODUCTION

Dynamic compression is widely used in hearing aids for the compensation of the recruitment phenomenon, that is, the reduced dynamic range between threshold and uncomfortable level, which is often observed in hearing losses of cochlear origin (sensorineural hearing loss). Dynamic compression maps a larger dynamic range of acoustic signals to the reduced dynamic range of hearing-impaired persons, enabling them to perceive signals that would otherwise be below their threshold, while still not exposing them to uncomfortably loud sounds in response to high-level input signals. Hearing aids implementing dynamic compression usually divide the input signal into a small number of frequency bands, and perform dynamic compression in each of these bands. Several rationales for fitting the compression char-

acteristics to the individual hearing loss are pursued. Generally, the approaches aim at restoring audibility, equalizing loudness, renormalizing loudness, maximizing speech intelligibility, or a combination of these approaches (cf. [1]). In combination with the gain characteristics, the time constants of the gain control have been extensively investigated [2, 3, 4, 5, 6, 7, 8, 9]. In terms of perceived speech quality, results generally reveal the best performance for time constants of about one second or more (automatic volume control)—that is, the system shows a largely linear transmission of the speech dynamics on the phoneme, syllable, and word levels. In terms of speech intelligibility, ambiguous results were found in the past regarding optimal time constants and number of frequency bands. The syllabic compression associated with fast-acting compression has a negative effect on the speech intelligibility index (SII) [10], which pursues

the idea of applying slow-acting compression for subband signal-to-noise ratio (SNR) optimization. Other authors argued that syllabic compression should lead to increased audibility of consonants and thus might improve the perception of consonants [11]. Current hearing aids use compression time constants in the range from a few milliseconds to several seconds. Compression usually sets in at low input levels, that is, a wide range of input levels is compressed.

All fitting rationales are mainly driven by empirical findings rather than by objective analysis of the physiological effects of hearing loss. Recently, the physiological basis of sensorineural hearing loss was largely clarified (cf. [12]), which might help to improve fitting rationales in the future. Physiological data show that an instantaneous compression of the basilar membrane motion takes place, which is attributed to the action of the outer hair cells. The compression is active on a wide range of input levels and provides compression of the dynamic range by a factor of about two to three. In combination with the frequency-place transform provided by the micromechanical properties of the basilar membrane, the system is effectively described as a nonlinear filterbank, with an approximately logarithmic time and frequency resolution, and level-dependent filter shapes.

Given the finding that recruitment results from a loss of instantaneous compression on the basilar membrane, recruitment patients might benefit from instantaneous compression performed in a hearing aid. Probably due to the ambiguous results for fast-acting compression, the considerations of the SII, and due to the apparent signal distortions, instantaneous compression for hearing aids has rarely been evaluated. The few previous studies investigating instantaneous compression for hearing aids have used, for example, broadband peak amplitude equalization [13] or octave band waveform compression [14] and reported increased speech intelligibility in quiet [13] or in noise [14], compared to more conventional hearing aids. In order to increase empirical findings in instantaneous compression, we assess the performance of a system which mimics the basic characteristics of the healthy cochlea. An auditory filterbank, that is, the gammatone filterbank [15], is used for frequency analysis and synthesis, which assures an auditory-like combination of time and frequency resolution for all frequencies, which might help reducing perception of distortions. Since the (partial) loss of nonlinear functionality in the impaired cochlea cannot be compensated for by introducing another nonlinear system in the form of an external compressor [16], we do not expect the combined system to behave like the healthy cochlea alone. Audible distortions are still to be expected from the combined system, as well as failure to restore other deficiencies of the impaired cochlea, like reduced frequency selectivity. To avoid excessive distortions, instantaneous envelope compression is used rather than waveform compression, as explained in Section 2.3.

The gain characteristics used for the instantaneous dynamic compression aims at restoring loudness and was deduced, individually for each subject, from adaptive categorical loudness scaling [17]. To investigate the effect of instantaneous multiband compression on speech intelligibil-

ity, tests were performed with nine sensorineural hearing-impaired subjects: The Oldenburg sentence test in noise [18], and a one-syllable rhyme test in quiet [19]. Tests were performed both with instantaneous dynamic compression and with linear amplification. Results for both amplification schemes were compared individually for each subject.

2. METHOD

2.1. Subjects

Nine sensorineural hearing-impaired subjects (4 males, 5 females; aged 49–78 years; median 65 years) participated in the experiment. Eight subjects were regular hearing-aid users, while one subject (AK) had never used hearing aids. Threshold and uncomfortable level were measured for each subject at the standard audiometric frequencies from 125 Hz to 8 kHz. The subjects showed different types and degrees of sensorineural hearing loss. Figure 1 shows the audiograms of all subjects, that is, the threshold of hearing and the uncomfortable level for sinusoids as a function of frequency. The gap between bone and air conduction thresholds was 10 dB or smaller, confirming true sensorineural hearing loss in all subjects. All subjects were paid for their participation on an hourly basis.

2.2. Procedure

Each subject participated in two experimental sessions, each lasting 1 to 1.5 hours. The two sessions were performed on different days within one week. In the first session, standard audiometry, adaptive categorical loudness scaling [17], the Oldenburg sentence test in noise [18, 20], and a one-syllable rhyme test in quiet [19] were performed. Speech tests were performed to train the subjects on the task. In the second session, speech tests were performed after processing with different amplification schemes.

Measurements were performed in two clusters: after finishing experiments with subjects AK, BE, and FL, measurements were reviewed and refined for the remaining subjects. Thus, the frequencies used for loudness scaling and the sampling of preprocessed presentation levels and SNRs differ slightly between these groups of subjects, as described in the following subsections.

2.2.1. Loudness scaling

Categorical loudness scaling was performed at the frequencies 500 Hz, 1 kHz, 1.5 kHz, 2 kHz, 3 kHz, and 4 kHz for the subjects AK, BE, and FL, using third-octave stationary noise (1 second duration). For all other subjects, categorical loudness scaling was also performed at these frequencies, and additionally at 6 kHz. Figure 2 shows equal-loudness-level contours extracted from the loudness scaling data for the subject BH and the mean values of normal-hearing listeners. Mean values were taken from [21]. The numbers next to the contours denote the loudness categorical units. The categories cu 5, cu 15, cu 25, cu 35, and cu 45 correspond to the verbal loudness categories “very soft,” “soft,” “medium,” “loud,” and “very loud,” respectively.

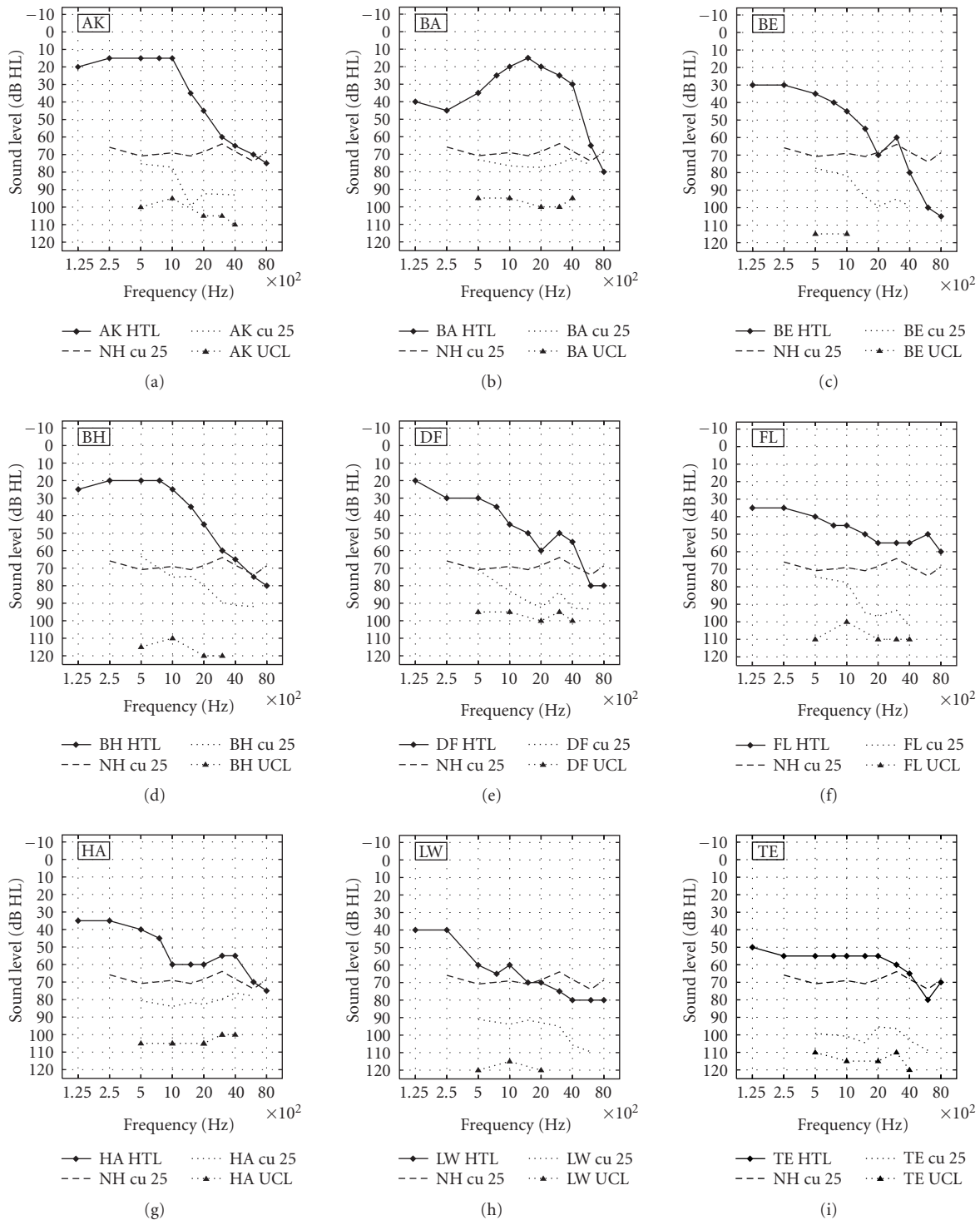


FIGURE 1: Audiograms showing air conduction hearing threshold (HTL) and uncomfortable level (UCL) of the subjects. Additionally, the sound level corresponding to medium perceived loudness (categorical unit 25—cu 25) is shown for each subject, and, for reference, also the mean sound levels for this loudness category in normal-hearing listeners (NH cu 25).

The data from the categorical loudness scaling measurements was used to individually fit the linear amplification and instantaneous compression schemes as described in Section 2.3.1.

2.2.2. Speech intelligibility

In the first experimental session, subjects performed two test lists containing 30 sentences of the sentence test in noise, and four to six test lists containing 47 words of the rhyme test in

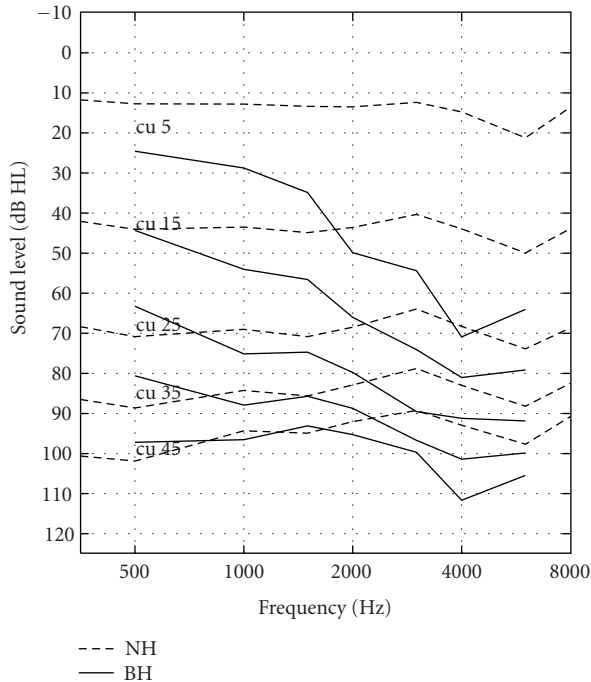


FIGURE 2: Equal-loudness-level contours for subject BH from categorical loudness scaling. For reference, the mean values for normal-hearing listeners are also shown.

quiet. The tests in the first session were performed unaided. Their purpose was to train the subjects on the intelligibility tests. Data were not evaluated.

In the second session, the subjects participated again in the one-syllable rhyme test in quiet and the Oldenburg sentence test in noise. No further training on the intelligibility tests was performed in the second session. For subjects BH, DE, and LW, dynamic compression was presented first, while for the other subjects, linear amplification was presented first. Preprocessed speech material from each of the two different amplification schemes was used, thus simulating either a linear hearing aid or a hearing aid with instantaneous dynamic compression. For the rhyme test, intelligibility in quiet was measured at several presentation levels in order to sample several points on the intelligibility function (percentage of correct words as a function of presentation level). The subjects were also asked for quality judgments and for their subjective preference regarding the speech material of the rhyme test processed with each amplification scheme. The questions asked were: “How would you describe the sound of this hearing aid simulation?” (free verbal description), “Would you accept this sound in hearing aids?” (yes/no), “Which hearing aid simulation do you prefer?” (first/second). The sentence test was performed in noise (speech simulating stationary noise) at a presentation level of 65 dB SPL (input level, i.e., before amplification) and the speech reception threshold, that is, the SNR corresponding to 50% word intelligibility, was measured.

Experimental details

To produce different SNRs, the presentation level of the speech was varied for the linear amplification scheme, while for the dynamic compression, the presentation level of the noise was varied. As confirmed in [20], this difference in presentation mode does not influence the result of this sentence test for normal-hearing or hearing-impaired subjects.

The speech test material was processed offline for each subject. For the linear amplification, the processing was done for all signals separately at a fixed level. Signal configuration, that is, presentation level and SNR, was set afterwards by amplifying and/or mixing preprocessed speech and noise for each subject. In case of the instantaneous compression scheme, this procedure was not possible due to the nonlinearity of the processing. In this case, mixtures of signals and noise at different SNRs and presentation levels needed to be preprocessed by the compressor. The sentence test material for the compression scheme was processed for a presentation level of 65 dB SPL input, with noise added at different levels. In this way, conditions with different SNRs were prepared with steps of 2.5 dB in SNR for the subjects AK, BE, and FL. For all other subjects, the step size was 1 dB. The speech material of the rhyme test was prepared at input presentation levels from 25 dB SPL to 75 dB SPL with steps of 10 dB for the subjects AK and FL. For all other subjects, including BE, input presentation levels of 15 dB SPL, 20 dB SPL, and 30 dB SPL were additionally prepared, since it became obvious that lower levels were needed. For each data point of the rhyme test, one test list consisting of 47 monosyllable words was presented with constant level. Level would be varied between lists to sample data points near and above 50% intelligibility, thus the number of lists per subject varies. Not all preprocessed presentation levels were actually presented during measurements.

For the sentence test in noise, two different procedures for measuring the speech reception threshold (SRT) for 50% word intelligibility were applied, depending on the amplification scheme used. For the linear amplification scheme, an adaptive measurement procedure [22] could be adopted, varying the SNR within a single test list and converging towards the SRT. Using this procedure, the SRT was measured twice in each subject, with each test list containing 20 sentences of five words. This procedure has a standard deviation of the SRT of approximately 0.5 dB [22]. Using dynamic compression, only preprocessed SNRs could be presented, and the adaptive SNR variation procedure could not be used. Instead, intelligibility (percentage of correct words) was measured for several test lists at different constant SNRs using test lists containing 10 sentences each. Several test lists were measured in order to bracket the 50% point of the intelligibility function. As a consequence, the number of test lists used with compression differs between subjects, because we did not know in advance the intelligibility resulting from a particular SNR. Data points could be all on the same side of the 50% point, or too far away from the 50% point to assess the point of 50% intelligibility with satisfactory accuracy.

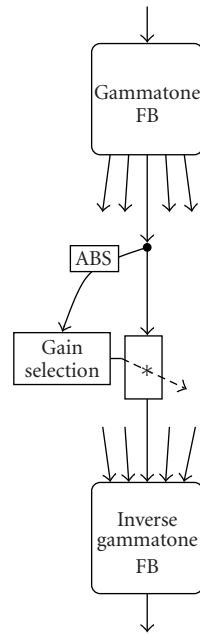


FIGURE 3: Block diagram illustrating the processing of speech data. The source signal is split into critical bands using a gammatone filterbank. In each band, instantaneous envelope compression is performed. A weighted sum, combined with different delays in different bands, is then used to resynthesize the signal.

A psychometric function was fitted to the data points, which was used to determine the SNR corresponding to 50% word intelligibility. The error of this procedure is also in the order of 0.5 dB [22]. According to [22], no systematic bias is introduced by the adaptive SNR variation. We may therefore assume that data from adaptive and nonadaptive procedures can be compared.

2.3. Algorithm

Amplification schemes were implemented using the algorithm sketched in Figure 3. The input signal was split into critical bands using the gammatone filterbank. In each filterbank band, the signal was multiplied with a gain factor and then the subband signals were resynthesized. In the case of the linear amplification scheme, the gain factor was the same for all input levels; it only varied across frequency. For the dynamic compression, the gain factor varied with the input level and the filter's center frequency. The calculation of the gain factor was implemented as follows. The instantaneous envelope of the signal was computed for each sample as the absolute value of the complex-valued output¹ of the respective gammatone filter. The envelope value was then used to lookup the corresponding gain factor from a gain table (gain as a function of input level). The compression was thus in-

¹The gammatone filters used here feature complex-valued filter coefficients and provide complex-valued output which is, in the filter's pass-band, a good approximation of the analytical signal [15]. The gammatone filters thus provide a causal, nonideal Hilbert transform.

stantaneous on the signal's envelope in each band. The inherent time constant is equal to the filter ringing time. Instantaneous envelope compression introduces fewer distortions than instantaneous waveform compression, since the signal's envelope changes more slowly than the waveform itself. Despite this, the distortions introduced by instantaneous envelope compression were clearly audible. The gammatone filterbank and the corresponding frequency resynthesis procedure were the same as in [15], except that they were adjusted to support a sound sampling rate of 44.1 kHz. The filterbank had 30 bands with center frequencies at 73.2, 107.7, 146.0, 188.7, 236.3, 289.4, 348.4, 414.2, 487.5, 569.1, 660.1, 761.4, 874.3, 1000, 1140, 1296, 1470, 1663, 1879, 2119, 2387, 2685, 3017, 3387, 3799, 4259, 4770, 5340, 5974, 6681 Hz.

2.3.1. Fitting of gain tables

Both amplification schemes were fitted with the objective of restoring loudness perception to that of a normal-hearing listener, using loudness functions derived from categorical loudness scaling. In order to individually fit the gain tables, the subject-specific loudness functions were compared with the mean loudness functions from normal-hearing listeners. For dynamic compression, the level difference between narrowband sounds evoking the same loudness perception in normal-hearing listeners and the hearing-impaired subject was used as the gain factor for this level and frequency for dynamic compression. For the linear amplification, the gain factor is constant regardless of the sound level and was chosen to restore loudness perception for the loudness category "medium" (cu 25). Sound levels corresponding to this category for all subjects (measured in this study) and mean sound levels for normal-hearing listeners for this category (taken from [21]) are displayed within the audiogram plots of Figure 1 as dashed curves. The gains used for linear amplification correspond to the level differences between these curves.

Because the center frequency of a gammatone filterbank band was usually between two frequencies of the loudness scaling, the gain factors for that frequency band were computed by first computing the gain factors at the adjacent frequencies where loudness scaling was performed. Then, the gain factors were interpolated to the center frequency of the filterbank band. The interpolation was linear on a logarithmic frequency axis and on a logarithmic gain scale. For filterbank bands with a center frequency outside the range of frequencies used in the categorical loudness scaling, the gain factors for the nearest loudness scaling frequency were used. For low input levels, that is, below 20 dB SPL in each subband, a noise gate was introduced. For 0 dB SPL input level, the output level was set to 0 dB SPL in each band. For 20 dB SPL input and above, the gain factor was deduced as previously described. Between 0 dB and 20 dB SPL, the gain factors were chosen to grow with a constant dB/dB rate so that the prescribed gain at 20 dB SPL input is reached.

2.4. Apparatus

All audiograms were measured using a Siemens Unity audiometer with Sennheiser HDA200 headphones. Adaptive

TABLE 1: Speech reception thresholds (SRT) for the Oldenburg sentence tests. The signal-to-noise ratios (SNR) for 50% word intelligibility are shown for linear processing (mean result from 2 adaptive runs with test lists of 20 sentences each) and for the instantaneous dynamic compression, together with the number of test lists of 10 sentences used to sample points of the intelligibility function. The last column shows individual differences in SRT with the different processing schemes. The difference given here is positive if compression resulted in a lower (better) SRT, and negative if linear processing resulted in a lower SRT. Given accuracies denote the accuracies of the procedures [22]. Accuracy of SRT difference stems from error propagation. Accuracy for mean SRT difference is the standard deviation.

Subject	Compr. SRT (dB SNR) ± 0.5 dB	Lin. SRT (dB SNR) ± 0.5 dB	SRT difference (dB) ± 0.7 dB
AK	-1.98 (4 lists)	-2.25	-0.27
BA	-2.46 (3 lists)	-2.40	+0.06
BE	-0.29 (3 lists)	-2.00	-1.71
BH	-1.72 (2 lists)	-3.55	-1.83
DF	-4.93 (5 lists)	-5.35	-0.42
FL	-3.75 (3 lists)	-3.05	+0.70
HA	-4.02 (3 lists)	-3.95	+0.07
LW	+2.66 (5 lists)	+3.70	+1.04
TE	-0.28 (3 lists)	-1.25	-0.97
Mean	—	—	-0.37 ± 0.99

categorical loudness scaling, Oldenburg sentence test, and one-syllable rhyme test were performed using a personal computer running self-developed audiological testing software. The sounds were presented using an RME DIGI96 sound card, an RME ADI-8 pro D/A converter, a Tucker-Davis HB7 headphone driver, and Sennheiser HDA200 headphones.

Calibration was done using Brüel & Kjær equipment: Artificial ear 4153, 0.5-inch microphone 4192, microphone preamplifier 2669, and measuring amplifier 2610. For calibration of each intelligibility test, a speech-shaped stationary noise was used which had the same long-term spectrum as the speech material of the respective test. The mean root-mean-square (RMS) of the speech material, without pauses at the beginning and the end of the sentences (sentence test) or words (rhyme test), was the same as the RMS of the noise used for calibration.

3. RESULTS

The SRT results of the Oldenburg sentence test in noise (50% word intelligibility) are shown in Table 1. For the dynamic compression scheme, the number of test lists used during measurements is also indicated. Subjects show a large interindividual variation in SRT between +2.7 dB SNR and -4.9 dB SNR when using instantaneous dynamic compression and between +3.7 dB SNR and -5.4 dB SNR for linear amplification. The reference value of normal-hearing listen-

ers is about -6.2 dB SNR [22]. This indicates a significantly reduced intelligibility in noise, which cannot be compensated by the amplification schemes employed. The difference in speech intelligibility thresholds resulting from both amplification schemes is given in the last column. A negative difference value indicates a lower threshold when the linear amplification was used, and vice versa. The mean of the threshold differences between amplification schemes is -0.4 dB. The threshold differences between amplification schemes are not significant (paired t-test, $df = 8$, $\hat{t} = 1.12 < t_{8,0.05} = 2.31$, $p > 0.05$).

Asked for their subjective preference, subjects AK, BA, BH, FL, and TE expressed a strong preference for the linear amplification scheme over the instantaneous dynamic compression scheme and would not accept the instantaneous dynamic compression scheme in hearing aids. Subjects BE, DF, HA, and LW had no preference for either scheme, and would accept both schemes in hearing aids. No subject preferred the instantaneous dynamic compression scheme.

Figure 4 shows results from the one-syllable rhyme test in quiet for linear amplification and for instantaneous dynamic compression. Data show the intelligibility, in percentage, of correct words over the presentation level. Triangles show the data for instantaneous compression, and circles show the data for linear amplification. All intelligibility data are depicted twice in each subfigure. The curves are plotted as functions of input sound levels (open symbols), and additionally as functions of output sound levels (closed symbols), that is, the mean presentation level of a list after either linear amplification or instantaneous dynamic compression have been applied to the speech material. Vertical error bars denote the accuracy of the intelligibility measurement with a rhyme test list of 47 words (Bernoulli experiment with $n = 47$ trials, $\sigma_p = \sqrt{p(1-p)/n}$, cf. [22]). Horizontal error bars show the standard deviation of the gains (broadband) applied to individual words of the rhyme test. These gains were computed by comparing the RMS values of the rhyme test words before and after amplification.

Comparing intelligibility with respect to input levels, the data show that the dynamic range of intelligible speech is generally extended to lower sound levels by instantaneous dynamic compression, as it is expected from dynamic compression in general. However, this is not true for subject BA, which will be further discussed in Section 4. From the curves plotted with respect to the output levels, the effectivity of the additional gain from dynamic compression for speech intelligibility can be inferred by comparing the output levels from both amplification schemes resulting in the same intelligibility. Subjects LW and TE exhibit a good agreement for both amplification schemes. Subjects FL and BE show a slightly better intelligibility when linear amplification is used, whereas subject DF shows a slightly better intelligibility when instantaneous dynamic compression is used. Data from subject BA clearly indicate better intelligibility using linear amplification. For subjects AK, BH, and HA, the intelligibility curves with respect to output level differ, but do not exhibit a bias for any of the amplification schemes.

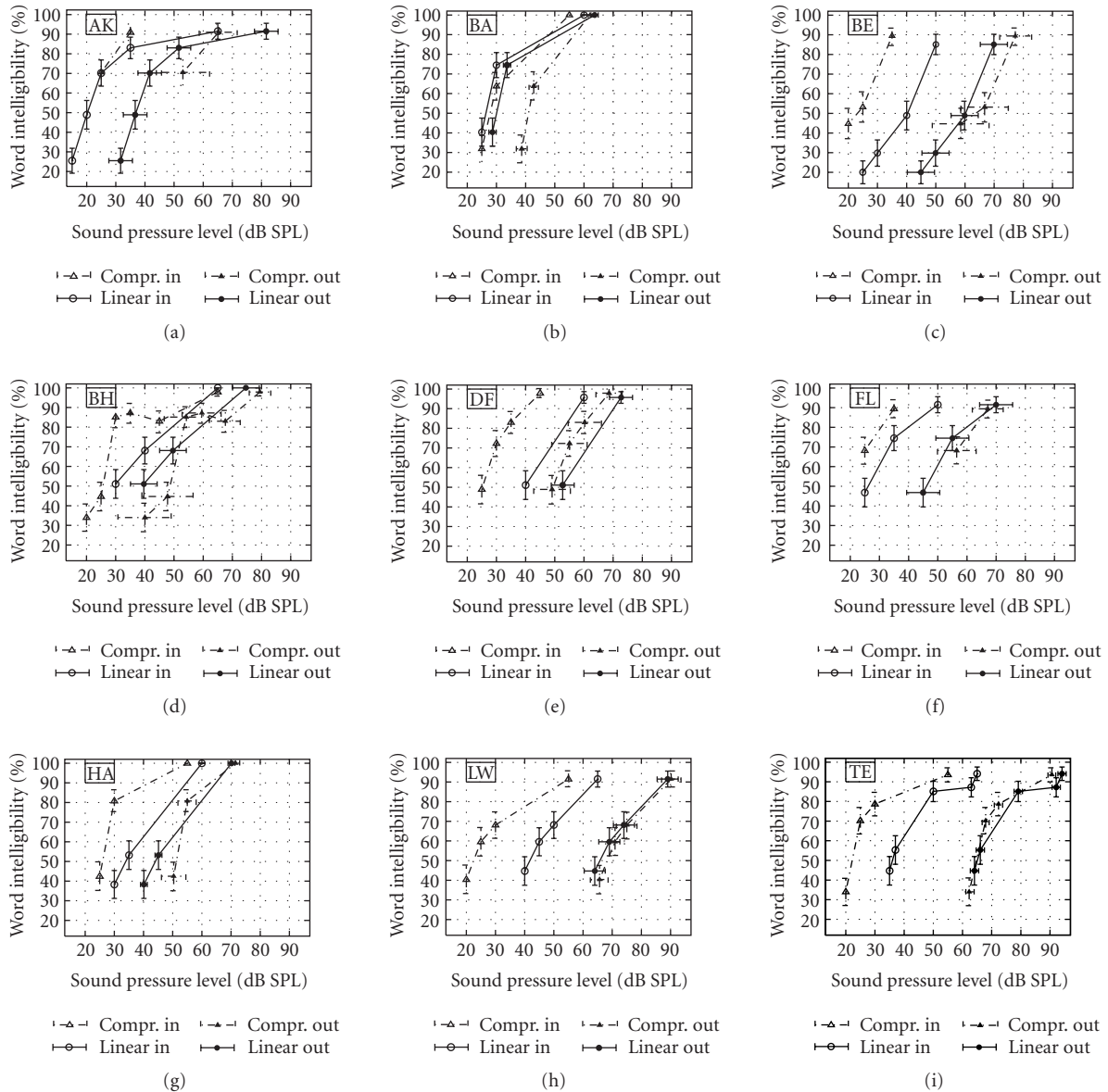


FIGURE 4: Results of the intelligibility test in quiet for all subjects. Diagrams show the percentage of correct words as a function of the presentation level. Triangles show the data of the instantaneous compression, and circles show the data of the linear amplification scheme. Open symbols refer to the input level of the amplification systems, whereas closed symbols refer to the respective output levels. Vertical error bars denote the accuracy of the test for one list, while horizontal error bars show the standard deviation of the average gains applied to individual words of the rhyme test.

Although presentation levels were generally not high enough to reach 100% intelligibility for every subject, results do not indicate increased discrimination loss with either amplification scheme.

4. DISCUSSION

The data from the Oldenburg sentence test in noise exhibit a slightly lower speech intelligibility threshold for linear amplification than for instantaneous dynamic compression. The mean difference in thresholds between both schemes is

0.4 dB, which was found not significant. Because the exposition to the (unfamiliar) instantaneous compression was only short, we expect to see acclimatization effects in speech intelligibility for longer exposition times [23]. Whether these effects exceed the acclimatization effects for the linear amplification is not clear, though.

Despite the signal distortions and reduced modulations introduced by instantaneous dynamic compression, speech intelligibility in noise was not significantly worse in comparison to a reasonably well-fitted linear system. This could be a result of auditory-like time and frequency resolution.

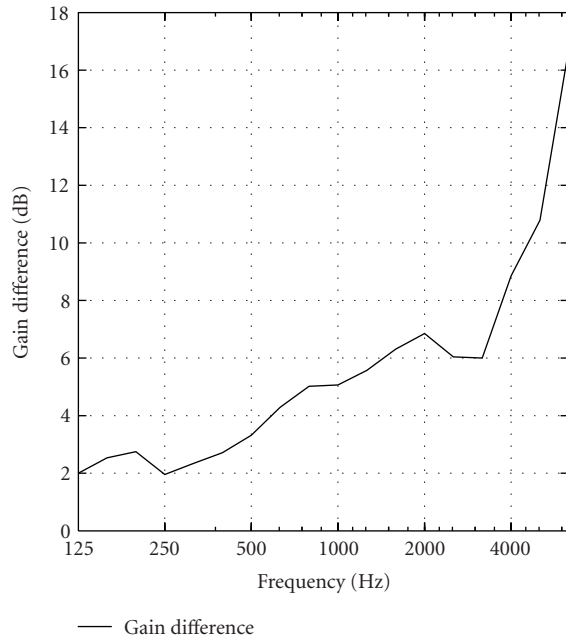


FIGURE 5: Gain differences seen in the sentence test between amplification schemes for -2 dB SNR, averaged over all subjects in third-octave bands. Positive gain difference means that dynamic compression provided higher gain.

To determine whether the fittings used for the different amplification schemes for the sentence test in noise were comparable, we compare average gain factors resulting from both schemes. Figure 5 shows average gain differences between amplification schemes in dB for different frequencies, computed for a typical SNR of -2 dB, that is, a speech level of 65 dB SPL and a noise level of 67 dB SPL. To obtain average gain differences between amplification schemes, intensity spectra of the output signals of a complete test list (10 sentences) were summed over third-octave bands, then divided and thereafter averaged over subjects (mean on a dB scale). The dynamic compression exhibits 2–7 dB higher gains at frequencies relevant for speech intelligibility. In the range of presentation levels used here, the SRT is independent of level [20], thus the comparison made here is reasonable.

Four subjects did not perceive any quality difference between amplification schemes, whereas five subjects clearly perceived the apparent distortion introduced by instantaneous compression. It is unclear whether perceived quality difference could diminish by acclimatization. The audiological data also do not provide evidence, why some subjects did perceive quality differences and other subjects with similar hearing loss did not. The obvious assumption that subjects with high speech reception thresholds in noise should also have difficulties to perceive distortions is not confirmed by the data. Quality factors thus seem to be largely independent of speech intelligibility and should be further investigated.

For the rhyme test in quiet, the similarity of the intelligibility curves with respect to output levels for subjects

LW and TE suggests that these subjects can make use of the gains from linear amplification and from instantaneous dynamic compression with equal effectivity, and are not obstructed by the distortions introduced by the nonlinearity of the dynamic compression scheme. For subjects AK, BH, and HA, the intelligibility curves for the dynamic compression scheme exhibit, for low input levels, lower intelligibility for the dynamic compression scheme than for linear amplification. This relation is reversed for higher input levels. It is possible that the expansion that was introduced in the dynamic compression scheme for low input levels plays a role in this effect and that the subjects would have had better intelligibility results in these cases without expansion. In subject BA, the output level difference between linear and compression schemes at the same intelligibility is about 12 dB, indicating that the subject cannot exploit the additional gain provided by instantaneous compression at all. The reason for this result might be that gain is mainly provided outside the main frequency range of speech due to the unusual hearing loss of this subject. In this case, compression increases the broadband output level, but does not increase audibility of relevant speech segments. Further analysis of this assumption could be carried out by narrow-band output level analysis in combination with speech-based frequency band importance weighting. This is not pursued here because of the clear-cut results found in the other subjects.

The aim of the fitting of the linear amplification was to restore loudness perception for medium loudness. However, the presentation levels used during the rhyme test in quiet were well below this category most of the time. This might raise the question as to whether the comparison between instantaneous dynamic compression and linear amplification was fair after all, since dynamic compression uses higher gains at lower levels. Even when comparing speech intelligibility at the same output level of both processing schemes, as was done here, the frequency shape of the output signals will be different, and part of the speech spectrum was most likely inaudible with linear amplification for low presentation levels. One could assume that the similarity of the results that we observe would stem from a tradeoff between inaudibility for linear amplification and distortions for instantaneous dynamic compression. But if audibility of these parts of the spectrum played a role, we would expect to see an advantage for linear amplification over instantaneous dynamic compression with increased level, where audibility is assured, but distortions are still present. Since we do not see such an advantage for linear amplification for high levels in the results, we may assume that the audibility question for linear amplification plays only a minor role in this setup. Regardless of the audibility question, results show that the distortions do not lead to a discrimination loss. However, further studies should use a linear reference response which is adjusted to the average response of the compression system for each specific condition and level, similar to what was done post hoc in Figure 5.

Although results show that distortions introduced by instantaneous compression do not hamper speech reception,

methods to further reduce apparent distortions should be further investigated. One possibility is to restrict the bandwidth of the compressed subband signals to the respective frequency band before resynthesis, as it is done in [14]. This postfiltering attenuates distortion components outside the frequency band in which they are generated and thus should lead to less audible distortions.

Gain characteristics were individually fitted to normalize loudness in this study. Other fitting rationales, for example, loudness equalization or SNR maximization, might lead to a slightly different absolute performance of the amplification schemes. However, the main finding of this study, that is, that instantaneous compression performs as good as linear amplification regarding speech intelligibility despite the distortions and the modulation reductions, is likely to be not altered when varying the fitting rationale.

Speech reception thresholds in noise were found to be higher in normal-hearing listeners than in all hearing-impaired subjects, which confirms other studies on speech reception in noise conditions in the hearing-impaired. Degraded speech reception in speech-simulating noise is generally attributed to loss of frequency selectivity and increased internal noise in the impaired auditory system. It can be assumed that this type of information loss cannot be compensated for by amplification or compression, which is the consequence of the fact that the combination of two nonlinear systems generally cannot be made equivalent to another nonlinear system, that is, the healthy cochlea. This stresses the necessity of combining proper amplification schemes and true SNR improvement by noise reduction when pursuing full rehabilitation of hearing impairment.

5. CONCLUSION

Sentence intelligibility tests performed with nine sensorineural hearing-impaired subjects showed that instantaneous dynamic compression has no effect on speech intelligibility in noise, when compared to reasonably well-fitted linear amplification.

Word intelligibility tests in quiet showed that, compared to a linear amplification system, additional amplification provided by instantaneous compression can be exploited and turned into increased speech intelligibility by the subject.

Distortions introduced by instantaneous compression have a negative impact on acceptance, but not on intelligibility.

ACKNOWLEDGMENTS

We thank Birger Kollmeier for his continuous support and all the members of the working group Medizinische Physik for fruitful discussions. We especially thank T. Brand, D. Berg, and K. Wagener for their help with the audiometric testing setup. We also thank three anonymous reviewers for helpful comments on earlier versions of the manuscript. This work is supported by the centre of excellence on hearing aid technology (BMBF 01EZ0212).

REFERENCES

- [1] H. Dillon, *Hearing Aids*, Thieme Medical Publishers, New York, NY, USA, 2001.
- [2] R. Plomp, "The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function," *Journal of the Acoustical Society of America*, vol. 83, no. 6, pp. 2322–2327, 1988.
- [3] I. V. Nábelek, "Performance of hearing-impaired listeners under various types of amplitude compression," *Journal of the Acoustical Society of America*, vol. 74, no. 3, pp. 776–791, 1983.
- [4] G. Walker, D. Byrne, and H. Dillon, "The effects of multichannel compression/expansion amplification on the intelligibility of nonsense syllables in noise," *Journal of the Acoustical Society of America*, vol. 76, no. 3, pp. 746–757, 1984.
- [5] V. Hohmann and B. Kollmeier, "The effect of multichannel dynamic compression on speech intelligibility," *Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1191–1195, 1995.
- [6] J. Verschure, E. J. Benning, M. Van Cappellen, W. A. Dreschler, and P. P. Boeremans, "Speech intelligibility in noise with fast compression hearing aids," *Audiology*, vol. 37, no. 3, pp. 127–150, 1998.
- [7] B. C. J. Moore, R. W. Peters, and M. A. Stone, "Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips," *Journal of the Acoustical Society of America*, vol. 105, no. 1, pp. 400–411, 1999.
- [8] M. Hansen, "Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility," *Ear and Hearing*, vol. 23, no. 4, pp. 369–380, 2002.
- [9] M. A. Stone and B. C. J. Moore, "Effects of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task," *Journal of the Acoustical Society of America*, vol. 114, no. 2, pp. 1023–1034, 2002.
- [10] ANSI S3.5—1997, "Methods for the calculation of the speech intelligibility index," American National Standards Institute, New York, NY, USA, 1997.
- [11] L. Hickson and D. Byrne, "Consonant perception in quiet: effect of increasing the consonant-vowel ratio with compression amplification," *Journal of the American Academy of Audiology*, vol. 8, no. 5, pp. 322–332, 1997.
- [12] B. C. J. Moore, *Perceptual Consequences of Cochlear Damage*, Oxford University Press, Oxford, UK, 1995.
- [13] K. Kaneko, K. Shoji, H. Kojima, et al., "Nonlinear digital hearing aid with near-instantaneous amplitude compression," *European Archives of Oto-Rhino-Laryngology*, vol. 258, no. 10, pp. 523–528, 2001.
- [14] J. L. Goldstein, M. Oz, P. Gilchrist, and M. Valente, "Signal processing strategies and clinical outcomes for gain and waveform compression in hearing aids," in *Proc. 37th IEEE Asilomar Conference on Signals, Systems, and Computers*, pp. 391–398, Pacific Grove, Calif, USA, November 2003.
- [15] V. Hohmann, "Frequency analysis and synthesis using a Gammatone filterbank," *Acustica / acta acustica*, vol. 88, no. 3, pp. 433–442, 2002.
- [16] C. Giguère and G. F. Smoorenburg, "Computational Modeling of outer haircell damage: implications for hearing and signal processing," in *Psychophysics, Physiology, and Models of Hearing*, T. Dau, V. Hohmann, and B. Kollmeier, Eds., pp. 155–164, World Scientific, Singapore, 1999.
- [17] T. Brand and V. Hohmann, "An adaptive procedure for categorical loudness scaling," *Journal of the Acoustical Society of America*, vol. 112, no. 4, pp. 1597–1604, 2002.

- [18] K. Wagener, T. Brand, and B. Kollmeier, "Development and evaluation of a German sentence test," in *Contributions to Psychological Acoustics - 8th Oldenburg Symposium on Psychological Acoustics*, pp. 439–466, bis-Verlag, Oldenburg, Germany, 2000.
- [19] E. L. v. Wallenberg and B. Kollmeier, "Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten," *Audiologische Akustik*, vol. 28, no. 2, pp. 50–65, 1989.
- [20] K. C. Wagener and T. Brand, "Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters," *International Journal of Audiology*, vol. 44, no. 3, pp. 144–156, 2005.
- [21] B. Kollmeier, *Hörflächenskalierung - Grundlagen und Anwendungen der kategorialen Lautheitsskalierung für Hördiagnostik und Hörgeräteversorgung*, Median-Verlag, Heidelberg, Germany, 1997.
- [22] T. Brand and B. Kollmeier, "Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests," *Journal of the Acoustical Society of America*, vol. 111, no. 6, pp. 2801–2810, 2002.
- [23] S. Gatehouse, "The time course and magnitude of perceptual acclimatization to frequency responses: Evidence from monaural fitting of hearing aids," *Journal of the Acoustical Society of America*, vol. 92, no. 3, pp. 1258–1268, 1992.

Tobias Herzke received his Physics degree (Dipl.-Phys.) from the University of Oldenburg, Germany, in 2003. Since 2003, he has been a faculty member of the Institute of Physics at the University of Oldenburg, Germany, and member of the Medical Physics Group (Prof. Dr. Dr. B. Kollmeier), where he works on the development, implementation, and evaluation of different signal processing algorithms for hearing aids.



Volker Hohmann received the Physics degree (Dipl.-Phys.) and the Doctorate in physics (Dr. rer. nat.) from the University of Göttingen, Göttingen, Germany, in 1989 and 1993, respectively. Since 1993, he has been a faculty member of the Institute of Physics at the University of Oldenburg, Germany, and member of the Medical Physics Group (Prof. Dr. Dr. B. Kollmeier), teaching activities in physics for undergraduate and graduate courses. His research expertise is in acoustics and digital signal processing with applications to signal processing in speech processing devices, for example, hearing aids. He is a Consultant with the Hörzentrum Oldenburg GmbH, and was a guest researcher at Boston University (Prof. Dr. Colburn) in 2000.

