Time constants of compression schemes – Less is more?

Matthias Latzel[1], Volker Hohmann[2] and Kirsten Wagener[3]

[1] *Audiological Research, Siemens Audiologische Technik, D-91058 Erlangen, Germany*

[2] *Medical Physics Section, University of Oldenburg, D-26129 Oldenburg, Germany*

[3] *Hörzentrum Oldenburg GmbH, D-26129 Oldenburg, Germany*

In recent years, Wide Dynamic Range Compression (WDRC) has been established in the listening device industry so that it became the most utilized tool in modern hearing instruments. However, although the general system has become status quo, the question of a correct fitting of the necessary parameters is still open. The developed fitting rules only calculate the target gain and perhaps the compression ratio, the number of channels and/or the compression kneepoints. Unfortunately, the time constants are generally not considered.
The present work presents investigations into this area. Two compression systems operating with completely different time constants were compared. One system works instantaneously with very short time constants. Additionally, the applied gain and compression depends on the distance of the instantaneous frequency from the centre frequency of the particular frequency band. The second system is a compression system which is already available in commercial hearing devices and which allows limited access to the time constants.
For the evaluation, both subjective and objective speech tests were used. In addition, a static and a dynamic loudness scaling method were integrated to provide information regarding how the loudness is normalized. In the third part of the evaluation, several sound samples were presented in different level ranges to be judged (i) absolutely using questionnaires and (ii) relatively in complete paired comparisons. In all cases, the compression systems were evaluated in level ranges relevant for real life listening situations.
The results found with the instantaneous compression system do not confirm the assumption of improved speech intelligibility in modulated noise but they show a better restoration of loudness than in conventional compression systems without deteriorating the sound quality.

MOTIVATION

In recent years, Wide Dynamic Range Compression (WDRC) has been established in listening devices and become one of the most utilized methods in modern hearing instruments to compensate for sensorineural hearing loss. In several papers the number of frequency bands, compression ratio and/or compression kneepoint and time constants and their respective effects on speech intelligibility and sound quality have been extensively investigated (see Souza, 2002 for an overview). The parameter "time constant" evokes conflicting views. E.g. Hansen (2002) disbelieves fast time constants to be the right mean to compensate for recruitment. Contrary to Hansens findings, Ver-

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.).
ISBN: 87-990013-1-4. Print: Centertryk A/S.

schuure *et. al* (1995) found out that fast time constants are advantageous to improve speech intelligibility which is considered to be a direct consequence of a successful compensation of recruitment. For Giguere and Smoorenburg (1999) it seems reasonable to use extremely fast time constants in order to compensate for outer hair cell loss. Herzke and Hohmann (2005) picked up this general idea and studied the effect of instantaneous compression (IC) on speech intelligibility with ambivalent results. The IC approach has been extended by controlling the effective compression by the sub-band instantaneous frequency to consider the two-tone suppression effect (Hohmann, 2006). It was evaluated by Bisitz and Hohmann (2006) who found promising results in terms of speech intelligibility improvements in situations with fluctuating noise (see also Hohmann 2007).

This idea of an auditory-model-based *instantaneous compression* scheme has been investigated further in the present paper by comparing a simplified implementation of the algorithm proposed by Hohmann (2006) and Bisitz and Hohmann (2006) with a *standard compression* scheme already available in commercial hearing instruments.

This led us to the following research questions:

- How does instantaneous compression perform in terms of speech intelligibility in comparison to the WDRC scheme integrated in the commercial hearing device?
- How is the sound quality affected by instantaneous compression?
- How does a linear setting perform in comparison to both compression schemes?

SET-UP

Subjects

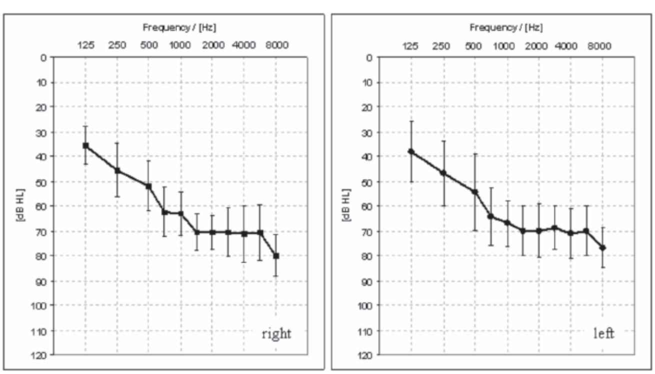

Fig. 1: Average (inter-individual) hearing loss (HTL) for all subjects, right and left ear.

Eight subjects with sensorineural hearing loss took part in the study. Fig.1 illustrates that the constraints regarding their hearing loss, e.g. symmetric and moderate to severe were achieved. Furthermore, the subjects were experienced hearing aid users to assure reliable results. They were paid for their participation on an hourly basis.

Methods

The choice of the audiological tests used in this study depended on one hand on the level ranges which are relevant for hearing aid processing in real life, in particular, speech processing. On the other hand it was important to sweep through the whole dynamic range to challenge the compression schemes under test.

Based on these considerations the following audiological test methods and measuring conditions were selected and conducted during three sessions:

- Pure-tone audiogram
- Categorical *loudness scaling* (broadband, binaural, unaided ⇔ aided) with stationary - noise taken from Oldenburg Sentence Test (Wagener *et. al,* 1999) and with „dynamic/transient noise“ using very short signals comprising transient components.
- *Rhymetest* (v. Wallenberg and Kollmeier: 1989) in quiet with subsequent absolute rating of subjective speech intelligibility and sound quality for speech level of 50 dB and 80 dB
- *Oldenburg Sentence Test* (OLSA) in modulated noise (Wagener and Brand, 2005; Wagener *et al.*, 2006) with 65 dB and 75 dB noise level
- Complete *paired comparison* of pre-processed sound samples
- *Absolute Rating* of pre-processed sound samples on a 7-point scale.

Sound samples were five pre-processed signals with presentation levels relevant for real life listening situations (soft, middle, loud, dynamic). Pre-processing of the sound samples was performed by presenting all samples in the free-field to a KEMAR manikin wearing hearing aid dummies with closed ear molds. Microphone signals were recorded and replayed via headphones during the absolute rating and paired comparison tasks and stored for subsequent processing.

Non-parametric tests (Wilcoxon and Friedman Test) were used for the *statistical analysis* of the data due to the small number of subjects.

Research Platform

All compression schemes were realized in real-time on the same platform (see Fig.2): the PC-based Master Hearing Aid System (MHA, Grimm *et al.*, 2006). The input signals were taken from hard disk, processed by the MHA and presented to the subjects via hearing aid dummies comprising real hearing aid receivers and closed ear molds.

The IC approaches (for two parameter settings IC1 and IC2) were fitted considering the “individual” loudness function estimated from the hearing thresohold and uncomfortable level. The resulting target gains were corrected based on the generic fitting rule LoudFit+ (Kiessling *et al.*, 2006).

The standard compression scheme is fitted regarding the common fitting rule which has been well established on the market (NAL-NL1 with acclimatization step 2). From now on this fitting scheme will be referred to as CEN.

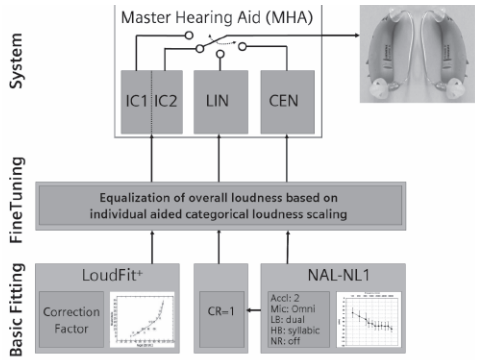

Fig. 2: Hardware Set-up and fitting/fine-tuning procedure for test conditions.
IC1 and IC2: Settings of instantaneous compression approach
CEN: WDRC: state-of-the-art compression
LIN: WDRC: state-of-the-art compression in linear setting

For reference the linear version of CEN is realized with the (linear!) gain set to the gain for 65 dB input of CEN. This condition will be referred to as LIN in the following.

Fine tuning: To avoid differences in overall loudness perception for the different approaches used in this study, the overall gain was interactively readjusted for all conditions by equalizing the aided categorical loudness scaling results for the stationary speech-shaped noise at an categorical loudness (CU) of 25 ('Medium').

RESULTS AND DISCUSSION

Categorical Loudness Scaling

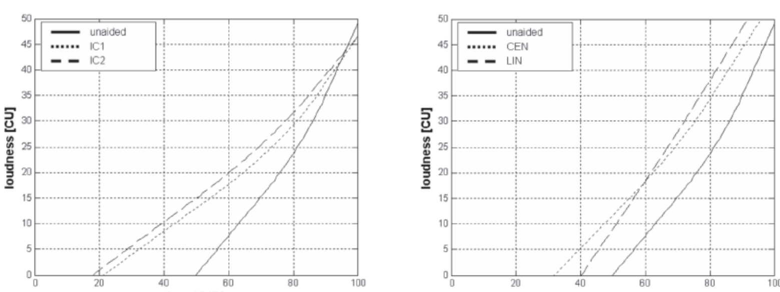

Fig. 3: Individual categorical loudness functions for one subject in all conditions. IC1 and IC2(left panel) and conditions LIN and CEN (right panel) Signal: Stationary - noise

Categorical Loudness Scaling (Stationary noise): Fig.3 shows exemplarily (based on one subject's data) the scaling data of broadband noise for all conditions. Data demonstrate that IC1 and IC2 (Fig. 3, left panel) provide the necessary compression in order to almost normalize loudness. In contrast the conditions CEN and LIN (Fig. 3, right panel) give less compression, especially at high input levels where these settings prescribe too much gain.

Additionally the data show that IC1 and IC2 prescribe more gain at low input levels than LIN and CEN. This aspect directly implies a better speech intelligibility for low speech levels (see next section).

Furthermore, the figure illustrates the correctness of the fine-tuning procedure: We can observe that the aided loudness functions are close together for CU = 25 meaning that all settings provide the same loudness perception for almost the same broadband input level at 'Medium' loudness.

Categorical Loudness Scaling (Dynamic/transient noise): It is generally assumed that very short time constants (conditions IC1 and IC2) are better at restoration of loudness perception of very short sounds with abrupt level changes than compression schemes with long time constants (LIN and CEN). The data of the categorical loudness scaling with "transient noise" does not confirm this hypothesis since for high input levels the instantaneous approach provides too much compression (see Fig.4, left panel)). This is possibly due to an overestimation of the suppression effect in the current approach. Thus the state-of-the-art compression approach with conventional time constants is proven to be advantageous for this signal (see Fig. 4, right panel). Neither LIN nor CEN restored normal loudness perception as well which is presumably not the goal for a compression schedule as a complete compensation would lead to the rejection of the hearing instrument by the user. However, the loudness function in condition CEN is parallel to the normal loudness function which implies the correct adjustment of the compression parameter for this particular test-design.

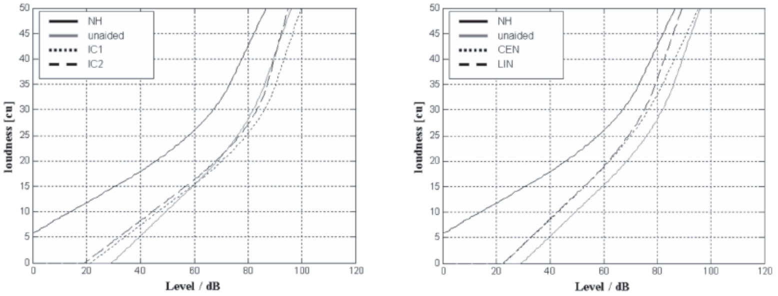

Fig. 4: Mean (all subjects) categorical loudness functions for conditions IC1 and IC2 (left panel) and conditions LIN and CEN (right panel) Signal: Dynamic/transient noise.

Speech Tests

Speech Tests in modulated noise (OLSA): The major finding of Bisitz and Hohmann (2006) is the improvement of speech intelligibility in modulated noise using instantaneous compression. These results could not be confirmed in this study for none of the noise level conditions. Fig. 5a shows the objective speech intelligibility test results in modulated noise. The performance for both settings with instantaneous compression for speech in noise (noise level 65 dB) is worse compared to the two other settings tested (LIN, CEN). The speech reception threshold (SRT) is reduced by about 1,5 dB for IC2 and by about 2,5 dB for IC1 compared to setting CEN. Both results are statistically significant. The speech test results for the noise level of 75 dB show a similar trend but – possibly because of the saturation effects – no statistical significance was achieved.

Speech tests in quiet (Rhymetest): The additional gain at low levels provided by IC1 and IC2 (as already discussed in section “categorical loudness scaling”) is reflected directly in the speech intelligibility performance in quiet. At a speech level of 50 dB, IC1 and IC2 show higher speech scores than CEN and LIN with the differences between IC2 and CEN being statistically significant (Fig.5b). For high speech levels of 80 dB the results are not so clear. The best result was achieved for the fastest compression condition (IC1) but still not statistically significant. All other conditions result in almost equal performance. The assumption that too little compression or even the absence of compression (see in particular condition LIN) leads to a decrease in speech understanding with increasing speech level could not be confirmed. This is true for all conditions.

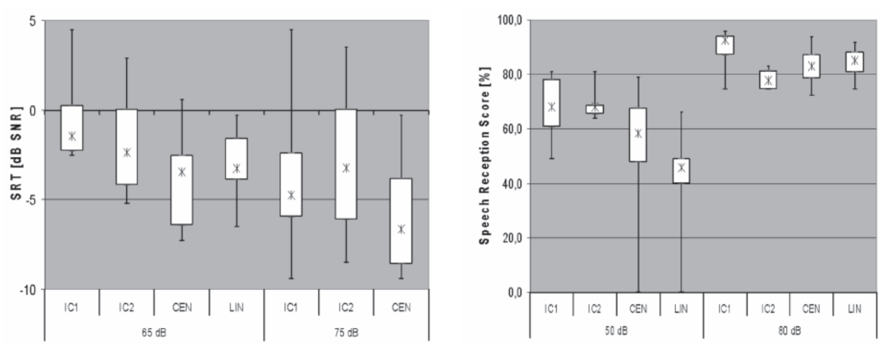

Fig. 5: Boxplots (Mean, Maximum, Minimum, 25- and 75 quartile) for speech intelligibility test results (OLSA = Speech Test in modulated noise, Rhyme Test = Speech Test in quiet) for all tested conditions (IC1, IC2, CEN,LIN).

Absolute Rating and Paired Comparison

The correct adjustment of loudness perception of broadband (static) noise using the IC approach has already been described in sections “categorical loudness scaling” and “speech test”. Here, the loudness perception proved to be better restored in the instantaneous compression condition than in the CEN and LIN condition, particularly for soft and loud input levels. This outcome is supported by the absolute ratings for attribute “loudness” as the sound samples in extreme loudness regions like “soft” (whispering at 45 dB) or “loud” (speech in traffic noise at 80dB) are rated best for both IC settings (Fig. 6a).

Looking at the results for the item “sound quality” two surprising effects can be observed:

Firstly, it is generally assumed that minimizing the time constant (IC1 and IC2) will reduce the sound quality substantially. Secondly, one would expect the same result if the loudness is successfully restored by a large amount of compression. The results of this study are in complete contradiction to these assumptions as neither the very short time constants nor the good performance in restoring loudness in case of both instantaneous compression approaches resulted in deterioration of the sound quality (Fig. 6b). In fact, the sound quality for “loud” and “soft” input levels is rated best for the IC conditions compared to CEN and LIN.

The results for middle level regions and both sound samples with dynamic level changes show no preference for one of the settings for “loudness” and “sound quality” except the LIN setting, which was rated badly mostly due to the large gain at high input levels.

For the items “overall impression” and “distortion” there are no differences between the conditions.

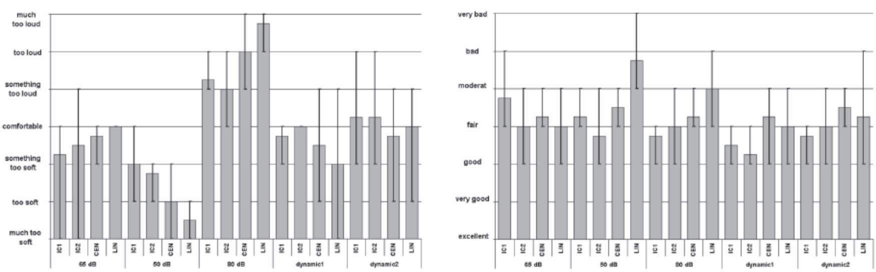

Fig. 6: Mean, Maximum and Minimum of the absolute subjective ratings of 5 sound samples and 4 test conditions for items “loudness” (a) and “sound quality (b) (all subjects).

The paired comparison tests showed no clear preference for one of the settings so that the results are neither discussed nor displayed in this paper.

Further interesting fact: the subjective rating of subjective speech intelligibility was conducted directly after performing the objective speech test (in quiet). This timing facilitates a good correlation between the subjective rating and the objective data.

CONCLUSIONS

The major finding of Bisitz and Hohmann (2006) was that speech intelligibility in modulated noise with instantaneous compression improves in comparison to the results achieved with linear processing. This finding could not be confirmed with the current implementation of the algorithm.

However, in principle, the concept of instantaneous compression allows for the restoration of loudness for a broader class of signals than traditional types of compression. The gain prescribed at low levels is more adequate than the gain provided by the state-of-the-art compression.

The better loudness restoration does not negatively affect the sound quality as predicted especially for sensorineural hearing impaired persons. These findings are in complete contradiction to the results of further studies (Souza, 2002) showing that the linear approaches generate the best ratings of sound quality.

Thus, the general assumption that instantaneous compression provokes unpleasant distortions has to be reconsidered.

The general idea of compensating for the loss of instantaneous compression in the cochlea with a similar system in the hearing aid, is supported by our results.

Further general findings:

The method to equalize loudness of different fittings by measuring the broadband loudness function and readjusting the overall gain accordingly works quite well and seems to allow for a better comparison of compression systems.

Future steps:

The present realization of instantaneous compression approach does not work for dynamic sounds with very abrupt level changes, especially with higher intensity levels. This may be due to an inaccurate implementation of the suppression effect used in the current approach to calculate the effective gain.

The results for the item "sound quality", which, contrary to expectations not only meet the same rating as the conventional compression schemes but also improve the ratings for soft and loud input levels, encourage continuing development of the novel compression approach with instantaneous processing.

REFERENCES

Bisitz, T., Hohmann, V. (**2006**). "Instantanfrequenzgesteuerte Multikanal-Dynamikkompression für Hörgeräte," In Fortschritte der Akustik - DAGA 2006, DEGA e.V., Berlin, 131-132.

Brand, T., and Hohmann, V. (**2002**). "An adaptive procedure for categorical loudness scaling," J. Acoust. Soc. Am. **112**, 1597-1604.

Giguère, C., and Smoorenburg, G. F. (**1999**). "Computational modeling of outer haircell damage: impülications for hearing and signal processing," Psychophysics, Physiology, and Methods of Hearing, T. Dau, V. Hohmann, and B. Kollmeier, Eds.: 155-164.

Grimm, G., Herzke, T, Berg, D., and Hohmann, V. (**2006**). "The Master Hearing Aid: A PC-Based Platform for Algorithm Development and Evaluation," Acta Acustica **92**, 618-628.

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

Herzke, T., and Hohmann, V. (**2005**). "Effects of instantaneous multiband dynamic compression on speech intelligibility," EURASIP JASP 2005(18), 3034-3043.

Hohmann, V., and Kollmeier, B. (**2007**). "A nonlinear auditory filterbank controlled by sub-band instantaneous frequency estimates. International Symposium on Hearing

– ISH 2006, Cloppenburg, Germany, Springer.
Hohmann,V. (**2007**). “Compensation of hearing deficiencies in the inner ear,” proceedings of EFAS/DGA joint congress, Heidelberg, DGA - Deutsche Gesellschaft für Audiologie.
Kiessling, J., Müller, M., and Latzel, M. (**2006**). “Fitting strategies and candidature criteria for unilateral and bilateral hearing aid fittings,” Intern. J. of Audiol. **45,** 53-62.
Souza, P.E. (**2002**). “Effects of compression on speech acoustics, intelligibility, and sound quality,” Trends Amplif .**6,** 131-165.
Verschuure, J., Maas, A. J. J, Stikvoort, R.M., de Jong, R.M., Goedegebure, A., and Dreschler, W.A . (**1995**). “Compression and its effect on the speech signal,” Contribution at Third International Hearing Aid Conference Iowa.
Wagener, K., Kühnel, V., and Kollmeier, B. (**1999**). “Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests,” Zeitschrift für Audiologie/ Audiological Acoustics, **38**, 4-15.
Wagener, K.C., and Brand, T. (**2005**). “Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters,” Intern. J. Audiol **44**, 144-157.
Wagener, K.C., Brand, T., and Kollmeier, B. (**2006**). “The role of silent intervals for sentence intelligibility in fluctuating noise in hearing-impaired listeners,” Intern. J. Audiol. **45,** 26-33.
v. Wallenberg, E.L., and Kollmeier, B. (**1989**). “Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten,” Audiologische Akustik, **38**., 50-65.