The effects of compression ratio and release-time on loud speech and noise signals, processed by a simulated non-linear hearing aid

ERIK SCHMIDT[1,2]

[1] *Widex A/S, Ny Vestergårdsvej 25, 3500 Værløse, Denmark*

[2] *Dept. of Acoustic Technology, Technical University of Denmark, 2800 Lyngby, Denmark*

This study investigated the effect of compression ratio and release time on hearing aid wearers' impressions of loud input signals. Two speech and noise signals, differing in spectra and signal-to-noise ratio, were processed in a model compressor with sixteen combinations of compression ratio and release time. The RMS input level of the signals was 75 dB SPL. Subjects rated the processed signals on categorical scales, in regard to loudness, speech clarity, noisiness and overall acceptance. With a compression ratio between 1.5:1 and 2:1 and a release time of 4000-ms, the highest degree of speech clarity and the lowest possible noisiness was achieved, while still maintaining a positive rating on the acceptance-scale. When shorter release times of 40 and 400-ms were used, ratings of acceptance declined when the compression ratio was 3:1 or greater. Thus, the preferred setting appears to be long release times in combination with a low compression ratio - providing the listener with a realistic loudness for the signal. When faster regulation is needed the compression ratio should not exceed 3:1.

INTRODUCTION

Most fitting rationales for non-linear hearing aids apply Wide Dynamic Range Compression (WDRC) to compensate for loudness recruitment. Fitting rationales may be based on models predicting loudness (e.g. Moore *et al.,* 1997), or on empirical studies of loudness perception in a hearing impaired population (e.g. Pascoe, 1988).

From a sound quality-point of view, it makes sense to provide the hearing aid user with some degree of level-variation in the output from the hearing aid, which reflects the overall level fluctuations in the surroundings. In earlier studies, investigating preferred listening levels for soft and loud sounds, a preference has been found for placing such sounds closer to the most comfortable loudness level (Neuman *et al.,* 1995; Smeds *et al.,* 2004). A consequence of this would be to use a rather high compression ratio, at least for high input levels. In this way the level range of loud sounds would be narrowed in, before being presented to the hearing aid user.

But the sound processing of level differences not only depends on the compression ratio, but also on the the attack- and release-times, the compression threshold and the number of channels in the hearing aid. In combination with the type of input signal, the setting of these parameters will affect the effective compression ratio, the overall

output level and the short-term spectral and temporal level differences of the processed signal (Kuk and Ludvigsen, 1999). This again could affect the listeners' impressions of loudness, sound quality and speech intelligibility. In the present study the combined effects of three different release times and six compression ratios were investigated.

Specifically, the study aimed at answering the following research questions:

- Do differences in speech spectra and signal-to-noise ratios between signals cause a significant difference in listeners' perceptions of the signals, when processed with the same compression settings?

- Which combination of compression ratio and release-time provides the "best impression" of speech intelligibility and user acceptance (and the lowest noise nuisance) in the two signals?

METHOD

Test signals

Two different input signals containing speech and party noise were prepared:

(1) Male speaker at loud vocal effort and party noise (0 dB SNR).

(2) Male speaker at normal vocal effort and party noise (+15 dB SNR).

The choice of vocal efforts and signal-to-noise ratios (SNR) in the two signals was made to simulate two realistic situations, where the hearing aid user listens to loud speech in background noise. In the first signal, the listener engage in a conversation with a person standing in front of him, while being at a noisy party. The second signal represents a situation where the HA-user listens to a speaker's voice, coming from a radio or TV-set at a high volume setting. Fig. 1 and 2 shows the long term spectra for speech and noise in signal (1) and (2).
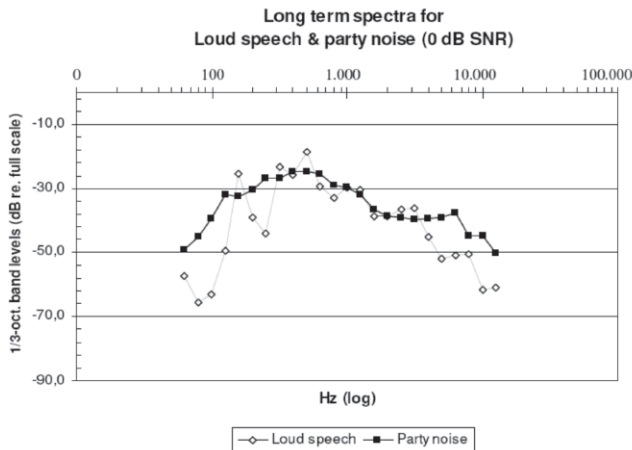
Fig. 1: Long-term spectra (1/3-octave levels) for male speech at loud vocal effort and for the party noise used in signal (1). The SNR between speech and noise signals is 0 dB.
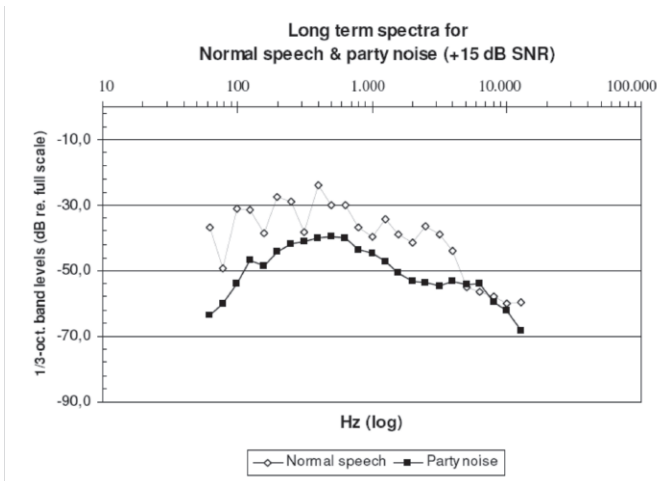
Fig. 2: Long-term spectra (1/3-octave levels) for male speech at normal vocal effort and for the party noise used in signal (2). The SNR between speech and noise signals is +15 dB.

The two input signals were compressed off line in the three channel compressor implemented in MATLAB. signals were compressed with six different compression ratios (CR): 1:1 (linear condition), 1.5:1, 2:1, 3:1, 5:1 and 10:1 and three different release-times (RT) of 40, 400 and 4000 ms. The attack-time was always held fixed at 10 ms. The three release-times were chosen to reflect a fast syllabic compressor, a slow compressor and a very slow compressor. Release-times in the range from 40 to 4000 ms are also seen in typical commercial hearing aids.

The input levels to the compressor for speech in the two test-signals were adjusted to 75 dB SPL RMS-level (i.e. the level specified for loud vocal effort in ANSI-S3.5). The anchor points in each channel were adjusted, such that a speech input of 62 dB SPL would receive the same gain, independent of the given compression settings.

This set-up of the compressor simulated a commercial hearing aid with a handle that varies the degree of gain (or compression ratio) for high input levels - while always keeping the same gain for a normal speech input of 62 dB SPL. The gain applied to the two test signals would vary depending on the given compression ratio – that is, gain would be reduced with increasing ratio.

The broadband static input-output characteristics of the compressor are shown in Fig. 3. The overall RMS input level of the two input signals (75 dB SPL), and for the normal speech signal used for adjusting the anchor-points (62 dB SPL), are encircled on the abscissa.

The compressor produced eighteen versions of each signal, with different output RMS-levels and compression characteristics, depending on the setting of ratio and release time in each condition.
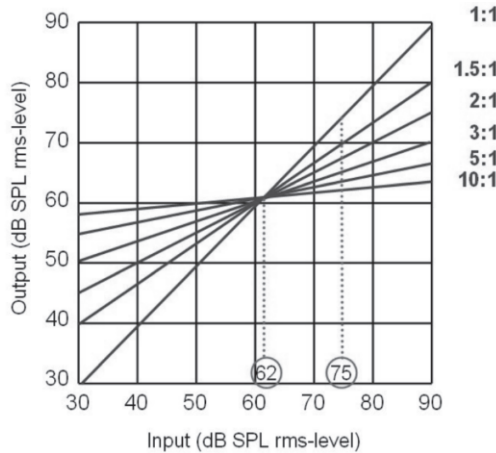
Fig. 3: Broadband input-output characteristics of the compressor. The RMS input levels for speech in the two test signals was 75 dB SPL. The anchor-points in each channel were adjusted relative to a normal speech input at 62 dB SPL RMS-level.

Setup for listening eksperiment

Test signals were presented to seven hearing impaired listeners (mean age of 74.6 years) in a free field set up. The presentation level was calibrated such that the RMS-level for speech in the uncompressed signals (1:1) was 75 dB SPL at the position of the listener. In all other signals, the level would be lower depending on the compression settings used.

All listeners had moderately sloping hearing losses and were experienced hearing aid users. For this experiment they wore binaural BTE hearing aids (Widex Senso Diva), fitted linearly according to NAL-RP for the individual hearing loss (Byrne *et al.,* 1991). In this way the compressed signals would be presented above the Most Comfortable Level, in the upper part of the auditory range in each listener.

The eighteen compressed versions of the two input signals were presented three times in randomized order to each test subject, giving a total of 48 trials. Subjects rated each trial on four categorical scales: (1) the overall loudness of the signal, (2) the clearness of the speech in the signal, (3) the noisiness in the signal and (4) how acceptable the to judge how acceptable this hearing aid setting would be, if they needed to listen to the sound for 10 minutes. The four scales are shown in Fig. 4.
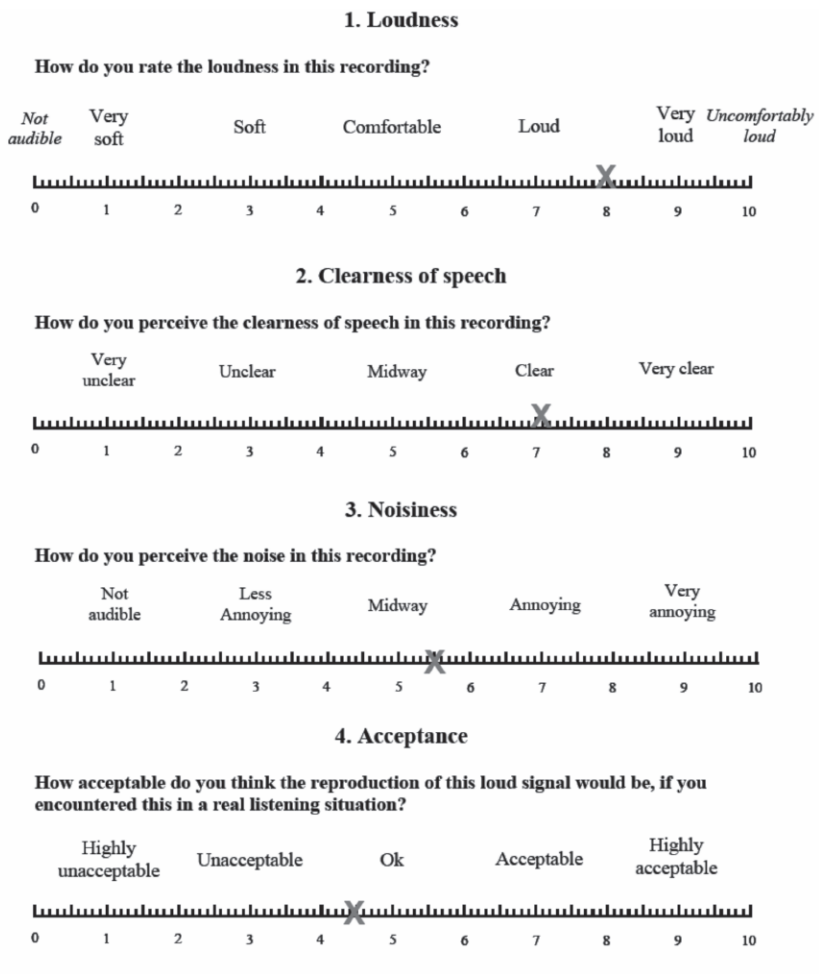
Fig. 4: The four categorical scales used for the subjective rating of the processed test signals (English translation).

RESULTS

In Fig. 5 (a-g), mean ratings on the four scales are shown for the three release times, as a function of compression ratio. Brackets indicate no significant difference in means (p > 0.05) between release times.

On each of the four scales, a mixed model analysis of variance was carried out. The ANOVA showed a significant effect of the compression setting (i.e., the combination of compression ratio and release-time) on all scales and in both signals.
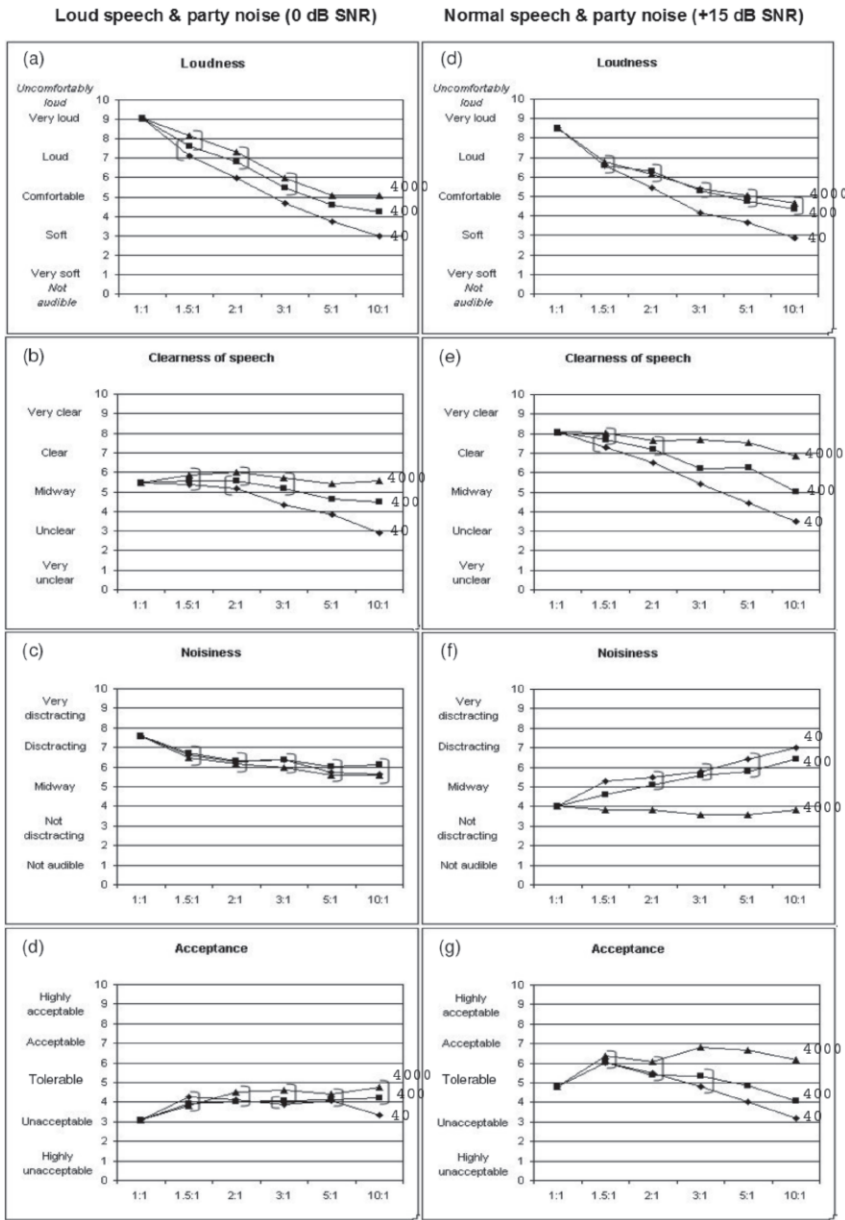
Fig. 5 (a-g): Mean ratings of loudness, clearness of speech, noisiness and acceptance.

DISCUSSION

From Fig. 5 it is seen how the shortest release time of 40 ms generally produce the lowest ratings of *loudness* (a, d) and *speech clearness* (b, e). At the higher compres-

sion ratios, the means obtained at the three release times generally differ significantly from each other. Highest ratings of speech clearness are obtained in both signals at the 4000 ms release time. Ratings of *noisiness* (c, f) do not differ from each other in signal (1) *Loud speech and party-noise* (0 dB SNR), whereas in signal (2), *Normal speech and party-noise* (+15 dB SNR), the longest release time stands out as the one producing the lowest noise nuisance.

The ratings of *speech clearness* made for signal (1) may be linked to the loudness, in case of the short release time of 40 ms. But at RT = 400 ms and 4000 ms, the more even ratings close to the "Midway" category seem to be related to the influence of the noise. The generally lower impression of speech clearness in this signal may also influence listeners' overall acceptances of the signal – in this case receiving a rather low acceptance in all conditions. This relationship was also noted by Preminger and Van Tasell (1995), who found that changes in sound quality can only be measured separately from changes in speech intelligibility when the speech is clearly audible above the noise - i.e. when the signal-to-noise ratio is positive.

In signal (2), the influence of the release time on the speech and noise signals is more clearly seen. In the ratings of *speech clearness*, the reduction in means with increasing ratio is most prominent at 40 ms and 400 ms. Even though the speech ratings may be linked to the loudness of the signal, they also seem related to the release time and its influence on the speech signal. In combinations of a short release time and high compression ratio, the dynamic range of signal (2) becomes narrower. This could mean that the natural intensity relationship between soft and loud speech components becomes distorted. At the same time, the ratings of noisiness increase with compression ratio, which is the opposite of the pattern in signal (1). This may be due to the increased gain for noise at the lower side of the anchor point in the compressor (fig. 3). This makes the noise increasingly audible in the speech pauses and may thereby also affect the intelligibility of the speaker. At the longest release times of 4000 ms, the noise-level in speech pauses is re-established to the original SNR, and ratings of speech clearness remains high whereas noisiness-means remain at a low level. This is also reflected in the ratings of acceptance where equally high ratings are given in all condition at this release time.

Two different research questions were investigated in this experiment. Firstly, it was found that difference in signal-to-noise ratio between signals did result in significant differences in ratings on the four scales. The favorable SNR of +15 dB in signal (2) *Normal speech and party-noise*, made the influence of the release time clearer compared to signal (1), *Loud speech and party-noise*, where the poor SNR seemed to even out ratings in all conditions. Secondly, the most optimal combination of compression ratio and release time appeared to be 2:1 for signal (1) and 1.5:1 for signal (2) – in both cases when using a long release time of 4000 ms. With this combination, the highest degree of speech clarity and lowest possible noisiness was achieved, while still maintaining an acceptance of "Tolerable" and a realistic loudness for the two signals.

The results of this experiment are in accordance with earlier studies that investigated

the effects of compression ratio and release time. Arguments against the use of high compression ratios in combination with shorter time constants have been put forward in studies focusing on different attributes of sound quality. Neuman *et al.* (1998), using a single channel compressor, found that compression ratio had the greatest impact on subjective ratings of sound quality made by hearing-impaired listeners. And when the compression ratio was 3:1, a short release time of 60 ms gave significantly lower ratings of sound quality, compared to release-times of 200 and 1000 ms. This was especially the case for signals with poor signal-to-noise ratios. Overall, good sound quality was preserved when the compression ratio was below 3:1.

CONCLUSION

In the present experiment, ratings of acceptance for the *Normal speech and party-noise* fell below the "Tolerable"-category at the shorter RT's of 40 and 400 ms, when the ratio was 3:1 or greater. On the contrary, acceptance-ratings made at RT = 4000 ms was high in all cases, down to a ratio of 1.5:1. Thus, for a loud speech and noise signal, even at a favorable SNR, the preferred setting seem to be a long release time in combination with a low compression ratio - providing the listener with a realistic loudness for that signal.

In case a faster regulation is needed (e.g. when a sudden increase in the input level occurs), the compression ratio should not exceed 3:1, because this will negatively affect the perceived noisiness and clearness of speech and thereby decrease user-satisfaction.

REFERENCES

Byrne, D., Parkinson, A., Newall, P. (**1991**). "Modified hearing aid selection procedures for severe/profound hearing losses," In The Vanderbilt Hearing Aid Report II (ed. Studebaker *et al.*), Parkton, MD: York Press.

ANSI-S3.5 (**1997**). "American National Standard methods for the calculation of the Speech Intelligibility Index," New York, American National Standards Institute, Inc.

Kuk, F. K., Ludvigsen, C. (**1999**). "Verifying the output of digital nonlinear hearing instruments," Hearing Review, **6**(11), p. 35, 36, 38, 60, 62, 70.

Moore, B. C. J., Glasberg, BR., Baer, T. (**1997**). "A model for the prediction of threshold, loudness and partial loudness," J. Audio Eng. Soc., **45**(4), 224-240.

Neuman, A., Bakke, M. Mackersie, C., Hellmann, S., Levitt, H. (**1998**). "The effect of compression ratio and release time on the categorical rating of sound quality", J. Acoust. Soc. of Am., **103**(5), 2273-2281.

Neuman, A. C., Bakke, MH., Hellman, S., Levitt, H. (**1995b**). "Preferred listening levels for linear and slow-acting compression hearing aids," Ear & Hearing, **16**(4), 407-415.

Pascoe, D. P. (**1988**). "Clinical measurements of the auditory dynamic range and their relations to formulas for hearing aid gain," In Proceedings of the 13th Danavox Symposium, edited by J. Hartvig Jenson, 129-151.

Preminger, J., Van Tasell, DJ. (**1995**). "Quantifying the relation between speech quality and speech intelligibility," J. Speech Hear. Res., **38**, 714-725.

Smeds, K., Keidser, G., Zakis, J., Dillon, H., Leijon, A., Grant, F., Convery, E., and Brew, C. (**2004b**). "Preferred overall loudness. II: Listening through hearing aids in field and laboratory tests," in Smeds K., Less is more? Loudness aspects of prescriptive methods for nonlinear hearing aids. Doctorial Thesis, Royal Institute of Technology, Stockholm, Sweden.