Ear & Hearing
Issue: Volume 17(2), April 1996, pp 162-175
Copyright: © Williams & Wilkins 1996. All Rights Reserved.
Publication Type: [From the Third International Hearing Aid Conference, 1995, The University of Iowa]
ISSN: 0196-0202
Accession: 00003446-199604000-00008

[From the Third International Hearing Aid Conference, 1995, The University of Iowa]

Compression and its Effect on the Speech Signal

Verschuure, J.; Maas, A. J. J.; Stikvoort, E.; de Jong, R. M.; Goedegebure, A.; Dreschler, W. A.

Author Information

Audiological Centre, Erasmus University Rotterdam, and University Hospital Rotterdam (J.V., A.J.J.M., R.M.J., A.G.), Rotterdam, The Netherlands; Philips Research Laboratory (E.S.), Eindhoven, The Netherlands; Audiological Centre, Academical Medical Centre (W.A.D.), Amsterdam, The Netherlands
Address for correspondence: J. Verschuure (Hans), Audiological Centre, ENT Department, University Hospital Rotterdam-Dijkzigt, Dr. Molewaterplein 40 NL-3015 GD Rotterdam.
Received June 23, 1995; accepted December 4, 1995

Abstract

Compression systems are often used in hearing aids to increase the wearing comfort. A patient has to readjust frequently the gain of a linear hearing aid because of the limited dynamic hearing range and the changing acoustical conditions. A great deal of attention has been given to the static parameters but very little to the dynamic parameters. We present a general method to describe the dynamic behavior of a compression system by comparing modulations at the output with modulations at the input. The use of this method resulted in a single parameter describing the temporal characteristics of a compressor, the cut-off modulation frequency. In this paper its value is compared with known properties of running speech. A limitation of this method is the use of only small modulation depths, and the consequence of this limitation is tested. The use of this method is described for an experimental digital compressor developed by the authors, and the effects of some temporal parameters such as attack and release time are studied. This method shows the rather large effects of some of the parameters on the effectiveness of a compressor on speech. This method is also used to analyze two generally accepted compression systems in hearing aids. The theoretical method is next compared to the effects of compression on the distribution of the amplitude envelope of running speech, and it could be shown that single-channel compression systems do not reduce the distribution width of speech filtered in frequency bands. This finding questions the use of compression systems for fitting the speech banana in the dynamic hearing range of impaired listeners.

Many patients complain of problems with hearing resulting from reduced dynamic range. Listening to a discussion involving multiple speakers, each talking at a different level, or listening under different acoustical conditions requires patients frequently to readjust the volume control on their hearing aids. In some patients with large losses, and thus small dynamic hearing range, even the dynamics of the speech signal itself cause problems; amplifying the weak parts of the speech to audible levels causes the strong parts to be uncomfortably loud.

Compression systems have been used to help patients in this respect for many years. Recently the number of available automatic hearing aids has surged with a number of different algorithms of automatic gain control (AGC), such as Multifocus and K-amp. The algorithms have sometimes opposing effects on the sound, and all manufacturers claim positive effects of better speech intelligibility and improved comfort. However, research on compression systems with hearing impaired people has shown that improved wearing comfort is often reported but quite often is accompanied by poorer measured speech intelligibility (see for discussion Verschuure et al., 1993; Verschuure & Dreschler, 1993). The manufacturers claim good sales, and this surge in sales of compression hearing aids without research supporting their efficacy is, therefore, cause for concern. The development of protocols to adjust the compression systems to a patient without clearly stating the limitations of such systems should be questioned in particular because patients cannot always distinguish between wearing comfort and speech intelligibility. During the Third International Hearing Aid Conference in Iowa City, some of these procedures were discussed, including the procedure developed by the Independent Hearing Aid Fitting Forum (IHAFF) and methods referred to as Desired Sensation Level(Input/Output) (DSL [i/o]), Visual Input/Output Locator Algorithim (VIOLA),FIG 6 as developed by Killion, and the RAB procedure(Ricketts & Bentler).

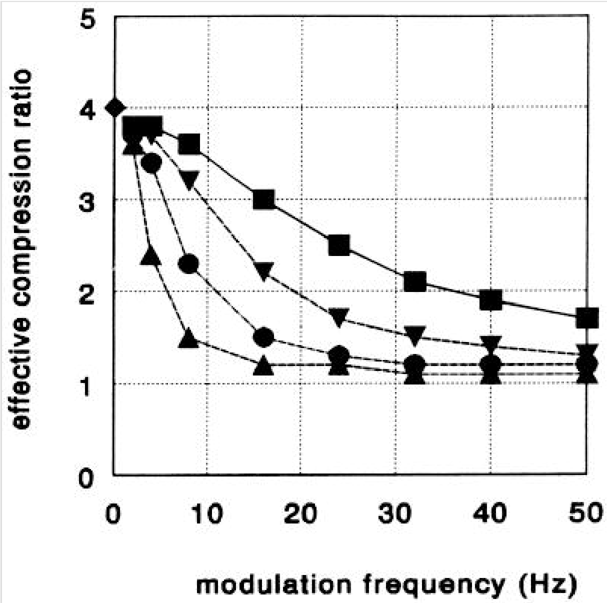

Figure 6. The effect of release time on the effective compression ratio for a smoothed phonemic compressor. The release times are 15 msec (squares), 30 msec (triangles down), 60 msec (dots), and 120 msec (triangles up).

It is the purpose of this paper to analyze what a compression system does to a speech signal, and in particular to the speech parts relevant for speech recognition, and to specify the parameters of a compression system that are relevant for different compression goals. Our approach involves a theoretical method to measure the effectiveness of compression systems and the comparison of the outcome measure of this measurement with that of the level distribution of a "normal" speech signal.

Description of Speech and Compression Systems

Speech Signal

A speech signal can be described in physical terms as a modulated spectrum. Both aspects, spectral information and temporal information, are relevant. The two aspects are not equally important for all parts of speech (Verschuure et al., 1993). Vowels, semivowels, and nasals require good spectral resolution (separate detection ofF_1 and F_2) while there is little information in the (almost absent) modulations. Fricatives and plosives, on the other hand, are strongly modulated signals differing mainly in time structure (e.g., the gap before the plosive); only crude spectral analysis in required.

The average spectrum of speech seems well established, even as a function of level (Pavlovic, 1992). It can be described roughly as having a peak around 400 Hz and falling off above 500 Hz at a rate of about 10 dB/oct.

The modulations can also be represented by a spectrum, the amplitude-modulation spectrum (Plomp, 1983). The relevant frequencies of this spectrum range roughly from 0.1 to 40 Hz. The modulation spectrum can be determined per octave band of the speech spectrum. Such an analysis shows that the modulation spectrum hardly depends on the octave band of the speech spectrum from which it is taken. The frequency of maximum modulation is around 3 Hz, and the maximum amount of modulation is found in the frequency band around 1 kHz (Fig. 1). For the high-frequency band (4 kHz), the maximum shifts somewhat toward a higher modulation frequency. (5 Hz).

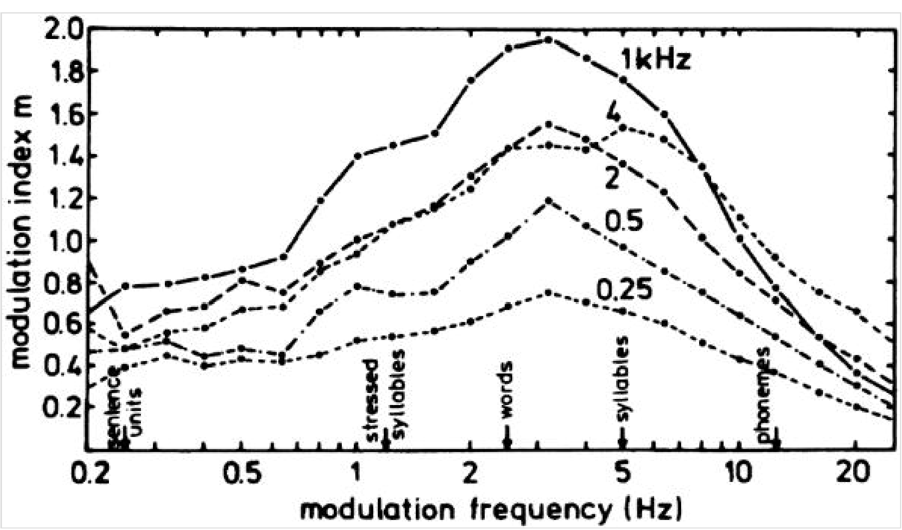

Figure 1. Modulation spectrum of running speech per octave frequency band (Plomp, 1983). Note that$m$ has a maximum value of 2 because they defined m as peak level/average level. In the more traditional definition, all values would be halved.

The figure also shows the modulation frequency that can be associated with phonetic entities. The stress pattern has a modulation frequency of up to about 1 Hz, words cause a modulation of around 2.5 Hz, syllables of around 5 Hz, and phonemes of around 12 Hz. The detection of the silent gap in a plosive (duration 30 to 60 msec) requires an even higher modulation frequency(up to about 30 Hz).

In hearing-impaired individuals, frequency resolution is reduced through excessive upward spread of masking, and the temporal resolution is reduced because speech is presented closer to the threshold where the effective temporal resolution is poorer than at higher levels. These facts and the possible consequences for speech recognition have been discussed by Verschuure et al. (1993).

Compression systems have an effect on the temporal patterns of speech. They may produce distortion by overshoot (temporal distortion) and may affect the frequency response (spectral distortion) if multi-channel systems are used. These effects may distort the speech signal and the distortion may affect the intelligibility of speech for a hearing-impaired person. We therefore need a detailed description of the effects of a compression system on speech and an analysis of how the distortions may or may not be detected by the hearing-impaired listener.

Compression System

Compression systems in hearing aids can be characterized in terms of a number of parameters (ANSI S3.22.1987):

1. Compression ratio. This measure represents the effectiveness of the compressor: in hearing aids it is usually given as the ratio between a rise in input and the resulting rise in output level (both in dB). A compression ratio of 4 means that the output is raised by 1 dB for each rise in input level of 4 dB.
2. Bandwidth. This measure specifies the frequency range over which the compressor is active.
3. Number of bands. This measure specifies the number of frequency bands into which the signal is split. Each frequency band may have different compression parameters.
4. Compression threshold. This is the knee-point of compression, giving the input or output level below or above which a particular compression circuit is activated.
5. Control signal. This measure is derived from either the input or the output signal, and it describes the feedback signal that controls the compression.
6. Attack time. The attack time is a measure of the speed with which the amplification is adjusted after a rise in input signal level. It may be specified as a 1/e, 10 to 90% decay time, or any other time constant. For hearing aids, attack time is defined as the time required for the output of a hearing aid to reach 2 dB of the steady-state level after a sudden increase in level from 55 to 80 dB SPL when stimulated by a sine wave of 2 kHz.
7. Release time. Release time is a measure of the speed with which the amplification is adjusted after a drop in input signal level. It can also be specified as a 1/e, 10 to 90% decay time, or as the time required for a 20 dB drop from a certain level. It is defined for hearing aids as the time required for the output of a hearing aid to reach 2 dB of the steady-state level after a 2 kHz sine wave is decreased in level from 80 to 55 dB SPL.Parameters 1 to 5 clearly deal with the amplitude of the signal. Parameters 6 and 7 deal with the temporal pattern. In the compression system with suppressed overshoots that we designed (Verschuure et al., 1993, Verschuure, Prinzen, & Dreschler, 1994), there is one additional parameter to be dealt with:
8. Delay time. This parameter specifies the delay in milliseconds between the speech input signal and the compression control signal. Such a delay may be used to anticipate events in the signal such as jumps in level.

It is clear that a great number of parameters are required to characterize the function of a compression system. However, knowledge of these parameters is insufficient to predict the effectiveness of a system or how the system will affect the speech signal. In fact, we propose in this paper a simple method to describe the effectiveness of a compressor on the speech signal.

For steady-state signals, only spectral effects are important, and the spectral effects can be computed from the input spectrum, the compression threshold, and the compression ratio in each frequency band. However, speech is a modulated signal, and temporal effects should also be taken into account. A compression system is designed to reduce modulations to the extent given by the compression ratio. The effectiveness in modulation terms can be given by modulation frequencies for which the modulations are effectively reduced and by modulation frequencies for which the reduction fails. The effectiveness can be measured by determining the modulation of a signal at the output given a certain modulation at the input (Verschuure, Maas, Stikvoort, & Dreschler, 1992). The method is similar to the one independently developed by Stone and Moore (1992), and the approach is similar to the approach used by Steeneken and Houtgast (1980) for measuring speech intelligibility in transmission lines, except that it now is used to measure the desired modulation reductions produced by hearing aids.

Spectral Analysis of Effective Compression

The effectiveness of a compressor can be tested by using a sine wave with carrier frequency ([omega]$_c$) and a modulating frequency([omega]$_m$). The amplitude of the carrier is a_o and the amount of modulation is given by the modulation index *m*. Comparison of the modulation index at input and output of the compressor gives a measure of the effectiveness of the compressor. The amount of modulation can be determined by accurately analyzing the spectrum of the signal. The theoretical issues are presented in the Appendix; here only the main points necessary to understand the method are presented.

An amplitude modulated sine wave can be described by:Equation (1) The spectrum of the signal consists of a carrier and two side bands. The difference in level between the two side bands and the carrier is in dB: Equation (2)Equation 2 shows that the comparison of the spectral level of the central component with that of the side bands provides a direct measure of the modulation depth. We now can determine the spectrum of a modulated sine wave at the input and at the output. If we now define the modulation depth at the output by *Rm*, we can compute the effective compression ratio R_{eff} as:Equation (3) There is a limitation to the method. The compressor works in the logarithmic (dB) domain, whereas spectra are determined as linear measures. A sinusoidal modulation is distorted by a logarithmic compressor. The distortion shows up in the spectrum as higher-order side bands and extra contributions to the carrier and first-order side bands. The distorations are small and negligible for small modulation depths, and for m < 0.45 (see Appendix) the errors are smaller than 5%. For this value the harmonics of the modulation side bands are more than 40 dB down for the second harmonic and more than 50 dB down for the third harmonic. On the other hand, for these modulation depths, the differences in level can be determined with a high enough accuracy to warrant the use of the method.

$$x(t) = a_0(1 + m\cos(\omega_m t))\cos(\omega_c t)$$

Equation 1

$$\Delta S = 20\ \log m - 6$$

Equation 2

$$R_{eff} = 10^{\frac{\Delta S_0 - \Delta S_i}{20}}$$

Equation 3

Analysis of the Method

In the present paper, we analyze the usability of the method to show its applicability in evaluating the temporal characteristics of compressors and to point to the possible translation into speech features as shown in Figure 1. We do so in four parts.

1. We study the theoretical limitation of the method by applying the method beyond its limitations.
2. We study the effects of parameter changes on the effectiveness of a compressor. We use an experimental digital compressor design allowing for independent changes of parameters. The effectiveness of the compressor design is also studied.
3. We compare designs of compressors as they are commercially available in hearing aids with our experimental compressor.
4. We compare the modulation technique measurements with the results of the amplitude distribution of running speech.

Compressor

We used a compressor implemented on a digital signal processor DSP 56001 for testing the method. The design has been described in detail by Verschuure and Dreschler (1993),Verschuure et al. (1993), and Verschuure, Prinzen, and Dreschler (1994); the block diagram is shown in Figure 2. The principle of the design is a two-channel processor. The compressor works only on the signal in the second channel, which is a filtered part of the speech signal(FIR-filter), usually the high-frequency part of the signal. The filter shape is computed as the inverse of the hearing loss. The signal passing through the filter is then delayed. The compressor control signal is taken directly from the input signal (it may be additionally filtered) and rectified. The rectified amplitude signal is compared with a table to determine the required amplification of the signal passing through the FIR-filter and delay. The signal passing through the unfiltered linear path is added to the compressed and filtered signal after a proper delay and in such a way that the frequency response of the total system for a signal presented near the level of maximum intelligibility is a half-gain response. The design is made in such a way that this is also the response for linear amplification. A stylized frequency response is shown in Figure 3. The delay is used to suppress overshoots at the onset of a louder part of the signal. A peak-hold circuitry is used to suppress overshoots at the offset of the signal.

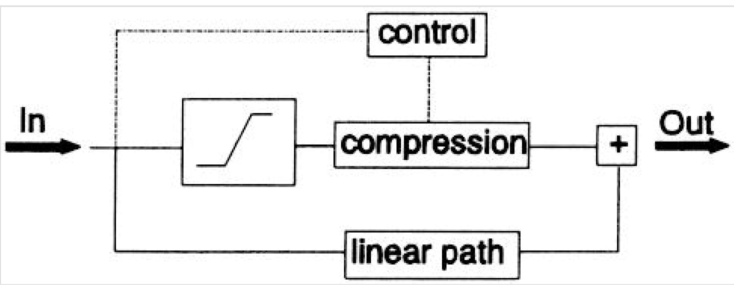

Figure 2. Block diagram of the smoothed compressor implemented on a DSP 56001. The smoothing is a realized by an extra time delay in both signals' paths to facilitate reduction of the overshoots.

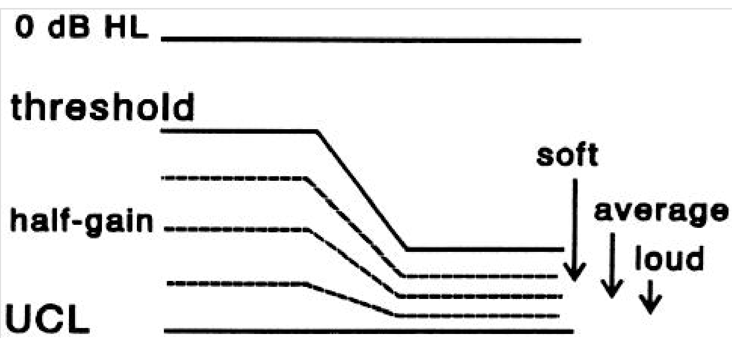

Figure 3. Stylized frequency responses obtained with smoothed compression as a function of level. The half-gain response coincides for all compression conditions including linear amplification.

The action of this compressor results in a system in which the higher-level parts show no overshoot (temporal distortion). All temporal distortion is transferred to the low-level parts of the signal where it most probably will not be detectable for a hearing-impaired person because of poorer temporal resolution and threshold effects (Verschuure et al., 1993).

Very high level sounds will flatten the frequency response. The compressor reduces the amplitude of the filtered part of the signal to such an extent that the signal has a lower level than a signal passing through the linear path. This may cause problems with the upward spread of masking. To overcome these problems, an anti-upward-spread-of-masking filter (anti-USOM filter) is implemented that puts a 6 dB/oct slope to the frequency response for high-level input sounds for frequencies above 0.5 kHz

The compression ratios of the implemented system can be set at 1 (linear), 2, 4, and 8. If not stated otherwise, the attack time was set at 5 msec, the release time at 15 msec, the delay time at 3 msec, and the compression threshold at -70 dB versus maximum input level. The static input-output characteristics showed the compression ratio to be correct over a range of at least 65 dB. Speech was presented at an RMS level of -22 dB versus maximum input level, leaving enough headroom to avoid high distortion.

The effects of the compressor and the choice of parameters on speech intelligibility have been described elsewhere (Verschuure & Dreschler, 1993; Verschuure et al., 1993; Verschuure, Prinzen, & Dreschler, 1994) and will not be repeated here.

Measurement Procedure

We used a Hewlett Packard Dynamic signal analyzer type 35665A for the analysis. The span was adjusted to 50 Hz or 100 Hz with a resolution of 400 lines. The dynamic range in fast Fourier transform mode was 72 dB. We analyzed the effectiveness of the compressor for modulation frequencies of 2, 4, 8, 16, 24, 32, 40, and 50 Hz.

Limitation of Method: Modulation Depth

This method of spectral analysis can be used effectively only if the test signal has a small modulation depth (Appendix). We wanted to check the theory by determining the effective compression ratio for various modulation depths. The carrier frequency was set at 2 kHz, a frequency in the middle of the band where the compressor was active, and the compression ratio was set at 4. In Figure 4 we show an example of a set of curves determined for modulation depths of 20 (squares), 40 (triangles down), 60 (dots), and 90% (triangles up). The theory predicts that only the smallest two modulation depths give reliable results within a 5% error (absolute error of 0.2). The point at 0 Hz was taken from the static compression curve.

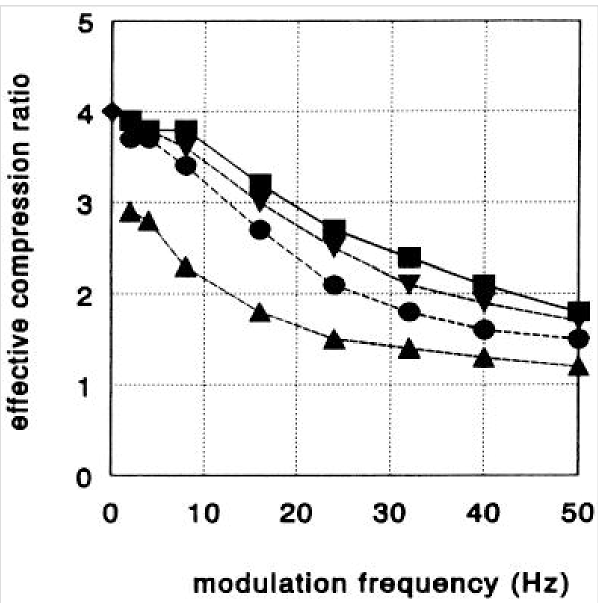

Figure 4. Analysis of the limitation of the method by determining the effect of modulation depth on the measurement of effective compression. The modulation depths are 20% (squares), 40% (triangles down), 60% (dots), and 90% (triangles up).

We saw that the difference between the 20% and the 40% curve was within the margin of error and that the other curves deviated more, as was expected. We concluded that the method works well within the theoretical limitations of using modulation depths smaller than 45%. In all further experiments, we used a modulation depth of around 40%.

Speech involves much higher modulation depths than 45%. It is usually assumed that modulations of up to 30 dB are relevant for speech intelligibility for speech spoken without much stressing and intonation(Speech Transmission Index, Articulation Index, and Speech Intelligibility Index calculations use this dynamic speech range). The method only allows the determination of the effective compression ratio around a certain level. If we want to see whether the total range of speech levels is effectively compressed, we have to determine the effective compression ratio at levels over the range relevant to speech, at least 30 dB apart. The result of the measurement of the effective compression ratio at 2 kHz for a modulation depth of 40% at the levels -10 and -40 dB versus maximum level is shown in Figure 5. The levels represent the maximum and minimum relevant levels of speech processed by the compressor. The compression threshold was set at -70 dB versus maximum level, so both levels are well within the compression range.

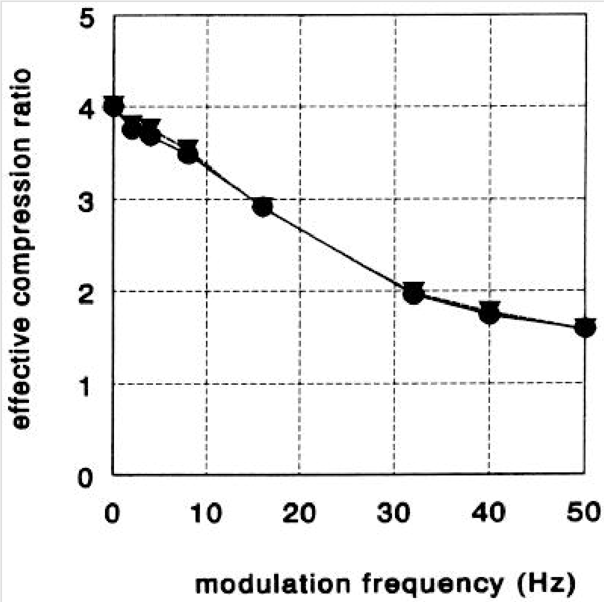

Figure 5. The effect of level on the measurement of the effective compression ratio. The measurements were done at -10 dB versus maximum level (dots) and at -40 dB versus maximum level (triangles), levels 30 dB apart. Speech usualy is presented at an RMS level of -22 dB versus maximum level, placing the upper and lower limit of relevant speech signals near -10 and -40 dB versus maximum level.

The analysis method shows that with a logarithmic compressor (a compressor working in the dB domain), distortions of the waveform do appear. This implies that compressors in hearing aids distort the amplitude envelope in a way showing up in the described analysis as higher harmonics of the side band for modulation depths comparable to the modulations in speech. These distortions may influence speech intelligibility in an unpredictable way, requiring the use of speech intelligibility tests when compressors are used. The limitation of the method to 45% modulation reflects the amplitude envelope distortion of logarithmic compressors but is no indication of limited use of the described analysis method.

Figure 5 shows that there is no effect of level within this range. The method is not sensitive to level changes as was expected as long as the compression threshold is sufficiently low in level.

We have to define a simple measure of effectiveness if we want to make comparisons between systems. The figures show a kind of low-pass filter response for modulation frequencies with the effective compression ratio as a kind of amplitude. In analogy to spectral filter theory, we define a cut-off modulation frequency as the half-value (-3 dB) point. Figures 4 and 5 show that the effective modulation bandwidth of this compressor is about 40 Hz.

The cut-off modulation frequency may be interpreted in terms of its effect on speech features. Comparison with Figure 1 suggests that the whole speech signal is effectively compressed with this compressor. In terms of time constants of signals, this value can be interpreted as reducing changes between signal parts occurring within 25 msec; such a speed is enough to make the compressor adjust its amplification within the silent gap before a plosive. We shall use the term phonemic compressor for such a fast compressing system.

We conclude that the theoretical limitation of the method should be taken into consideration. We further see that this type of analysis seems to make it possible to compare different designs of compressors with respect to their effectiveness and translate this measure to speech features. Compression systems can be described by two parameters, the effective compression ratio for low-modulation frequencies and the cut-off modulation frequency that describes the upper limit to which modulations in speech are suppressed.

Effect of Compression Parameters on the Effective Compression Ratio

We determined the effects of some of the parameters on the effective compression ratio. Our compressor allows for a rather easy change of the parameters without the effect being confounded by unpredictable interactions. The testing was done using our compressor.

Release Time of Compressor

The attack and the release times of the compressor have a pronounced effect on the effectiveness of the compressor. In hearing aids, the attack time is usually short, but the release time in commercial hearing aids shows large differences from values of about 70 msec up to 150 msec, typically around 125 msec.

We measured the effective compression ration as a function of the attack and the release time for a 2 kHz carrier frequency in the middle of the frequency band in which the compressor is effective. The level was set to the RMS level of speech (-22 dB versus maximum level). The modulation depth was 40%. The effect of varying the release time is shown in Figure 6 for an attack time of 5 msec. A similar result was obtained from varying the attack time.

We see the expected reduction of effectiveness of the compressing system. The cut-off modulation frequency changes from about 40 Hz to 20, 10, and 5 Hz. In terms of speech, the reduction of the cut-off modulation frequency means that for a release time of about 100 msec, as is often used in hearing aids, only differences between words, not between syllables, are compressed. It can be interpreted that only the level variations present in the intonation pattern are effectively filtered out, but that the amplitude distribution would hardly be affected.

Effect of Smoothing the Amplitude Envelope by Delay

Verschuure et al. (1993) and Verschuure and Dreschler (1993) argue that overshoots can be expected to reduce the effectiveness of the compressing system quite drastically. If a change in level occurs (e.g., at vowel onset or offset, at onset of plosive, etc.), the system will transmit the jump in level without any reduction, even if the system is fast. To make the system more effective, we introduced smoothing of the amplitude envelope by time delay and peak hold. The effect of delay on the effective compression ratio was tested next.

The conditions were a 2 kHz carrier modulated at 45% and set at a level 22 dB below the maximum level (the RMS level of the speech signal) with an attack time of 5 msec and a release time of 15 msec.

Figure 7 shows the determined effective compression ratio for various delays. It is striking to see that a delay of just 3 msec increases the effectiveness of the compressor by a factor of about 2. The cut-off frequency moves up from about 18 Hz to 35 Hz. A delay that is too long is counterproductive and even leads to an expansion of faster modulations (above 30 Hz).

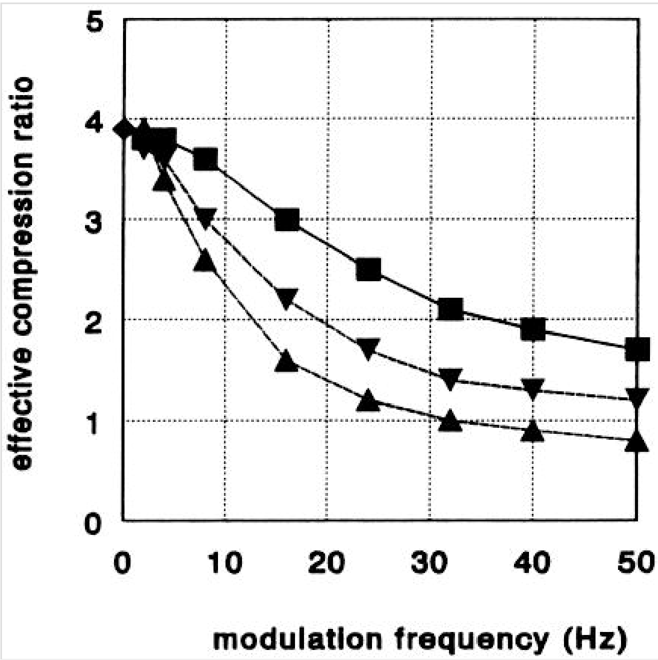

Figure 7. Effect of delay time (smoothing) on the effectiveness of the compressor. The delay times are optimal smoothing by a 3 msec delay (squares), no smoothing by a 0 msec delay (triangle down), and a misfit delay of 10 msec (triangle up).

Frequency Dependence

Thus far we have analyzed the effectiveness of the compressor at a certain signal frequency. The method also allows the determination of the frequency dependence of the compressor. Two implementations of the smoothed phonemic compressor are shown, just to demonstrate the kind of effects small changes in parameters may have on the effectiveness of the compressor.

Figure 8 shows the effective compression ratio for different carrier frequencies for a setting of the compression threshold at-70 dB versus maximum input signal and with the anti-USOM filter switched off. The static compression ratio was 4. The filter was chosen as the widest and flattest band filter possible which was a high-pass filtered at 0.5 kHz. The modulation frequency used was 4 Hz.

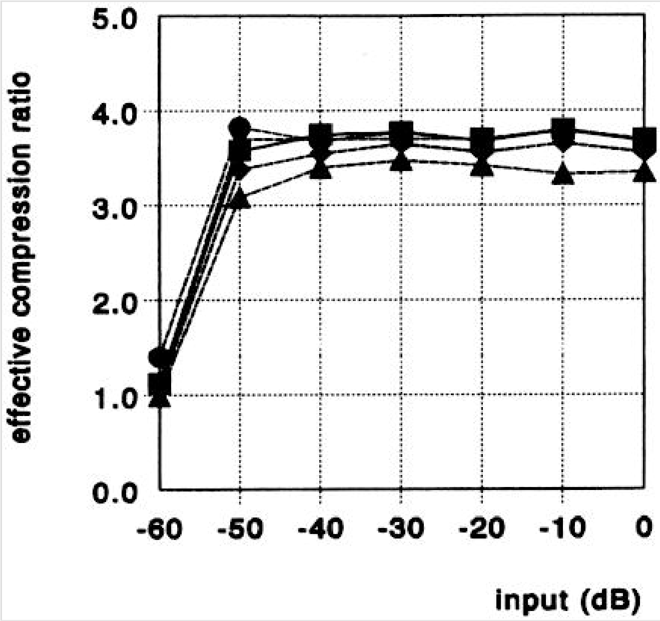

Figure 8. Effective compression ratio for various frequencies for a static compression ratio of 4. The compression threshold is set at -70 dB. The carrier frequencies are 8 (square), 4 (triangle down), 2(dot), 1 (diamond), and 0.5 kHz (triangle up).

We see that the dynamic range is limited to about 50 dB. The effective compression factor is just below 4, and the effectiveness goes down near the low cut-off frequency of the filter.

Next we determined the effectiveness of an older implementation of the compressing system. The filter was chosen in a very similar way. The lowest compression threshold in this system was -50 dB and this setting was chosen. Also, the anti-USOM filter was switched on. With this system, we analyzed the effective compression ratio of the isolated compressor of this implementation and of the total system. We found the isolated compressor to be as effective as the one used in Figure 8, except for a smaller dynamic range. The total system, however, was less effective, as can be seen in Figure 9. The dynamic range became rather limited, the compressor was less effective at higher levels because the signal in the linear path took over at high levels (at most frequencies) and because of the effect of the anti-USOM filter (at 0.5 kHz). The anti-USOM filter started to work as a compressor at high levels because it reduced the amplification in the low frequencies.

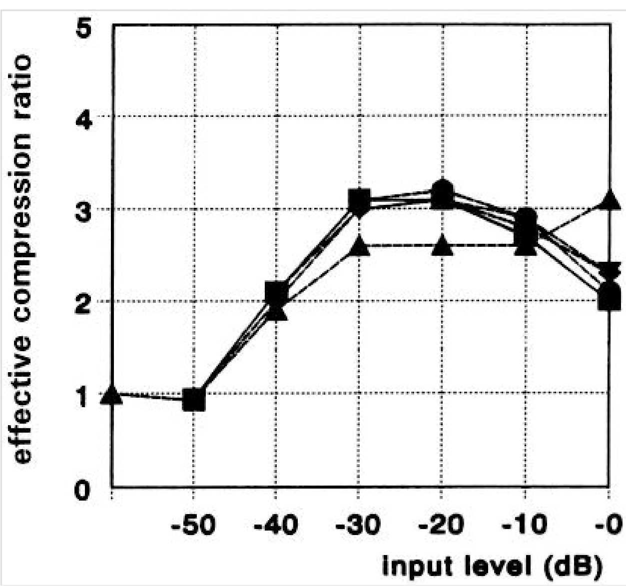

Figure 9. Effective compression ratio for various frequencies as in Figure 8, except that the compression threshold was set to -50 dB, the compressor was somewhat slower, and the anti-USOM filter was switched on.

We conclude that:

1. Our compression system is highly effective irrespective of the carrier frequency. It allows compression of rather fast changes in the speech signal over a wide dynamic range.

2. Changes in the parameters of the compression system have large effects on the effectiveness of the compression system with possible consequences for the compression of speech. It is necessary to analyze the effectiveness of compressors because the parameters interact in a rather complicated way.

Comparison of Commercial Hearing Aids and Experimental Compressor

It seems useful at this point to apply the same measurements to commercial hearing aids. The analysis was done for a rather common type of AGC as found in many standard hearing aids and for a K-amp compression system.

AGC-Hearing Aid

We used a Philips S45-I hearing aid to represent the conventional AGC hearing aids. The parameters of the Philips hearing aid were a compression ratio of 5 for input levels above 65 dB SPL, an attack time of 5 msec, and a release time of 110 msec. The static responses were checked and found to be correct.

The effective compression ratio of the S45-I is shown in Figure 10. The figure shows that the compression system has a cut-off modulation frequency of less than 4 Hz. This indicates that the hearing aid only suppresses the level information on the intonation pattern and slower changes in level. The system is not effective at evening out differences in level between words.

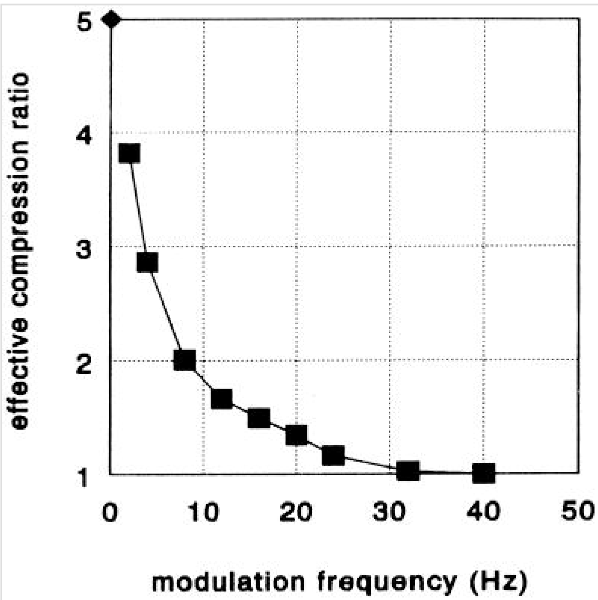

Figure 10. Effective compression ratio of a typical commercial hearing aid (Philips S45-I). Diamond indicates static value.

K-Amp Hearing Aid

Next we tested a K-amp amplifier. The K-amp has a level-dependent frequency response that is similar to that of our compression system. The compression system is an adaptive compression circuit. The attack time is 5 msec, the release time depends on the amount of time a signal has been on. Without the modulation analysis approach used here, it is difficult to determine how highly modulated signals such as speech are affected by such complex processing systems.

The time constants are usually not included in the specifications of hearing aids. Killion (1993) stated that the K-amp's release time for short signals such as slamming doors is around 20 msec; the release time rises sharply for signals longer in duration than about 100 msec, such as speech, to a release time of about 600 msec for signals longer than about 1 sec. In view of our analysis method, it was expected that the system would nevertheless show a short time constant in the range of 20 msec. The compression ratio is just above 2. The measuring conditions were a modulation depth of 40% and a carrier frequency of 2 kHz.

The effective compression ratio of the K-amp is shown in Figure 11 (as squares) together with the effectiveness of our smoothed compressor (diamonds) for compression ratio 2.Figure 11 shows that the cut-off modulation frequency of the K-amp is about 12 Hz, which is much higher than that found with conventional compressors. On the other hand, given the similarity in time constants of the K-amp and our smoothed compressor, the K-amp is less effective. If we take the effect of smoothing into account (Fig. 7), the difference can be understood to be the result of a small difference in release time and of the overshoots.

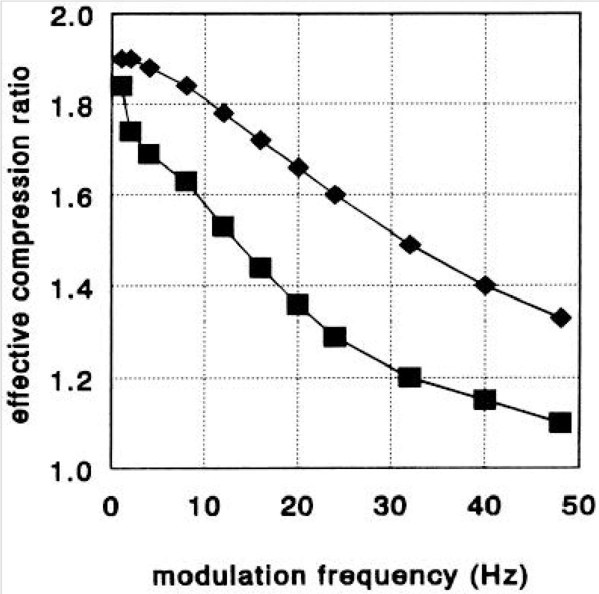

Figure 11. Effective compression ratio of a K-amp(squares) as compared to that of the smoothed phonemic compressor(diamonds).

Interpreting the data in terms of speech modulations, the K-amp seemed to be effective in evening out differences between words and syllables but did not seem to be fast enough to enhance consonant recognition such as the detection of a gap before a plosive. The short time constant may further cause overshoots, leading to false articulation cues. The overshoots interfere with the speech signal, and speech intelligibility should be determined when using such a system.

Conclusion

We conclude that there is a great difference between conventional AGC hearing aids and K-amp hearing aids. Translated into speech features, the former type of aids shows no effective compression for frequencies associated with words, syllables, and phonemes. The latter type of hearing aids shows effective compression for frequencies of up to about 13 Hz, including frequencies at which words and syllables are delivered. The K-amp could thus interfere with speech features.

Amplitude Distribution of Speech Signal

All of our remarks about the effectiveness of compressors were based on the comparison of the effective amplitude cut-off frequency for compression and the prevalent modulations in speech (Fig. 1). It is necessary to measure the effect of compressors on real speech.

Uncompressed Speech

For this analysis, we used a 32 sec sample of recorded continuous speech without pauses or hesitations spoken by one Dutch actor. The recording was made on DAT-tape by a radio engineer of the Dutch Broadcasting Enterprises without using any form of compression. The actor has a general accent and good articulation and intonation.

Wide-Band Control

The RMS level was adjusted to the -22 dB versus maximum level of our compressor, and this level was about 68 dB in the following wide-band plots. To diminish the background noise, the compression threshold was set at -45 dB, coinciding with 45 dB in the wide-band plots. The speech signal was rectified, and the amplitude was sampled with a cut-off frequency of 50 Hz at a rate of 1024 samples per second. Linear amplitudes were transformed in dB.Figure 12 shows the results of analyzing the output from the smoothed phonemic compressor as shown in Figure 2 using the flattest and widest filterband possible. The four line types are for compression ratio 1 (linear; solid line), 2 (dashed), 4 (dash-dotted), and 8 (dotted). To compare the widths of the patterns, the number of counts was scaled for the peak values to coincide; no shifts in the level direction were performed.

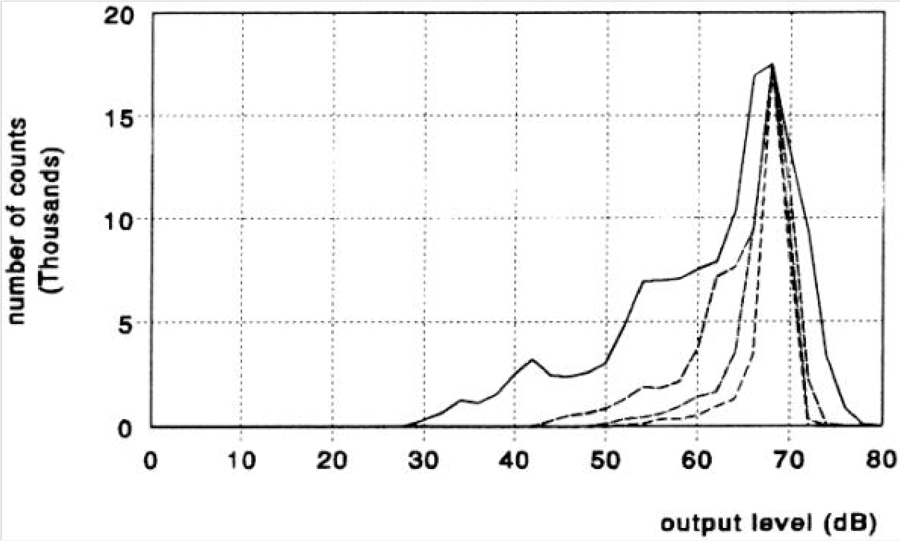

Figure 12. Level distribution of continuous speech for linear amplification (drawn line) and for compression ratio 2(dash), 4 (dash-dot), and 8 (dots) with the smoothed phonemic compressor of Figure 2.

The width of the uncompressed speech was about 45 dB, not uncommon for normally intonated speech. The compressor reduced this width substantially. If we analyze the width of the distribution pattern at a count number 20 dB below the peak value (0.1 × maximum number), we get for the compression ratios 2, 4, and 8-a reduction of the width by a factor of 1.9, 3.8, and 5.2. The compressor is indeed reducing the distribution of speech amplitudes.

Manufacturers of hearing aids tend to claim in their brochures that compression systems reduce the distribution of speech levels in all octave bands resulting in a much narrower "speech banana," which should then be fitted into the narrow dynamic hearing range. A simple analysis of what a single-channel, wide-band compressor does contradicts this idea. Consider a vowel (high-level sound with spectrum falling off at 10 dB/oct over 0.5 kHz) followed by a high-frequency consonant (spectrum around 2 kHz, level -20 dB versus peak spectral level of the vowel). The vowel-consonant combination gives only a small modulation in the 2 kHz band for the uncompressed signal. With wide-band compression the level of the consonant is raised, actually causing more modulation in the high-frequency band with compression than without compression. The example shows that unpredictable effects may occur.

We also analyzed the level distribution of the same segment of speech in octave bands. The compressor control was the same as wide-band control. The only difference compared with Figure 12 is that now only the distribution of signal levels in the 2 kHz octave band are analyzed. The results are given in Figure 13.

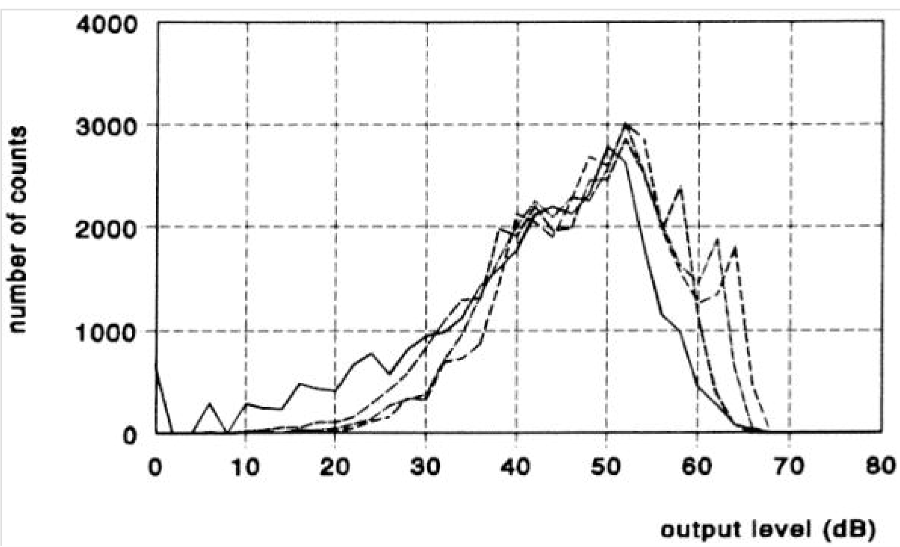

Figure 13. Distribution of speech levels in a 2 kHz octave band for different compression ratios: linear (drawn line), compression ratio 2 (dash), 4 (dash-dot), and 8 (dot) as in Figure 12.

We see that compression results in a little more gain for the high frequencies for higher compression ratios and a very small reduction in the amplitude distribution. The distribution seems to shift upward, in particular, the low-level sounds below 30 dB. The effective reduction for all compression ratios varies between 1.3 and 1.4, showing no effective reduction of the level distribution, or in other words, no "squashing of the banana."

Figure 14 shows a similar analysis of the amplitude distribution of the same speech sample in the octave band around 0.5 kHz with the same legend as the Figures 12 and 13. We see no clear effect of gain. The reduction factors for the compression ratios 2, 4, and 8 are 1.5, 1.9, and 2.0, respectively. This shows that there is some reduction in this frequency band but far less than is given by the compression ratio.

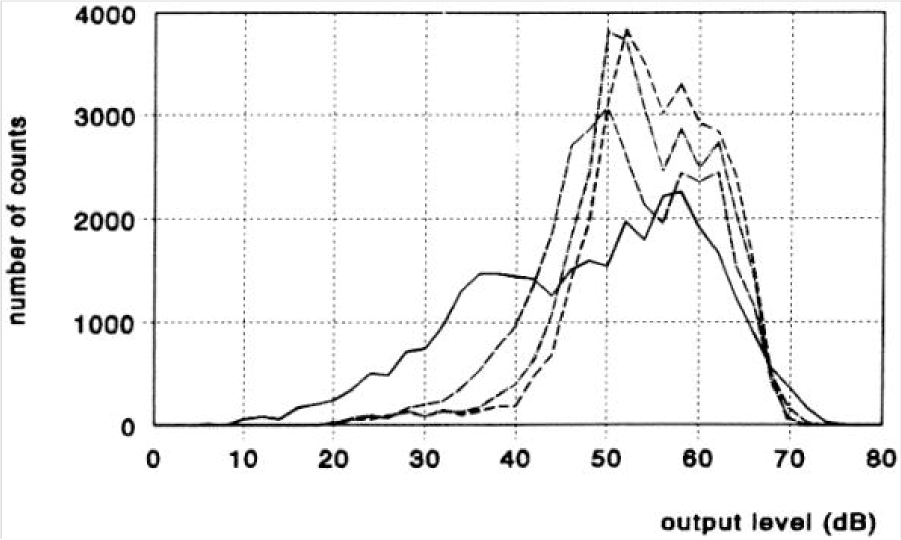

Figure 14. Distribution of speech levels in a 0.5 kHz octave band for linear amplification (drawn line) and for compression ratios 2 (dash), 4 (dash-dot), and 8 (dot).

We conclude that the single-channel, wide-band compressor is effective in reducing level differences. The compression does not lead, however, to a similar reduction of the width of the speech level distribution in the octave bands (no squashing of the speech banana). Theories about fitting the speech banana into the dynamic range are not justified for single-channel compressors. The usual compression systems are slower than our system, and in that case even a reduction of the distribution width in the wide-band analysis could be absent.

High-Frequency Control

We also analyzed the speech distribution when the control signal was taken from the output of the FIR-filter in Figure 2, the filter having a low-frequency cut-off frequency of 1850 Hz. In the linear filter path, a filter was added giving a high-frequency emphasis of 6 dB/oct. In this case the higher-frequency components of speech controlled the compression system and not the lower-frequency components. We expected the high-frequency band to be more effectively compressed.

The wide-band analysis (Fig. 15) now shows a somewhat less effective reduction of the amplitude distribution. The distribution shows an effective reduction by a factor of 2.1, 3.3, and 4.4, respectively, for compression ratios 2, 4, and 8. This shows that the system is also effective using this type of control for smaller compression ratios but not for higher compression ratios.

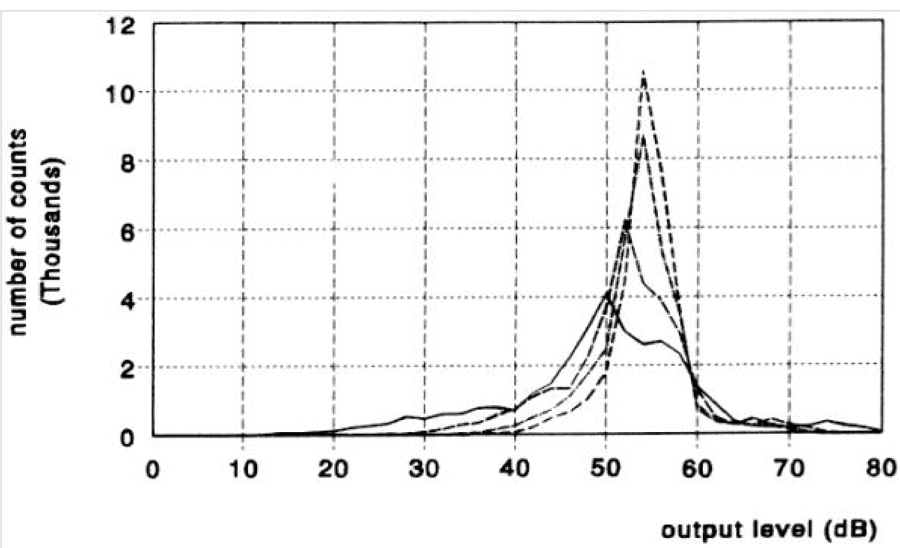

Figure 15. The distribution of unfiltered speech levels with high-frequency control of the compressor and high-frequency emphasis of the speech. Legend as in Figure 13.

In a way similar to that mentioned above, we studied the level distributions of speech in octave bands. The 0.5 kHz octave results showed that there was no change in the amplitude distribution at all. This result was expected because the signal was bypassing the compressor through the linear path.

Figure 16 shows the 2 kHz octave filtered amplitude distribution of the same speech signal as was used for the analysis in Figure 15. We see a reduction of the distribution. The effective reduction of the distribution width at the -20 dB level is now 2.0, 2.5, and 3.0, respectively. Also, the reduction in width of the amplitude distribution is now less than was found in the wide-band analysis but more effective than when a wide-band control signal was used.

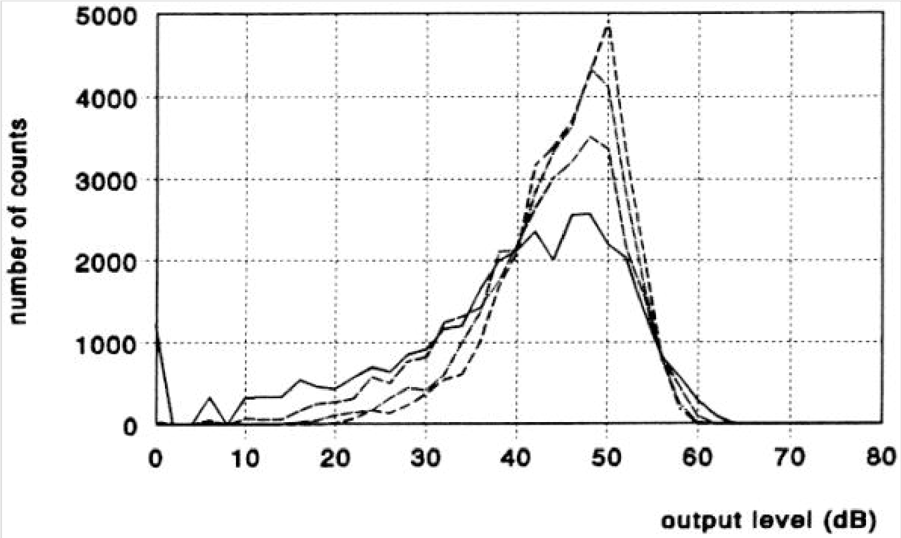

Figure 16. The distribution of 2 kHz octave filtered speech for linear amplification (drawn line), compression ratio 2(dash), 4 (dash-dot), and 8 (dot). The compressor is high-frequency controlled.

We conclude that the single-channel, wide-band compressor really does what it is expected to do for lower compression ratios. A reduction of the overall level, however, does not imply a reduction of the amplitude distribution in all frequency bands nor a reduction in one frequency band. If we want to achieve a squashed speech banana, the signal has to be split up into a number of channels with independent compression control by the speech signal in that particular band resulting in spectral distortion.

Commercial Hearing Aids

In a similar way, we analyzed the distribution of speech levels of the same speech sample passing through the S45-I hearing aid and a hearing aid with a K-amp circuit. The distributions were interpreted as above in an effective reduction of the level distribution. In the wide-band analysis, the reduction of the level distribution with the S45-I was 1.5 (static compression ratio is 5) and with the K-amp was 1.9 (static compression ratio is supposed to be just above 2). In the octave band analysis, we found in the 0.5 kHz band a reduction of the distribution width by a factor of 1.3 for the S45-I and no reduction of the distribution for the K-amp; in the 2 kHz band, the effective reduction of the distribution width was a reduction by a factor of 1.3 for the S45-I and 1.1 for the K-amp.

The analysis showed that the S45-I was not reducing the distribution width very effectively, although the static compression ratio was 5. The K-amp actually behaved as a single-channel syllabic compressor with wide-band control because it effectively reduced the wide-band level distribution but not the distribution in the high-frequency band.

Discussion

We have seen that for compressors to be effective for speech, the time constants should be very short. We argued that if we want to improve the detection of certain speech features of consonants, amplitude modulations of up to about 30 Hz should be compressed. Our speech intelligibility data show that for low compression ratios (less than 3), the speech score does improve for such fast (phonemic) compressors in comparison to linear amplification with the same frequency response (Verschuure et al., 1993; Verschuure, Prinzen, & Dreschler, 1994). Analysis of which speech features contributed to the improved score showed that consonant recognition indeed improved because of an improved plosive-fricative distinction (Verschuure, Reference Note 1). If slower compression systems are used, no improvements in the maximum score are reported, but the speech score still remains higher over a wider range of presentation levels than it does for linear amplification. If scores are averaged over levels, compression can be shown to improve average intelligibility even while maximum intelligibility may be somewhat reduced (de Gennaro, Krieg, Braida, & Durlach, 1981; Verschuure & Dreschler, 1993).

The theoretical calculation of the Appendix shows that a logarithmic compressor distorts the temporal pattern by introducing side bands into the modulation spectrum. This made us decide to limit the modulation depth in the analysis. However, speech of a single talker has a modulation depth between 30 and 45 dB. The temporal shape of these modulations is distorted by a logarithmic compressor. Such distortions reduce the effectiveness of the compressor (Fig. 4) and may interfere with consonant recognition and thus reduce speech intelligibility. The method itself could be used to quantify the distortion, but we did not study this speculation.

There are many reports in the literature that multi-channel compressors or even single-channel compressors do not improve, and may reduce, speech intelligibility (for a discussion of these matters, with references, see Verschuure et al., 1993). Bustamante and Braida (1987) and de Gennaro, Braida, and Durlach(1986) tried to fit the speech banana within the dynamic hearing range by multi-channel compression, and they found rather negative effects of multi-channel compression on the speech score. In these cases, the speech banana was effectively reduced, but spectral and temporal distortion could well reduce speech intelligibility. De Gennaro, Krieg, Braida, and Durlach (1981) suggested that the ineffectiveness of compressors could contribute to this finding. Our analysis, however, suggests that small compression ratios may be favorable and thus cannot be blamed for the poor results. Distortions of both types, on the other hand, may result in false articulation cues. Maré, Dreschler, and Verschuure(1992) showed that a multi-channel compressor with overshoot reduction and intelligent interaction between the compression in the different bands is not reducing the speech intelligibility in patients.

In typical commercial hearing aids, those with adaptive compression excluded, we could expect from the spectral analysis method that the amplitude distribution of words, syllables, and phonemes is hardly affected by the compression system. The measured level distributions of speech, either wide-band or octave filtered, were reduced by a factor of about 1.4. Both methods show that the compressor system is hardly effective. This type of hearing aid will most probably suppress the level information on the intonation pattern, and one could well ask whether that is desirable. The claim of reducing the amplitude distribution of speech (the speech banana) is unwarranted. Given these observations, a longer time constant could be preferable, giving the patient the possibility of profiting from the full perception of the intonation pattern. The advantage of such a system would still be that it adapts the speech level to different acoustical conditions without actually changing the speech signal itself.

Adaptive compression reduces the distribution of the amplitude levels of speech far more effectively. However, because of the overshoots, the effectiveness is reduced as is also shown by de Gennaro, Krieg, Braida, and Durlach (1981) for a multiband compressor with a compression ratio of 3. Additionally, we think that the overshoots may result in false articulation cues. Such a system, even if it is a single-channel compressor, changes the temporal characteristics of speech and may interfere with speech intelligibility. If such systems are used, speech intelligibility should be tested. In those cases, wearing comfort and speech intelligibility could become opposing effects, and one should ask from which setting a particular patient would profit most, the one providing better speech intelligibility or the one providing more comfort.

We see that the low effectiveness of the S45-I and the higher effectiveness of the K-amp are reflected in both measurements, the spectral analysis method, and in the distributions of speech levels. This suggests that the spectral analysis method can be used with some predictive value.

In view of the data on speech intelligibility and the effectiveness of compressors, we would propose to use slow-acting compressors (release times above 1 sec) to adapt to different acoustical conditions without the compressor having any effect on the speech signal. Single-channel compression systems have to be considered irrelevant for the reduction of the amplitude distribution of speech per octave band and thus should not be used for that purpose. The only possible positive effect is a better detectability of phonetic features. In that case we have to make sure that the detectability is not confounded by temporal distortions, and thus fast compression systems should include overshoot reduction.

We used special speech material for the analysis of the speech amplitude distribution. We wanted to use normal continuous discourse. Our first idea was to use a radio interview for that purpose. However, a speech signal recorded from the radio showed unwanted compression effects because the signal passes through many stages of compression: at the recording site, at the input to the transmitter, and in the receiver. The material, therefore, could not be used for our purposes. On the other hand, we did not want to make a recording ourselves because such recordings often do not have the characteristics of natural speech, as they are read. We obtained a recording on DAT-tape from a radio engineer working at the Dutch Broadcasting Enterprises. He assured us that the recordings had been made without any compressors being used. The amplitude distribution showed a width of about 45 dB, and this width was wider than the width of speech often reported in the literature. The difference was due to the amount of intonation (Verschuure & Dreschler, 1993). We also used another sample with more hesitations in it and found that the distribution of the levels was very similar except that the number of counts of intervals with a very low level (below 20 dB in the Figs. 12-16) increased dramatically. The width and shape of the level distributions above 25 dB were very similar.

Conclusions

1. A modulation technique can be used to assess the effectiveness of compressing systems at a certain level.
2. The modulation depth should not exceed 45% for an accurate analysis. The effect over a wider range of levels can be determined by repeating the measurement at different levels.
3. Smoothed compression increases the effectiveness of the compressor by a factor of about 2
4. Nonadaptive compression systems in commercial hearing aids usually are too slow to affect the speech signal except for the intonation pattern; adaptive compression systems compress differences between words and syllables, not within phonemes.
5. A phonemic compressor can effectively reduce level differences between sounds with the exact nature of the reduced amplitude width determined by the compressor control signal.
6. An effective compressor reducing overall level differences does not reduce the width of the amplitude distribution of speech in octave bands(squash the banana); there is no support for a claim of fitting the dynamic speech range to the dynamic hearing range.
7. Hearing aids with adaptive compression (adaptive release time) change the temporal characteristics of speech and may change speech intelligibility.

Acknowledgments:

Many people have contributed to this study. It is the result of years of discussions among numerous people on the issue of effectively describing compression. We want to thank explicitly B. Geerdink, M.Sc and P. Termeer, B.Sc from Philips Hearing Instruments.

This research has been supported by the Innovatief Onderzoeks Programma Hulpmiddelen Gehandicapten (IOP-HG) and Stichting Innovatief Programma Technologie (StIPT), both programs financed by the Dutch Ministry of Economical Affairs, the Heinsius Houbolt Trust Fund, and the Technological Innovative programme for the Disabled and Elderly (TIDE) of the European Union in cooperation with the hearing aid departments of Philips in Eindhoven, the Netherlands and Siemens in Erlangen, Germany.

References

Bustamante D. K., & Braida L. D. (1987). Multiband compression limiting for hearing-impaired listeners. *Journal of Rehabilitation Research and Development, 24*, 149-160. [Context Link]

Gennaro S. V. de, Braida L. D., & Durlach N. I. (1986). Multichannel syllabic compression for severely impaired listeners. *Journal of Rehabilitation Research and Development, 23*, 17-24. [Context Link]

Gennaro S. V. de, Krieg K. R., Braida L. D., & Durlach N. I. (1981). Third-octave analysis of multichannel amplitude compressed speech. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 125-128. [Context Link]

Killion M. (1993). The K-Amp hearing aid: An attempt to present high fidelity for persons with impaired hearing. *American Journal of Audiology, 2*, 52-74. Full Text @ EUR | Bibliographic Links | [Context Link]

Maré M. J., Dreschler W. A., & Verschuure H.(1992). The effects of input-output configuration in syllabic compression on speech perception. *Journal of Speech and Hearing Research, 35*, 675-685. [Context Link]

Pavlovic C. V. (1992). Statistical distribution of speech for various languages. *Journal of the Acoustic Society of America, 88*(Suppl. 1), 8SP10. [Context Link]

Plomp R. (1983). Perception of speech as a modulated signal. *Proceedings of the 10th International Congress of Phonetic Sciences, Utrecht*, 29-40. [Context Link]

Steeneken H. J. M., & Houtgast T. (1980). A physical method for measuring speech transmission quality. *Journal of the Acoustic Society of America, 67*, 318-326. [Context Link]

Stone A. M., & Moore B. C. J. (1992). Syllabic compression: Effective compression ratios for signals modulated at different rates. *British Journal of Audiology, 26*, 351-361. Full Text @ EUR | Bibliographic Links | [Context Link]

Verschuure J., & Dreschler W. A. (1993). Present and future technology in hearing aids. *Journal of Speech-Language Pathology and Audiology/Revue d'Orthophonie et d'Audiologie, JSPLA Monograph*(Suppl. 1), 65-73. [Context Link]

Verschuure J., Dreschler W. A., Haan E. H. de, Cappellen M. van, Hammerschlag R., Maré M. J., Maas A. J. J., & Hijmans A. C.(1993). Syllabic compression and speech intelligibility in hearing impaired listeners. *Scandanavian Audiology, 22*(Suppl. 38), 92-100. [Context Link]

Verschuure J., Maas A. J. J., Stikvoort E., & Dreschler W. A. (1992). Syllabic compression in hearing aids: Technical verification of nonlinear signal processing. *Acustica, 74/4*, s.14. [Context Link]

Verschuure J., Prinzen T. T., Dreschler W. A. (1994). The effects of syllabic compression and frequency shaping on the speech intelligibility in hearing impaired people. *Ear and Hearing, 15*, 13-21. [Context Link]

Reference Notes

1. Verschuure, J. Unpublished data. [Context Link]

APPENDIX

Theory of Spectral Modulation Analysis

The effectiveness of a compressor can be tested by using a sine wave with carrier frequency ([omega]$_c$) and a modulating frequency([omega]$_m$). The amplitude of the carrier is a_0, and the amount of modulation is given by the modulation index *m*. The amount of modulation can be determined accurately by analyzing the spectrum of the input or output signal.

An amplitude modulated sine wave can be described by:Equation (4) The spectrum of the signal of Equation 4 consists of three frequency components:Equation (5) We see that the difference in amplitude between the two symmetrical side bands and the carrier can be translated in dB. The equation for the difference in level in dB is:Equation (6) The comparison of the spectral level of the central component with that of one of the side bands provides a direct measure of the modulation depth.

$$x(t) = a_0(1 + m\cos(\omega_m t))\cos(\omega_c t)$$

Equation 4

$$\frac{1}{2}ma_0\cos((\omega_c - \omega_m)t)$$

$$a_0\cos(\omega_c t)$$

$$\frac{1}{2}ma_0\cos((\omega_c + \omega_m)t)$$

Equation 5

$$\Delta S = 20\log m - 6$$

Equation 6

Let us now consider how a sinusoidally-modulated sine wave (Eq. 4) is processed by a compressor. Compressors usually work in the logarithmic domain (dB versus dB) and not in the amplitude domain. We distinguish between linear amplitude measures denoted by small italics and logarithmic amplitude measures denoted by capital italics. The output signal (*U* in dB) depends on the input signal above the compression threshold (I_k) according to:Equation (7) where *I* is the input signal and*R* is the compression ratio. Transforming Equation 7 to the linear amplitude domain, gives:Equation (8) Equation 8 can be rewritten in linear amplitudes as: Equation (9)

$$\Delta S = 20\ \log m - 6$$

Equation 7

$$20\ \log \frac{u}{u_0} = R\left(20\ \log \frac{i}{i_0} - 20\ \log \frac{i_k}{i_0}\right) + 20\ \log \frac{i_k}{i_0}$$

Equation 8

$$\frac{u}{u_0} = \frac{i^R}{{i_0}^R}\left(\frac{i_k}{i_0}\right)^{1-R}$$

Equation 9

Substituting the time signal *x(t)* (Eq. 4) for *i* and assuming that the signal is well above the compression threshold reduces i_o and i_k to amplitude factors. We get: Equation (10) The first part of Equation 10 is a simple amplitude factor describing the action of the compressor. Substituting a new compressed amplitude,b_0, gives: Equation (11) The second part of Equation 10 or 11 shows that both the carrier sine wave and the modulation sine wave can be distorted by the compressor. This effect can be understood easily:

$$u = u_0\left(\frac{i_k}{i_0}\right)^{1-R}\left(\frac{a_0}{i_o}\right)^R (1 + m\ \cos\omega_m t)^R \cos^R \omega_c t$$

Equation 10

$$u = b_0(1 + m\cos\ \omega_m t)^R \cos^R \omega_c t$$

Equation 11

1. If we have a fast compressor and a low-frequency carrier, and for a sufficiently large compression ratio, the compressor acts instantaneously on the amplitude of the carrier and no sine wave will appear at the output. In our case, we assume that the carrier frequencies, which are above 100 Hz, are much higher than the upper limit of the effective compression frequency, which is below 50 Hz. For this reason, the effective compression ratio is 1 and can be omitted from that part of the equation.

2. If we use a high modulation depth and sufficiently fast compression, the modulation wave-form at the output will no longer be a sinusoid because of the nonlinearity of the logarithmic (dB) scale. As a consequence the higher harmonics of the modulation frequency appear, and the modulation waveform will be distorted.

The modulation factor can be analyzed using a Taylor expansion. The general form of the Taylor expansion is: Equation (12) This equation may only be used if *x* is small compared with 1. In our application, this implies the use of small modulation depths.

$$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!}x^2 + \frac{\alpha(\alpha - 1)(\alpha - 2)}{3!}x^3 + \ldots$$

Equation 12

Computing the first three terms of the Taylor approximation, replacing*x* with *m cos [omega]t* and [alpha] with *R* and using geometric equations for the second and third powers of the cosine, we get: Equation (13) This equation gives the amplitude factors of each modulation side band. We also see that there is an additional DC-term.

$$\frac{u}{b_0} = \left(1 + \frac{1}{4}R(R-1)m^2\right) + \\ + \left(Rm + \frac{1}{8}R(R-1)(R-2)m^3\right)\cos\omega_m t + \\ + \frac{1}{4}R(R-1)m^2\cos 2\omega_m t + \\ + \frac{1}{24}R(R-1)(R-2)m^3\cos 3\omega_m t + \dots$$

Equation 13

We now can make error estimates of the additional terms, specifically those adding to the DC-term and the first side band.

Because *R* has a value between 0 and 1, we know that:Equation (14) and Equation (15) If we do not want the errors to become larger than 5% in amplitude, we have to limit the modulation depth to less than 45% ($m < 0.45$). For this value of *m*, the harmonics of the modulation side bands are more than 40 dB down for the second harmonic and more than 50 dB down for the third harmonic.

$$R(R-1) \leq 0.25$$

Equation 14

$$R(R-1)(R-2) \leq 0.75$$

Equation 15

The calculation shows that the logarithmic compressor causes distortion in the modulation amplitude envelope. The spectral analysis method gives a correct representation of its effectiveness as long as the modulation depth of the input signal is rather small (<45%). This implies that the high modulation depths of speech cannot be imitated by the signal because of the high amount of distortion. It does not rule out the use of the method for higher modulation depths because, in general, we are interested in the effectiveness of a compressor near a certain level. The range of levels that make up speech can be analyzed by determining the effective compression ratios at a number of levels within the speech range. This also implies that speech compressed by a logarithmic compressor shows temporal distortion.

Using modulation depths of less than 45%, we can determine the effectiveness of the compressor at that particular level directly from the spectrum. Equation 13 simplifies to:Equation (16) In spectral terms this leads to side bands with a level of ½*Rm*. Writing R_{eff} explicitly using Equation 6, we get:Equation (17) Within the limitation of small modulation depths ($m < 0.45$), we can now analyze the effectiveness of the compressor system simply by looking at the input and output spectrum. The determining factor is the level differences between the carrier and the first upper or lower side band. This method allows for an easy determination of the effectiveness of a compressor at a certain level for a given carrier frequency and for a given modulation frequency.

$$\frac{u}{b_0} = (1 + Rm\cos\omega_m t)\cos\omega_c t$$

Equation 16

$$R_{eff} = 10^{\frac{\Delta S_0 - \Delta S_t}{20}}$$

Equation 17

The temporal distortion of compressors appears as higher harmonics, and these harmonics could be used as a measure of total temporal distortion. [Context Link]

IMAGE GALLERY

Select All

Export Selected to PowerPoint

$$x(t) = a_0(1 + m\cos(\omega_m t))\cos(\omega_c t)$$

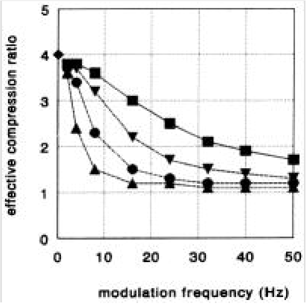

Figure 6

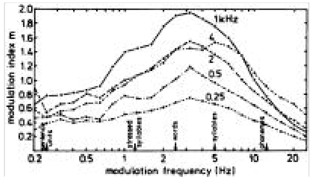

Figure 1

Equation 1

$$\Delta S = 20\,\log m - 6$$

Equation 2

$$R_{eff} = 10^{\frac{\Delta S_0 - \Delta S_i}{20}}$$

Equation 3

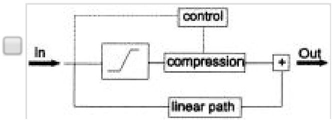

Figure 2

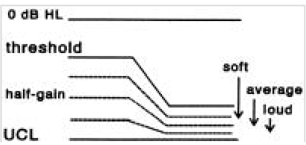

Figure 3

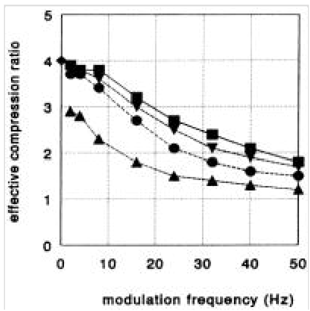

Figure 4

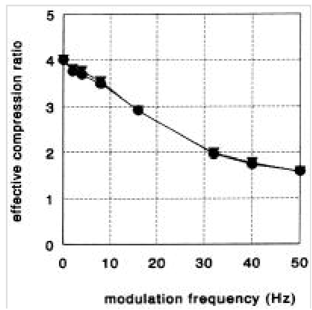

Figure 5

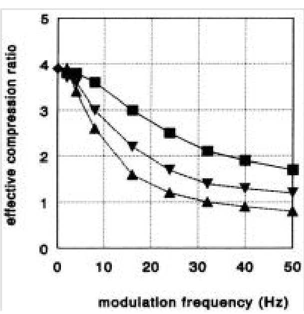

Figure 7

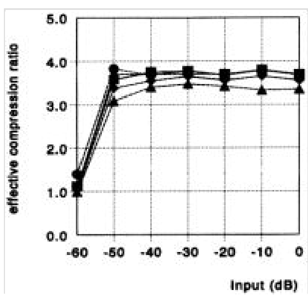

Figure 8

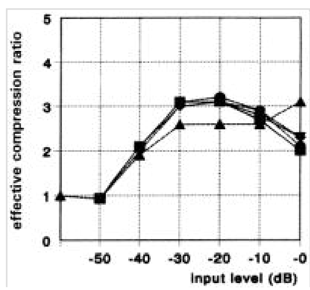

Figure 9

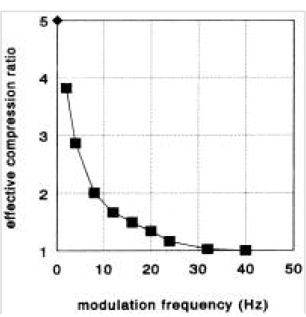

Figure 10

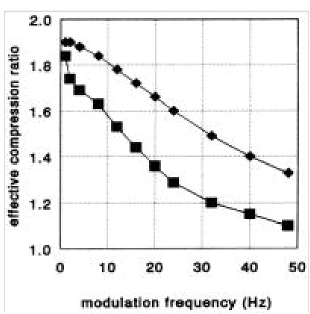

Figure 11

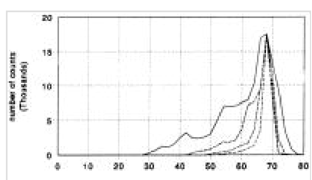

Figure 12

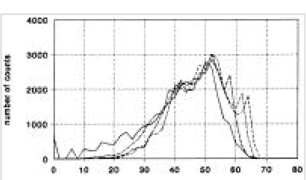

Figure 13

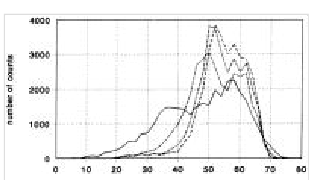

Figure 14

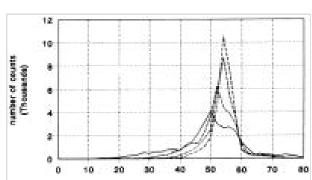

Figure 15

$$x(t) = a_0(1 + m\,\cos(\omega_m t))\,\cos(\omega_c t)$$

Equation 4

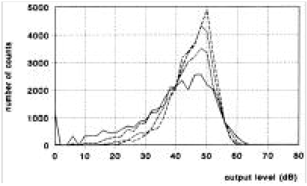

Figure 16

$$\frac{1}{2}ma_0\cos((\omega_c - \omega_m)t)$$
$$a_0\cos(\omega_c t)$$
$$\frac{1}{2}ma_0\cos((\omega_c + \omega_m)t)$$

Equation 5

$$\Delta S = 20\ \log m - 6$$

Equation 6

$$\Delta S = 20\ \log m - 6$$

Equation 7

$$20\log\frac{u}{u_0} = R\left(20\log\frac{i}{i_0} - 20\log\frac{i_k}{i_0}\right) + 20\log\frac{i_k}{i_0}$$

Equation 8

$$\frac{u}{u_0} = \frac{i^R}{{i_0}^R}\left(\frac{i_k}{i_0}\right)^{1-R}$$

Equation 9

$$u = u_0\left(\frac{i_k}{i_0}\right)^{1-R}\left(\frac{a_0}{i_0}\right)^R(1 + m\ \cos\omega_m t)^R\cos^R\omega_c t$$

Equation 10

$$u = b_0(1 + m\cos\ \omega_m t)^R\ \cos^R\omega_c t$$

Equation 11

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \ldots$$

Equation 12

$$\frac{u}{b_0} = \left(1 + \frac{1}{4}R(R-1)m^2\right) + \left(Rm + \frac{1}{8}R(R-1)(R-2)m^3\right)\cos\omega_m t + \frac{1}{4}R(R-1)m^2\cos2\omega_m t + \frac{1}{24}R(R-1)(R-2)m^3\cos3\omega_m t + \ldots$$

Equation 13

$$R(R-1) \leq 0.25$$

Equation 14

$$R(R-1)(R-2) \leq 0.75$$

Equation 15

$$\frac{u}{b_0} = (1 + Rm\ \cos\omega_m t)\cos\ \omega_c t$$

Equation 16

$$R_{eff} = 10^{\frac{\Delta S_0 - \Delta S_t}{20}}$$

Equation 17

Back to Top

© 2015 Ovid Technologies, Inc. All rights reserved. OvidSP_UI03.18.00.124, SourceID 69505

About Us | Contact Us | Terms of Use