J Am Acad Audiol 20:161–171 (2009)

Long-Term Signal-to-Noise Ratio at the Input and Output of Amplitude-Compression Systems

DOI: 10.3766/jaaa.20.3.2

Graham Naylor*
René Burmand Johannesson*

Abstract

We present measurements showing that the long-term signal-to-noise ratio (SNR) at the output of an amplification system that includes amplitude compression may be higher or lower than the long-term SNR at the input, dependent on interactions among the actual long-term input SNR, the modulation characteristics of the signal and noise being mixed, and the amplitude compression characteristics of the system under test.

The effects demonstrated with the measurements shown here have implications for choices of test methods when comparing alternative hearing aid systems. The results of speech-recognition tests intended to compare alternative systems may be misleading or misinterpreted if the above interactions are not considered.

Key Words: Amplification, compression, hearing aids, modulation, noise, speech

Abbreviations: CR = compression ratio; ICRA = International Collegium for Rehabilitative Audiology; RMS = root mean square; SNR = signal-to-noise ratio; WDRC = wide dynamic-range compression

Amplitude compression is a common feature in modern hearing aids. Various rationales for using compression exist (Dillon, 1996), and accordingly the choices of compression parameters to be applied (time constants, dynamic range, compression ratio, number of frequency channels, etc.) vary widely between systems in the field.

Meanwhile the research literature does not provide overwhelming evidence supporting the use of compression or clear guidelines regarding the optimal settings of compression parameters. Concerning speech intelligibility in noise, only a minority of studies show better performance with compression than with linear amplification (see, e.g., Souza, 2002, for a review). Studies comparing the benefits of alternative choices of compression parameters are widely divergent in their conclusions (see, e.g., Gatehouse et al, 2006a, for a review of studies on time constants). There are many possible reasons for this lack of consensus. Choices of signal and noise conditions, device configurations, listener groups, and domains of outcome measure are all sources of confounding effects, as the action of amplitude compression is dependent on complex interactions between signal properties and system configuration, and the benefits of compression are dependent on what question is being asked, and of whom (Gatehouse et al, 2003, 2006b).

The most challenging listening situations are those in which a desired signal is heard in the presence of noise. Given the complex properties of the typically desired signal (speech) and the variety of properties that noises may have, signal-plus-noise conditions are also those in which the most complex and varied effects of compression are likely to arise. Historically it has been possible to manipulate the characteristics of a signal-plus-noise mixture at the input to a compression system and observe both the sum signal at the output and listener performance (typically speech intelligibility).

Additional insights into the effects of a compression system could be gained by observing the signal and noise separately at the system output. Several investigators have recently begun to do this. Olsen et al (2005) used the separation technique of Hagerman and Olofsson (2004) to extract the signal and noise components from the mixture at the system output. As one part of their study, they examined the long-term signal-to-noise ratio (SNR) at the input and

*Oticon Research Centre ''Eriksholm,'' Denmark

Graham Naylor, Oticon Research Centre ''Eriksholm,'' Kongevejen 243, DK-3070 Snekkersten, Denmark; Phone: +45 4829-8915; Fax: +45 4922-3629; E-mail: gn@oticon.dk

Portions of this study were presented at the International Hearing Aid Research Conference, Lake Tahoe, August 2006.

output of a fast-acting compression-amplification system, when the input signals were speech against a background of "fully modulated noise." The compression system was unusual and artificial, in that the control signal for the compressor was always derived from the same signal-plus-noise mixture, even though the mixture being processed by the compressor varied in SNR. However, one of the measurement conditions (input SNR = −14.3 dB, control signal SNR = −15 dB [Olsen et al, 2005, Table 3]) corresponds closely to that of a normal compressor, where both of these SNRs would be equal. In this condition, the long-term SNR at the output was approximately 3.5 dB higher (less negative) than the long-term SNR at the input (see Olsen et al, 2005, Figure 4).

Souza et al (2006) used the same separation technique to quantify some effects of compression on speech in noise. They conclude that fast-acting wide dynamic-range compression (WDRC) degrades the SNR. This is in apparent contradiction to the findings of Olsen et al (2005). Souza et al (2006) also conclude that the speech envelope is less distorted by compression when there is noise present than when noise is absent.

Souza et al (2007) wished to determine the audibility of speech signals in a noise background, after processing through compression amplification, and again used the Hagerman and Olofsson (2004) separation technique to extract speech levels at the output. For the conditions tested, calculated speech audibility was found to be lower after compression amplification than after linear amplification, but the authors urge caution in this conclusion, due to the limited variety of experimental conditions.

Stone and Moore (2007) wished to investigate the potentially deleterious effects of fast-acting compression arising through its effects on the speech envelope. They used a technique whereby access to the "gain signal" arising within the compression system allows one to generate the compressed signal and noise components as they must have been at the output. One property of compression acting on a mixture of signals is that a degree of correlation arises between the envelopes of the mixture components at the output. Stone and Moore suggest that this may give rise to a form of perceptual fusion between the signal sources and thus degrade intelligibility.

Thus several approaches have been taken so far in studying how compression amplification affects the separate components of a signal-plus-noise mixture at the output and their mutual relationship. We chose to focus on long-term SNR for the following reasons:

First, there is some apparently contradictory data in the literature, which may be resolved by a systematic set of measurements covering a wide variety of conditions.

Second, the long-term SNR has a long history for quantifying perceptual conditions in linear systems, and it may also be useful in the case of nonlinear systems. Long-term SNR is a crude indicator of perceptual performance, as it cannot account for any of the temporal dynamics in either the signal or noise and, for example, is unable to predict that speech signal perception against modulated background noise (at least for normal-hearing listeners) is generally much better than against steady-state noise at the same long-term SNR (Wagener and Brand, 2005). However, amplitude compression has many diverse effects on signals. Some of these effects may be perceptually advantageous, while others may be perceptually deleterious. Indeed, the same phenomenon may produce advantageous effects under some signal or system conditions and deleterious effects under other conditions. For this reason also, it seems prudent to start by looking at a simple and often-used measure like long-term SNR.

In the first part of this study we compare the long-term SNR at the input and output of a hearing aid with amplitude compression by carrying out a series of measurements and using the Hagerman and Olofsson (2004) technique for separation of signal and noise at the hearing aid output. The "Signal" is always speech, and we systematically vary the amount of modulation in the "Noise" component, the parameters of the compression, and the long-term SNR at the input.

The measurements presented here offer a resolution of the apparent contradiction between the results of Olsen et al (2005) and Souza et al (2006). Explanations are provided for the observed effects, in terms of the dynamic properties of the signal and noise being mixed and the action of the amplitude-compression system.

Having proposed the basic pattern of objective effects of compression on long-term output SNR, in the second part of this article we briefly address the question of whether these effects are likely to have any perceptual relevance. Finally, in the light of the data presented here, we draw attention to the fact that some speech audiometry procedures often used for comparing the speech-in-noise performance of alternative amplification systems may suffer from serious confounding effects that could render them unreliable or misleading.

METHOD

Signal Demixing Technique

Measurement of long-term SNR at the input of a compression-amplification system ("Input SNR") is straightforward, as long as one has control over the activation of Signal (S) and Noise (N) separately. At the system output, the characteristics of S (or N) will differ, dependent on whether or not N (or S) is also present at the input. Therefore simply activating the sources of S and N alternately cannot be used when determining the SNR at the output ("Output SNR"). Instead, the separation method of Hagerman and Olofsson (2004) can be used. This off-line method requires two recordings to be made from the system output: in one of these, S or N (but not both) at the input must be given the opposite sign. Assuming that the system responds identically to S + N and S + (−N), and given a highly reproducible time alignment of signals and sampling, it is feasible to remove either S or N from the output mixture by adding or subtracting the two recorded output time series. The degree of removal can be verified objectively by combining S + N with (−S) + (−N), which ideally should give a perfect null. The unwanted residual signal was suppressed by more than 20 dB compared to the lowest of the S or N signal components for all conditions in the present experiments.

Apparatus

An electroacoustic setup was used, in which a real hearing aid in a test chamber was the system under test. Compared to using a software simulation, this approach provides challenges in calibration, programming of the system configuration, and signal acquisition but was used because we wished to be able to progress to increasingly complex and realistic hearing aid configurations (including a completely different hearing aid) without changing any of the signal conditions. A behind-the-ear hearing aid (Oticon Syncro II) was initially programmed using the manufacturer's fitting software, targeting a flat insertion gain frequency response (+1 dB/−5 dB from 300 Hz to 6 kHz for a 65 dB SPL pure-tone excitation) in the average adult ear, according to the principles of Bentler and Pavlovic (1989). The resulting program settings were then modified using special programming software to provide a constant compression ratio for pure-tone input levels in the range from 30 dB to 90 dB SPL. This ensured that the level distributions of all signals and noises used in the subsequent measurements were within the region of constant compression ratio and thus unaffected by any compression kneepoint or output limiting. The same software was used to configure compression parameters for all the measurements. The hearing aid was programmed to act as a single-channel device for most of the measurements. For the measurements reported here, all additional features (directional processing, noise reduction, etc.) were disabled, so we could observe the intended phenomena without confounds from others. Portions of the measurements were replicated with a different hearing aid (Bernafon Symbio 100), programmed likewise.

Signal generation and recording were done with MATLAB code on a Windows XP SP2 PC through a 24 bit Echo layla 3G sound card with a sample rate of 96 kHz, an ASIO 2.0 sound card driver, and a Tucker-Davis Technologies TDT SA1 power amplifier. A Brüel & Kjær anechoic test box type 4232 provided the test enclosure and source loudspeaker, with a Brüel & Kjær IEC711 ear simulator mounted on a type 4192 microphone for recording the hearing aid output and a Brüel & Kjær type 4192 microphone for recording the input signal mixture close to the hearing aid's microphone inlet. Both measurement microphones were connected to a Brüel & Kjær type 5935 dual microphone supply.

The measurement system simultaneously acquires Linear and A-weighted measurements and 1/3-octave spectra from 100 Hz to 10 kHz. During operation/acquisition it was most convenient to use the A-weighted level for calibration, due to its robustness against low-frequency fluctuations and disturbances. However the reported levels are Linear-weighted, calculated from summing the long-term 1/3-octave spectra from 100 Hz to 10 kHz.

Spectral Matching of S and N

Hearing aids typically apply a strongly frequency-dependent gain to the input signal. When comparing Input SNR and Output SNR summed across a given frequency band, it is important that the frequency dependence of gain within that frequency band does not combine with differences between the input spectra of S and N to bias the result in an uncontrolled fashion. An illustrative but extreme example would be the case with S being a speech signal (with typical roll-off toward high frequencies), N being a white noise signal, and a hearing aid with a typical high-frequency emphasis; SNR is to be measured in one broad frequency band. The value of Input SNR will be controlled by the low-frequency peak of S relative to the flat spectrum of N, whereas the Output SNR after the frequency shaping in the hearing aid will be controlled by the new high-frequency peak of N relative to the now-flattened spectrum of S. Any influence that the amplitude-compression system might have on Output SNR would thus be confounded with the effects of frequency shaping.

Pilot measurements indicated that inherent differences between the long-term spectra of our speech material and a commonly used "speech spectrum noise" could result in up to 4 dB of bias in the difference between Input SNR and Output SNR, if the hearing aid applied a typical high-frequency emphasis. To avoid this type of bias, the long-term 1/3-octave spectra of all the S and N signals at the hearing aid microphone (i.e., input) were matched within [−0.5 dB; +1.0 dB] at all frequencies between 100 Hz and 10 kHz. We verified that this removed any bias by making a set of measurements with the hearing aid programmed as a linear amplifier (for which Input SNR and Output SNR should always be equal).

Our requirement for close spectral matching of S and N distances this study from real-life hearing aid use, where S and N spectra are unmatched and often change over time. In such a situation it would be extremely difficult to disentangle SNR effects due to compression from the effects of differing spectra. When attempting to understand one set of phenomena, we have to minimize interference from other phenomena.

Signals Used

The same speech signal ("S1") was used in all measurements, consisting of sentence lists 1 and 3 from the published CD recordings of the Dantale II corpus (Wagener et al, 2003). Each list consists of ten five-word sentences. Pauses between sentences were removed, resulting in a running speech signal of approximately 46 sec duration. The root mean square (RMS) level of this signal was 60 dB SPL at the hearing aid microphone in all measurements.

Three noise (N) signals were used:

- Unmodulated noise "N1" was stationary noise with the same long-term spectrum as the Dantale II speech material (the R stereo channel from the Dantale II CD).
- Modulated noise "N2" was International Collegium for Rehabilitative Audiology (ICRA) two-speaker-modulated noise (track 6 from the ICRA noises CD [Dreschler et al, 2001]), filtered as previously described to have the same long-term spectrum as the Dantale II speech material.
- Speech noise "N3" was the same material as S1 but with the order of sentence lists reversed (3 followed by 1). Thus S and N3 are two speech signals with the same talker and linguistic characteristics but uncorrelated on a moment-by-moment level.

Each noise signal was available in two versions; the normal noise signal (N) and the inverted noise signal (−N), obtained simply by changing the sign of the normal noise signal. Measurements were carried out in pairs, with the input signal composed of S + N or S + (−N), to provide data for the demixing algorithm.

The noise signals N1, N2, and N3 were varied in RMS level between 50 and 70 dB SPL at the hearing aid microphone, thus providing Input SNRs in the range −10 to +10 dB. All signals had a minimum duration of 45 sec, whereas the measurements had a duration of 30 sec. A random time offset of up to 15 sec between S and N was applied for each measurement (the same offset in each recording in an S + N and S + [−N] pair).

Choice of SNR Measure

All SNRs quoted here are derived from ratios of simple long-term RMS levels in S and N, acquired over the entire duration of the measurement. Naturally occurring energy minima ("pauses") between words in a sentence were included, whereas the pauses of indeterminate duration between sentences were removed, as already mentioned. In many circumstances, all speech pauses are excluded when measuring the RMS level of speech signals, as the RMS level of the speech when present is otherwise underestimated. However, for the present purposes we decided to include natural pauses, primarily because it is during the pauses in the S and N signals that some of the most interesting effects occur. By including the complete signal, we make no assumptions about which parts are important, avoid treating S and N differently, and avoid having to define the boundaries of "pauses." The data values of Input SNR and Output SNR shown in the figures represent the ratios of the long-term RMS levels of the S and N signals (in the case of Output SNR, after separation).

Measurement Conditions

Input SNR was varied between −10 dB and +10 dB in steps of 2 dB, thus covering most of the relevant range of Input SNRs for real speech-in-noise listening conditions. Output SNR was measured at each value of Input SNR. This sequence was repeated for the combinations of S and N signals, compression parameters, and hearing aid models shown in Table 1, which also contains measurement ID numbers that are referred to in the legends of the figures showing the results.

In the case of the Oticon Syncro device, all possible combinations of all the compression parameters mentioned in Table 1 amount to 72 conditions (3 noise types × 3 compression ratios × 4 compression speeds × 2 configurations of channels = 72 conditions). Displaying all of these results would require an inordinate amount of space. Measurements were made covering nearly half of the conditions, and according to these data the interactions between the effects of different parameters appeared to be minimal. Thus in the following section we illustrate selected "cross sections" of the data, which show trends representative of the whole.

RESULTS AND DISCUSSION

The reproducibility of the measurements was assessed by repeating a random selection of the measurement conditions on a later occasion, after removal and replacement of the hearing aid in the

Table 1. Details of Measurement Conditions

Device	Signal	Noise	Compression Ratio	Compression Channels	Attack/Release Time (msec)[a]	Measurement ID	Figure Number
Syncro	S1	N1	2:1	1	5/26	1	1, 3
		N2				2	
		N3				3	
Syncro	S1	N2	1:1	1	5/26	4	4
			2:1			5	
			4:1			6	
Syncro	S1	N2	2:1	1	5/26	7	5
					10/100	8	
					22/400	9	
					50/1550	10	
Syncro	S1	N3	2:1	1	5/26	11	6
				8		12	
Syncro	S1	N3	2:1	1	0.6/5	13	7
Symbio					0.6/4	14	

[a]American National Standards Institute, 2003.

test chamber and recalibration. Test–retest differences in Output SNR were within 0.2 dB in all cases. The main contribution to the measurement variability comes from the random selection of the noise period. The N2 modulated noise causes the largest variability of all the signals used.

Effects of Relative Modulation in S and N

Figure 1 shows the results of a measurement in which S is speech (S1) and N is unmodulated noise (N1). The hearing aid is configured with a single channel of fast-acting compression and a compression ratio of 2:1. The Output SNR is lower than the Input SNR for all Input SNRs tested, the difference growing from approximately 0.5 dB at −10 dB Input SNR to almost 4 dB at +10 dB Input SNR. Given that the hearing aid at any given moment applies the same gain to the speech and noise in the mixture at the input, the altered long-term SNR at the output requires an explanation.

Figure 2 shows time series corresponding to the short-term input levels of S1 and N1, and the corresponding gain arising in the hearing aid, for a period of time roughly equal to the duration of one sentence. The measurement condition is with Input SNR = +10 dB. The long-term RMS level of S at the output is determined mainly by the level of the peaks of S. These peaks are always significantly above the level of N and thus determine the gain applied to those same peaks. The compression characteristic of the hearing aid ensures that these peak input levels receive lower-than-average gain. In the pauses of S, the N signal has greatest amplitude and determines the gain applied. As N is less intense than S, the gain will be higher in these periods than it is for the peaks of S. The high gain applied during speech pauses has almost no effect on the long-term output level of S, since S has almost no energy at these times. In sum, the S signal "penalizes itself" in terms of gain, relative to the N signal, and this reduces the Output SNR relative to the Input SNR.

This mechanism remains in force as Input SNR is lowered, but with decreasing effect as the peaks of S gradually become less prominent against the steady-state N. For very large negative Input SNR (well beyond −10 dB), the asymptote is equality of Input SNR and Output SNR, as the gain in such conditions will be entirely determined by the level of N at all

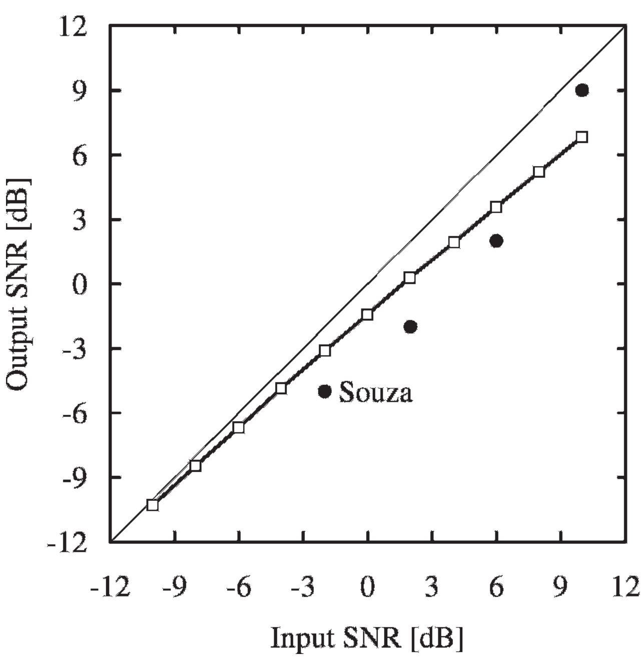

Figure 1. Results of measurement 1: signal-to-noise ratio (SNR) for single-channel fast-acting compression, compression ratio 2:1, speech S1 in unmodulated noise N1. Diagonal line indicates equality between Input SNR and Output SNR. Additional points ("Souza") show data for corresponding measurement from Souza et al, 2006, Table 1.

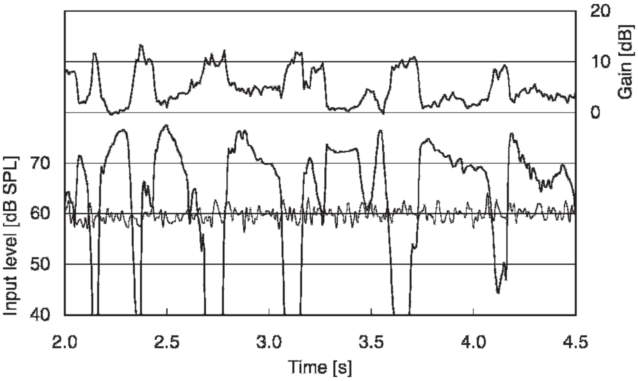

Figure 2. Segment of time course of signal levels in measurement 1 at Input SNR = +10 dB and corresponding gain values applied by the hearing aid. Level and gain values are derived from true root mean square for each 10 msec. Lower two curves: input signals speech S1 (thick line) and noise N1 (thin line). Upper curve: hearing aid gain.

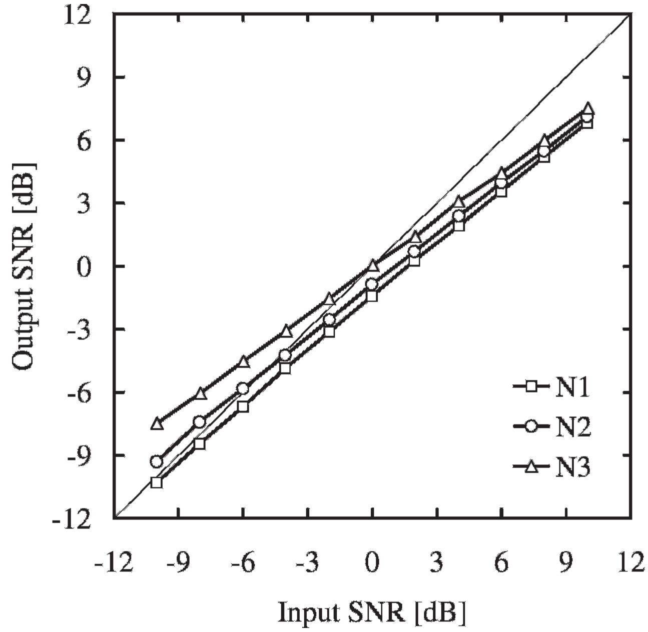

Figure 3. Results of measurements 1–3: signal-to-noise ratio (SNR) for single-channel fast-acting compression; compression ratio 2:1; speech S1 in unmodulated noise N1, modulated noise N2, and speech N3.

times and thus is essentially constant. As Input SNR increases toward large positive values, the difference between Input SNR and Output SNR will continue to increase until the momentary SNR at the input is always large enough for the gain to be determined entirely by S.

The above explanations implicitly assume extremely fast-acting compression. Bringing the finite time constants of the compressor's level detector into play, we may consider the effects commonly referred to as "overshoot." Since a level detector with finite time constants is sluggish in following level variations, the gain applied to the early part of an event with rapid onset or offset will deviate from the intended static input–output curve ("too much gain" for an onset and "too little gain" for an offset). Thus in the present measurement case, and with positive Input SNR, onset overshoot will allow strong onsets of S to pass with more gain than intended, and the "compression penalty" for peaks of S explained above will be reduced. Likewise, immediately after offsets in S, the N signal will briefly receive less gain than intended, and the "gain bonus" for N overall will be reduced. Overshoot effects tend to make the overall result more like linear gain, that is, pull the Output SNR curve toward equality with Input SNR. The results in Figure 1 indicate clearly that in the present case with fast-acting compression, the effects of overshoot are outweighed by the effects of the static compression characteristic.

Figure 1 includes data points from Souza et al for one of their measurement conditions that very closely resembles the present one (S = speech, N = spectrum-matched steady-state noise, single-channel compression with ratio 2:1, attack/release times = 5/50 msec [Souza et al, 2006, Table 1]). For Input SNRs of −2, +2, and +6 dB, Souza et al found Output SNRs between 1 and 2 dB lower than the present measurements indicate. Such deviations are thought to be within reasonable bounds of deviation arising from the inevitable differences in systems and signals tested (PE Souza and LM Jenstead, personal communication). At an Input SNR of +10 dB, Souza et al (2006) measured an Output SNR of +9 dB, almost 3 dB higher than the present measurements. In their study this off-trend result was tentatively identified as being due to the presence of a compression threshold in the system they measured. The present measurements, made on a system in which any effects of a compression threshold were carefully avoided (by ensuring a constant compression ratio for all input levels in the range 30 dB to 90 dB SPL), support this interpretation, insofar as they show a trend that is continuous and smooth over the whole range of Input SNR.

Figure 3 collects the results for measurements with noise types N1 (unmodulated), N2 (two-talker modulated), and N3 (speech) and the same hearing aid configuration as used in Figure 1. Consider first the data collected with the modulated noise N2. At large positive Input SNRs, the Output SNR is lower by an amount similar to that seen with N1 and for the same reasons as expounded above. However, as Input SNR decreases, the Output SNR approaches equality with the Input SNR more quickly than with N1. This is because the input level distribution of N2 is much wider than that of N1, so occasionally N will be dominant over S at times other than pauses in S, and now N will "penalize itself" in terms of gain applied, even at positive Input SNRs. For Input SNRs below −6 dB, the Output SNR with N2 is actually slightly higher than the Input SNR. The crossover point

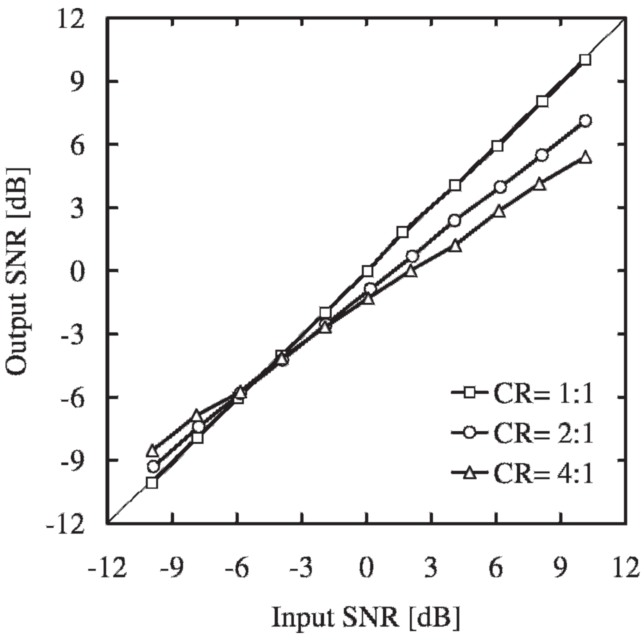

Figure 4. Results of measurements 4–6: signal-to-noise ratio (SNR) for single-channel fast-acting compression; compression ratios (CRs) = 1:1, 2:1, and 4:1; speech S1 in modulated noise N2.

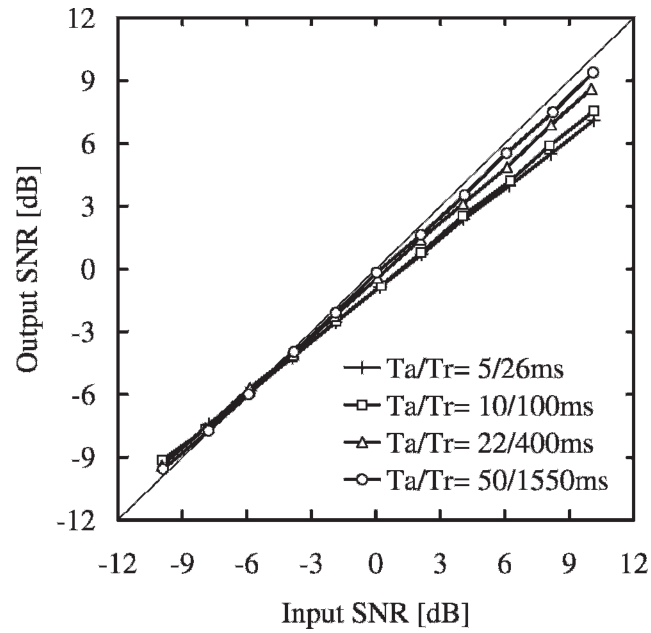

Figure 5. Results of measurements 7–10: signal-to-noise ratio (SNR) for single-channel compression; compression ratio 2:1; compression time constants (attack/release [Ta/Tr]) = 5/26 msec, 10/100 msec, 22/400 msec, and 50/1550 msec; speech S1 in modulated noise N2.

(−6 dB) represents the Input SNR "bias" required such that the long-term levels of both S1 and N2 (with their different-shaped level distributions) are affected equally when they are both present simultaneously.

Also shown in Figure 3 is the case where both S (S1) and N (N3) have the same level distribution statistics (S and N both being speech with the same sentence structure spoken by the same talker). This yields an antisymmetrical graph, with the crossover point at 0 dB Input SNR, as we should expect. The Output SNR is always closer to 0 than the Input SNR; at ±10 dB Input SNR, the effect seen here is about 2.5 dB.

In graphs of Input SNR versus Output SNR, the two principal effects of changing the modulation in N are thus to move the crossover point and vary the angle of the data line relative to the diagonal. The comprehensive measurements mentioned earlier indicate that a given N always yields the same crossover point (given that S is always the same), regardless of the compression parameters. Likewise, the angle of the data line varies for different Ns in a very similar way to that shown in Figure 3, regardless of the compression parameter being varied.

Effects of Compression Parameters

In this section we present the results of measurements in which several important parameters of the compression are varied. The variations used here are chosen to represent wide ranges that nevertheless may be found in currently available hearing aid systems.

Compression Ratio

Figure 4 shows the results of measurements in which only the compression ratio (CR) was varied. With CR = 1:1, Output SNR and Input SNR are equal for all values of Input SNR, as expected. As CR is increased the line of Output SNR versus Input SNR rotates away from the diagonal, such that Output SNR varies less and less as CR increases. The trend line crosses the diagonal at a point that remains more or less constant regardless of CR. With noise type N2 shown here, this point is approximately −6 dB.

Compression Time Constants

Figure 5 shows the results of measurements in which only the speed of compression was varied. Again trend lines are seen that rotate around a crossover point whose location is determined by the relative modulation in S and N. As compression becomes less fast-acting, the trend line moves toward the diagonal, indicating increasingly linear behavior as the variations in gain over time are reduced. Alternatively, in terms of output, overshoots become more and more significant. However, it is notable that even for the very slow compression with attack and release time constants of 50 and 1550 msec, respectively, there is still a noticeable deviation from linear behavior.

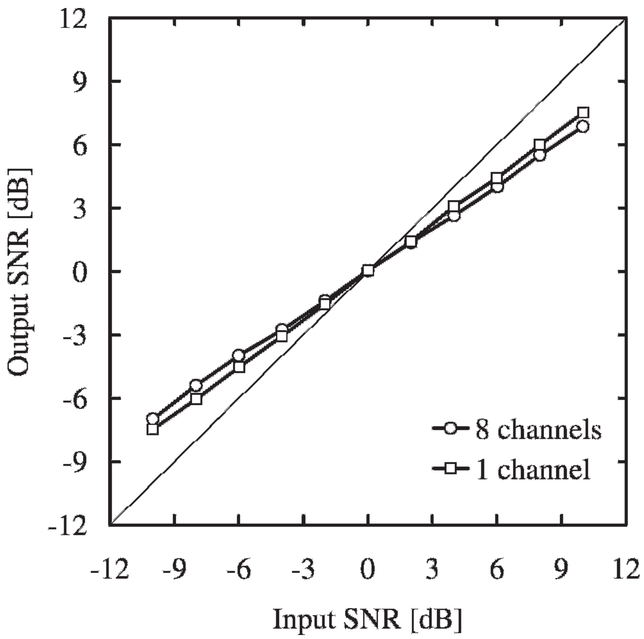

Figure 6. Results of measurements 11–12: signal-to-noise ratio (SNR) for single-channel and eight-channel fast-acting compression, compression ratio 2:1, speech S1 in speech N3.

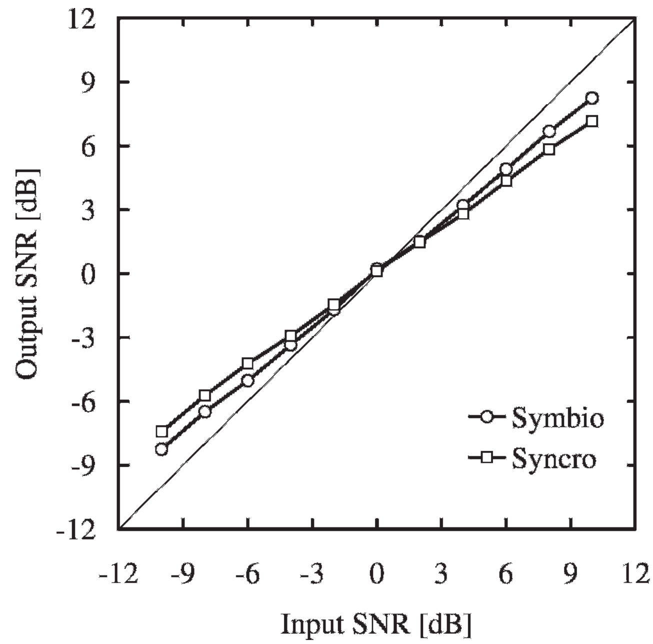

Figure 7. Results of measurements 13–14: signal-to-noise ratio (SNR) for single-channel very fast-acting compression, compression ratio 2:1, speech S1 in speech N3. Measured with Oticon Syncro II and Bernafon Symbio 100.

Number of Compression Channels

Figure 6 shows the results of measurements in which only the number of frequency channels for amplitude compression was varied. In contrast to the data shown in Figures 4 and 5, for these measurements the N signal was N3 (speech). As the number of compression channels increases, the trend line rotates away from the diagonal and Output SNR deviates increasingly from Input SNR. The effect of changing from one to eight channels is rather small, but it should be noted that the eight-channel configuration of the Syncro II device includes a degree of coupling between the level detectors of neighboring compression channels, so the compression action of neighboring channels is not fully independent.

Summary: Effects of Compression Parameters

Variations in the compression ratio, time constants, and number of compression channels do not affect the point at which the Output SNR versus Input SNR line crosses the diagonal, but they do affect the gradient of the line. Thus, more "aggressive" compression, however it is achieved, leads to greater deviations between Input SNR and Output SNR. For realistic variations in compression parameters, the effect on Output SNR can amount to several decibels.

Effects of System Architecture

There are many ways of arranging the signal-processing system in a hearing aid to provide frequency shaping and amplitude compression. Thus it is conceivable that all the effects presented so far are contingent upon the architecture of the particular hearing aid system under test. However, the phenomena we have observed are explainable on the basis of general properties of amplitude-compression systems, without recourse to specifics of architecture. Therefore we would hypothesize that the same phenomena would occur in other architectures. As a first test of this hypothesis we repeated several of the measurements with another hearing aid, having a fundamentally different architecture. The Oticon Syncro II device used so far has a relatively conventional WDRC design in which frequency-response shaping in a filter bank is followed by a quasi-true-RMS level detector and a compressor (Naylor, 1997). In contrast, the Bernafon Symbio 100 device is a so-called Channel-Free system in which the estimated overall loudness of the input signal determines the frequency response and thus the amount of gain from moment to moment (Chung, 2004). Figure 7 shows the results of one measurement with the Symbio device, along with a corresponding measurement with the Syncro device. The two devices were programmed such that all the summary parameters of their compression action were as similar as possible (see bottom row of Table 1). Figure 7 shows that the same phenomena with respect to Output SNR occur in both devices, although the deviation of Output SNR from Input

SNR is considerably diminished with the Symbio device. A similar trend of diminished effect in Symbio was observed for the effects of the compression ratio parameter (data not shown). Neither time constants nor number of channels were accessible for adjustment in Symbio.

As shown in Figure 1, Souza et al (2006) observed similar effects to the ones seen here. The architecture of their system is not described in detail, but it appears to produce effects of greater magnitude than those measured with the Syncro device. In conclusion, it seems likely that the effects seen with Syncro are transferable to systems with other architectures, although the magnitude of the effects may differ from system to system.

Perceptual Relevance

The perceptual relevance of the objective effects presented here remains to be investigated in detail. However, there is some evidence to suggest that the objective effects correspond to perceptual consequences. The literature on the perceptual consequences of compression is characterized by apparently contradictory data. Some of these contradictions might be due to confounding effects of noise types and SNRs used in different studies and may in principle be resolvable in terms of the phenomena presented here. Very few studies have contained conditions directly compatible with testing a correspondence between Output SNR effects and perception, and these remaining studies inevitably contain methodological differences. Olsen et al (2005) measured speech intelligibility for normal-hearing listeners in “fully modulated noise” with linear amplification and with fast-acting compression (using a nonstandard compressor structure, as mentioned above). Results indicated a “compression benefit” (i.e., intelligibility advantage of the compression system relative to the linear system) when the long-term SNR at the system input was negative. The compression benefit increased for increasingly negative input SNRs and was positively correlated to the difference between the long-term SNRs at the output and at the input of the compression system, as determined for their “prediction 2” (the compression system “improved” the long-term SNR when it was negative). This correlation should be viewed with some caution due to the unusual structure of the compressor. Nevertheless, inspection of Olsen et al’s (2005) Table 3 indicates that a substantial “compression benefit” would also have been measured for large negative SNRs at the input, if the compressor control signal had been derived in the normal way from the mixture actually present at the input.

Stone and Moore (2003) measured speech intelligibility for normal-hearing listeners against a single-talker background, when the signal mixture was passed unprocessed or processed by either a slow-acting or a fast-acting compressor (followed by a noise vocoder in all cases). At an Input SNR of +5 dB (where the measurements presented in this article would predict the Output SNR to be reduced by the fast-acting compressor), the speech intelligibility was significantly lower with the fast-acting compressor than with the other processing conditions. Results from this study must also be treated with caution, since the subjects were listening to noise-vocoded signals.

Souza et al (2007) measured speech recognition in noise after linear or fast-acting compression amplification, for signal and noise conditions in which Souza et al (2006) had found Output SNR to be lower than Input SNR. The subjects were hearing impaired, and they performed slightly worse with the fast-acting compression amplification (see Souza et al, 2007, Table 3). Thus there exist data in the literature showing both benefits and disbenefits of fast-acting compression relative to linear amplification or slow-acting compression, opposing trends that are nevertheless consistent with the measured trends in Output SNR versus Input SNR for the conditions we have tested.

We cannot with certainty state that a compression system that for a given condition of S and N at the input “improves” the long-term SNR also provides an improvement in the perception of S, since the system may also be generating other effects that have a negative effect on the perception of S (e.g., in the case of positive Input SNR, N occurring in a pause of S will be amplified more in a compression system than in a linear system and may cause increased forward masking of the following onset of S). The experimental evidence cited above suggests that the direct effects on long-term Output SNR outweigh the secondary side effects of compression, but more evidence is required before drawing any firm conclusions. In particular, the influence of hearing impairment on the balance between long-term SNR effects and other effects has yet to be studied directly.

Implications for Speech-in-Noise Testing

The results shown here raise some questions regarding the ways in which hearing aid–outcome testing is carried out using speech in noise, as choices of signal types and Input SNRs may interact with the compression parameters of the system being evaluated to generate misleading results. In this section we examine this issue, which could have implications for test procedures based on controlling a mixture of speech and noise.

Consider the hypothetical situation in which we wish to compare listener performance with two compression schemes: system A with fast-acting compression and system B with slow-acting compression.

Choosing to focus on speech-in-noise performance as the outcome domain, it would be normal practice in many laboratories to express performance differences between systems A and B in terms of the difference in Input SNR required to provide equal listener performance with both systems. Adaptive or fixed-stimulus procedures would be used to establish the Input SNR needed with each system to achieve, say, 50 percent correct on a standardized speech test (e.g., Hearing in Noise Test [Nilsson et al, 1994] or QuickSIN Speech-in-Noise Test [Killion et al, 2004]).

There are a number of problems with this approach, all concerned with the vital significance of the Input SNR as a determinant of system operation. It is in the nature of compression systems that they affect strong and weak portions of the input signal in different ways. As a consequence, the effects of the two compression systems on the S and N components of the mixture are dependent on the Input SNR at which the measurement is made, and as we have seen in the case of long-term Output SNR, the difference between them may have opposite signs at opposite ends of the Input SNR continuum. Thus the result of the comparison between systems A and B may depend on the Input SNR at which testing takes place. This in turn is highly affected by various factors in the testing protocol.

First, the construction of the speech test and its scoring will influence the SNR region at which listeners need to operate to achieve criterion performance. Altering any of the following common parameters may affect the SNR for criterion performance by as much as several decibels:

- changing the criterion performance level (e.g., from 50 to 80% [Pichora-Fuller et al, 1995; Bronkhorst et al, 2002; Wagener et al, 2003])
- changing from high-context to low-context sentences (Pichora-Fuller et al, 1995)
- changing from scoring words to scoring sentences (Boothroyd and Nittrouer, 1988; Bronkhorst et al, 2002; Wagener et al, 2003)
- changing from a highly intelligible to a less intelligible talker (Cox et al, 1987a, 1987b)
- changing from binaural to monaural listening (Plomp and Mimpen, 1979)
- changing the noise type, for example, from unmodulated to modulated (Wagener and Brand, 2005)
- changing the spatial separation of the speech target and masker (Bronkhorst and Plomp, 1989)

While these effects do not combine linearly, their summed effect is nevertheless substantial.

Different choices of test parameters thus may predetermine the modes of operation of the systems being compared and bias the outcome of the comparison. The reporting of the relative benefits of alternative amplification systems in terms of (input) SNR difference is thus an incomplete report, unless all the test parameters listed above are specified and it is made clear at what input SNR the reference system was tested.

Second, and even more serious, the inherent SNR requirements of hearing-impaired listeners differ over a wide range (e.g., Hagerman, 1984; Lunner, 2003; Wagener and Brand, 2005). Some listeners need a substantial positive SNR to reach criterion performance, while others (including many with impaired hearing) can perform equally well with a substantial negative SNR. Thus, in the test scenario proposed above, and with some plausible combinations of test parameters, we reach the surprising conclusion that some listeners might experience A as better than B (because they are tested at negative Input SNRs) while others might experience B as better than A (because they are tested at positive Input SNRs). For the individual listener, the contrast experienced is not "false" in any sense. The experiment (correctly) fails to produce a consensus result, but if the underlying reasons are not recognized, a misleading conclusion ("A and B are on average equally beneficial") may be drawn.

Effects of this kind may bear some of the responsibility for the general lack of consensus in the literature regarding the relative benefits of different forms of amplitude compression for speech recognition in noise (e.g., Gatehouse et al, 2006a). In addition it is clear that general conclusions about the relative benefits of alternative amplification schemes should be made with caution and be based on tests that sample a variety of listening conditions representative of real-life use.

It should be emphasized that we are not advocating that comparisons of the user benefits of nonlinear amplification systems should be made at equal Output SNR rather than equal Input SNR. According to the argumentation put forward here, the most reliable method would be to measure listener performance with the alternative systems at a number of appropriately spaced, fixed Input SNRs. As a postscript, the arguments in favor of testing with fixed rather than adaptive Input SNRs become even more pertinent if the systems being compared include more strongly nonlinear types of processing such as fine-grained noise reduction or features based on, for example, voice activity detection.

CONCLUSIONS

It is now possible to examine the long-term SNR at the output of a nonlinear system such as a hearing aid with amplitude compression. This provides new insights into the action of such compressors on signal mixtures. The measurements presented here demonstrate that under realistic acoustical conditions, the long-term Output SNR may differ from the Input SNR by several decibels in either direction.

In general terms it was found that the deviation between Input SNR and Output SNR depends on the modulation characteristics of the Signal and Noise and increases with more aggressive compression. The effects were observed in two radically different hearing aid signal-processing structures. More work is required to explore the perceptual significance of these effects, but there are grounds for believing that the difference between Input SNR and Output SNR will be informative about listener performance with a given amplification system.

Speech-in-noise testing to establish the relative benefits of different compression systems for hearing aid users requires careful thought. Effects due to the test's structure and the acoustical signals used, the compressors under test, and especially individual differences between listeners' baseline performance may interact to confound the results, unless care is taken in the design of the test.

Acknowledgments. We wish to acknowledge the significant contributions to the early stages of this work made by our colleagues Thomas Lunner and Thomas Behrens, and we thank Søren Laugesen and anonymous reviewers for their helpful comments on earlier versions of the manuscript.

REFERENCES

American National Standards Institute. (2003) *ANSI S3.22-2003. Specification of Hearing Aid Characteristics*. New York: Author.

Bentler RA, Pavlovic CV. (1989) Transfer functions and correction factors used in hearing aid evaluation and research. *Ear Hear* 10:58–63.

Boothroyd A, Nittrouer S. (1988) Mathematical treatment of context effects in phoneme and word recognition. *J Acoust Soc Am* 84:101–114.

Bronkhorst AW, Plomp R. (1989) Binaural speech intelligibility in noise for hearing-impaired listeners. *J Acoust Soc Am* 86:1374–1383.

Bronkhorst AW, Brand T, Wagener K. (2002) Evaluation of context effects in sentence recognition. *J Acoust Soc Am* 111: 2874–2886.

Chung K. (2004) Challenges and recent developments in hearing aids. Part II. Feedback and occlusion effect reduction strategies, laser shell manufacturing processes, and other signal processing technologies. *Trends Amplif* 8:125–164.

Cox RM, Alexander GC, Gilmore C. (1987a) Development of the Connected Speech Test (CST). *Ear Hear* 8:119S–126S.

Cox RM, Alexander GC, Gilmore C. (1987b) Intelligibility of average talkers in typical listening environments. *J Acoust Soc Am* 81:1598–1608.

Dillon H. (1996) Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times? *Ear Hear* 17:287–307.

Dreschler WA, Verschuure H, Ludvigsen C, Westermann S. (2001) ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology. *Audiology* 40:148–157.

Gatehouse S, Naylor G, Elberling C. (2003) Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol* 42(S1):77–85.

Gatehouse S, Naylor G, Elberling C. (2006a) Linear and nonlinear hearing aid fittings—1. Patterns of benefit. *Int J Audiol* 45:130–152.

Gatehouse S, Naylor G, Elberling C. (2006b) Linear and nonlinear hearing aid fittings—2. Patterns of candidature. *Int J Audiol* 45:153–171.

Hagerman B. (1984) Clinical measurements of speech reception threshold in noise. *Scand Audiol* 13:57–63.

Hagerman B, Olofsson A. (2004) A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. *Acta Acustica* 90:356–361.

Killion MC, Niquette PA, Gudmundsen GI, et al. (2004) Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 116:2395–2405.

Lunner T. (2003) Cognitive function in relation to hearing aid use. *Int J Audiol* 42 (Suppl) 1:49–58.

Naylor G. (1997) Technical and audiological factors in the implementation and use of digital signal processing hearing aids. *Scand Audiol* 26:223–229.

Nilsson M, Soli SD, Sullivan JA. (1994) Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 95:1085–1099.

Olsen HL, Olofsson A, Hagerman B. (2005) The effect of audibility, signal-to-noise ratio, and temporal speech cues on the benefit from fast-acting compression in modulated noise. *Int J Audiol* 44:421–433.

Pichora-Fuller MK, Schneider BA, Daneman M. (1995) How young and old adults listen to and remember speech in noise. *J Acoust Soc Am* 97:593–608.

Plomp R, Mimpen AM. (1979) Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18:43–52.

Souza PE. (2002) Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends Amplif* 6:131–165.

Souza PE, Jenstad LM, Boike KT. (2006) Measuring the acoustic effects of compression amplification on speech in noise. *J Acoust Soc Am* 119:41–44.

Souza PE, Boike KT, Witherell K, Tremblay K. (2007) Prediction of speech recognition from audibility in older listeners with hearing loss: effects of age, amplification, and background noise. *J Am Acad Audiol* 18:54–65.

Stone MA, Moore BCJ. (2003) Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *J Acoust Soc Am* 114:1023–1034.

Stone MA, Moore BCJ. (2007) Quantifying the effects of fast-acting compression on the envelope of speech. *J Acoust Soc Am* 121:1654–1664.

Wagener K, Josvassen JL, Ardenkjaer R. (2003) Design, optimization and evaluation of a Danish sentence test in noise. *Int J Audiol* 42:10–17.

Wagener KC, Brand T. (2005) Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters. *Int J Audiol* 44:144–156.