Department of Veterans Affairs

Journal of Rehabilitation Research
and Development Vol. 30 No. 1, 1993
Pages 82–94

Real-time multiband dynamic compression and noise reduction for binaural hearing aids

Birger Kollmeier; Jürgen Peissig; Volker Hohmann
Drittes Physikalisches Institut, Universität Göttingen, W-3400 Göttingen, Germany

Abstract—A multi-signal-processor set-up is introduced that is used for real-time implementation of digital hearing aid algorithms that operate on stereophonic (i.e., binaural) input signals and perform signal processing in the frequency domain. A multiband dynamic compression algorithm was implemented which operates in 24 critical band filter channels, allows for interaction between frequency bands and stereo channels, and is fitted to the hearing of the individual patient by a loudness scaling method. In addition, a binaural noise reduction algorithm was implemented that amplifies sound emanating from the front and suppresses lateral noise sources as well as reverberation. These algorithms were optimized with respect to their processing parameters and by minimizing the processing artifacts. Different versions of the algorithms were tested in six listeners with sensorineural hearing impairment using both subjective quality assessment methods and speech intelligibility measurements in different acoustical situations. For most subjects, linear frequency shaping was subjectively assessed to be negative, although it improved speech intelligibility in noise. Additional compression was assessed to be positive and did not deteriorate speech intelligibility as long as the processing parameters were fitted carefully. All noise reduction strategies employed here were subjectively assessed to be positive. Although the suppression of reverberation only slightly improved speech intelligibility, a combination of directional filtering and dereverberation provided a substantial improvement in speech intelligibility for most subjects and for a certain range of signal-to-noise ratios. The real-time implementation was very helpful in optimizing and testing the algorithms, and the overall results indicate that carefully designed and fitted binaural hearing aids might be very beneficial for a large number of patients.

Key words: *binaural hearing aids, impaired loudness perception, multi-signal-processor, noise reduction, recruitment phenomenon, sensorineural hearing impairment, speech intelligibility.*

INTRODUCTION

The most common complaints of patients with sensorineural hearing impairment are their reduced ability to understand speech in a noisy environment and their impaired mapping between the sound-pressure level of natural acoustical signals and the perceived loudness of these signals. The impaired loudness perception is often associated with the so-called "recruitment phenomenon," (i.e., the inability of the patient to perceive any sound at low to moderate sound-pressure levels and a steep increase in perceived loudness if the level increases from moderate to high values). Therefore, dynamic compression circuits have traditionally been incorporated in hearing aids (1). They operate on the full input frequency range and/or in several independent frequency bands in order to account for the frequency dependence of the hearing dysfunction. In the literature, however, there has been controversy over the benefit of multichannel compression algorithms (especially if short time constants are involved) in comparison with linear or broadband compression systems (2,3,4).

Address all correspondence and requests for reprints to: Birger Kollmeier, Drittes Physikalisches Institut, Universität Göttingen, Bürgerstrasse 42-44, W-3400 Göttingen, Germany.

Unfortunately, due to the computational expense involved in multiband algorithms, only short speech samples have been used so far to evaluate these systems empirically and to compare their performance with other systems. In addition, most of the compression systems developed so far only operate monaurally (i.e., on the signal for one ear). Thus, the systems can distort the spatial auditory impression, which is primarily determined by binaural hearing (i.e., by listening with both ears). Therefore, a real-time binaural multiband-dynamic-compression algorithm is described and evaluated in this paper that incorporates interaction between both stereo channels to preserve interaural intensity cues. Adjustable interaction between frequency bands is also provided which allows for a parametric transition from a broadband (single-channel) system to a multiband system where all frequency channels are processed separately.

Binaural hearing also contributes significantly to the ability of normal listeners to suppress disturbing noise and to enhance the signal coming out of a desired direction (i.e., the so-called "cocktail party effect"). In addition, a reduction of the perceived reverberation and its negative effect on speech intelligibility is performed by normal listeners who are able to exploit binaural cues (e.g., interaural time and intensity differences) with sophisticated signal-processing strategies in the central auditory system (5). To restore the speech perception abilities of the impaired listener in noisy and reverberant environments, the evaluation and processing of interaural differences might therefore be performed by a "binaural" hearing aid using an intelligent processing scheme that operates on two input signals and provides one or two output signals. Several algorithms of this type have been proposed in the literature that were not necessarily intended for use in hearing aids (6,7,8,9,10,11,12). However, they tend to be very sensitive to small alterations in the acoustical transfer functions, require a high computational complexity, or introduce disturbing processing artifacts.

The directional filter algorithm proposed by the authors (13) minimizes these disadvantages since it is rather insensitive to changes in the acoustical transfer functions and exhibits a limited computational complexity. A real-time implementation is therefore possible, which helps to reduce the artifacts. In non-reverberant acoustical conditions, the algorithm is successful in enhancing a "target speaker" in front of the listener with up to three interfering speakers distributed off the midline. When reverberation is added, however, the performance of the algorithm deteriorates due to processing artifacts. A combination with a scheme for suppressing reverberation is described here that also should extend to reverberant conditions the potential benefit obtainable from this algorithm.

In this paper, the implementation and first results with these algorithms on a multi-signal-processor set-up in real-time is described. After evaluating the binaural multiband-dynamic-compression algorithm, the combination of the directional filter with a dereverberation algorithm that operates on binaural input signals is evaluated. The real-time implementation facilitates the processing of large speech samples and allows for an interactive optimization of the processing parameters as well as an interactive fitting to the requirements of the individual patient.

METHOD

Hardware Set-Up

A block diagram of a hardware set-up is given in **Figure 1**. Three digital signal processors (AT&T WE DSP 32C), each a part of an Ariel PC-32 Digital Signal Processor (DSP) board in a PC-bus slot, are connected with serial high-speed interfaces. A stereo 16-bit A/D (analogue-to-digital) converter is serially connected to the first DSP, while a 16-bit stereo D/A (digital-to-analogue) converter is serially connected to the third DSP. The input microphone signals are either recorded with a dummy head or with miniature microphones located in the outer ear canal of an individual. These signals are amplified,

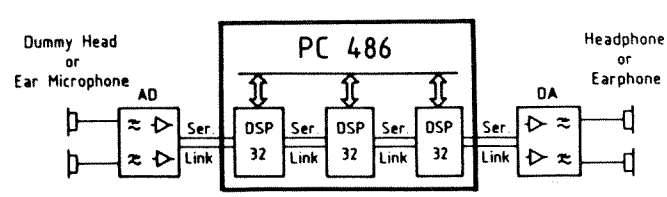

Figure 1.
Block diagram of the hardware set-up employing three Digital Signal Processor (DSP) chips with serial connections to external AD/DA converters.

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993

low-pass filtered and converted to digital. The output signals of the D/A converters are low-pass filtered, amplified, and presented to the subject via headphones or insert earphones. An overlap-add technique (14) is implemented with the three DSPs. The first DSP divides the incoming time signal into overlapping segments, multiplies each time segment with a Hamming window, and extends the segment with additional zeroes before performing a 512-point fast Fourier transform (FFT). The second DSP processes the signals in the spectral domain while the third DSP performs the inverse Fourier transform and the overlapping addition of the filtered time segments in order to reconstruct the time signal.

The three DSP boards are housed in a PC-compatible 486 personal computer. A program library was developed that reflects the high specifications of a multiprocessing system with respect to the coordination of the processors, the data transfer protocol, and the debugging options. To retain the flexibility and the simple structure of the whole software, the high-level routines that structure the whole program system were written in "C" language. On the other hand, to provide an efficient real-time realization of certain routines, the computational intensive parts of the program were written in assembly language. To ensure an effective and time-saving data transfer between the processors, each processor operates on alternating DMA input and output buffers, which may be accessed while simultaneously processing the data from the other data buffers.

Algorithms

Figure 2 gives the block diagram of the algorithm for multiband dynamic compression. Successive short-term spectra are calculated in both stereo channels using Hamming-windowed segments of 408 samples and an FFT length of 512 samples at an overlap rate of 0.5 (distance of successive frames: 204 samples; sample rate: 30 kHz). The subsequent processing is performed in the frequency domain. For each individual ear, linear frequency shaping is provided with a high spectral resolution by multiplying each FFT channel with a prescribed fixed value. In addition, a dynamic nonlinear weighting of the frequency channels is performed in 24 non-overlapping bands with a bandwidth according to the critical bandwidth of the ear (i.e., approximately

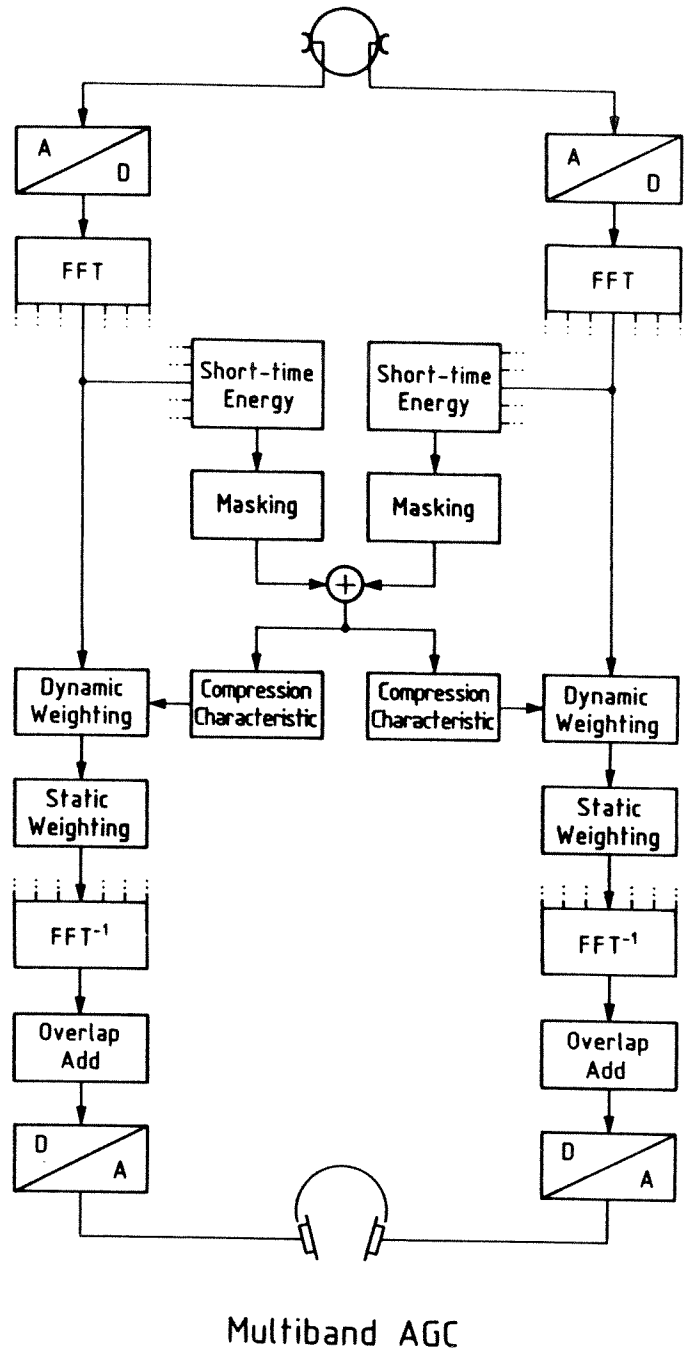

Figure 2.
Block diagram of the multiband compression algorithm.

100 Hz below 500 Hz center frequency and 0.2 × center frequency for frequencies above 500 Hz) (15). Thus, the nonlinear level adjustment is performed with less spectral resolution than the linear frequency shaping.

For each frequency band in each ear, a compression characteristic is prescribed that is computed

as follows: The input energy for each frequency band is obtained by adding up the energies of all FFT channels belonging to the respective frequency band. This value is low-pass filtered with an exponential time window employing different time constants for increasing and decreasing instantaneous energy (i.e., "attack" and "release" time). Subsequently, the masking effect of the energy within a frequency band on adjacent frequency bands is taken into account. Upward spread of masking is realized by attaching ramps to each frequency band with 10 dB per bark toward higher frequencies. Similarly, downward spread of masking is realized by ramps with 25 dB per bark toward lower frequencies. In each band, the respective maximum out of the instantaneous energy within the band and the energy originating from the ramps of adjacent frequency bands is adopted as "effective" input level. Therefore, the level adjustments in the different frequency bands are linked together and the processing artifacts are reduced. The degree of this linkage may be altered by changing the slope values of the ramps between 0 dB per bark (broadband compression) and 50 dB per bark (multichannel compression). Finally, the "effective" energy values from the left and the right stereo channel are added in order to simulate the binaural loudness summation.

The fitting of the compression characteristic to the hearing loss of each patient can be explained by **Figure 3**, which outlines the result of a loudness scaling procedure. The dashed curves denote the level of a narrow-band noise as a function of its center frequency, which produces for normal listeners the loudness sensations "very soft," "comfortably loud," and "very loud." The solid lines denote the respective curves of a listener with high frequency hearing loss. For low frequencies, a relatively high dynamic range is retained, whereas at high frequencies only about 10 dB remains between the impression of "very soft" and "very loud." The aim of the algorithm is to restore the perceived loudness of the individual impaired listener in each frequency band as closely as possible to the perceived loudness of an average normal listener. Therefore, the amplification within each frequency band is adjusted for each "effective" input level to compensate the level difference between the loudness impression of the average normal listener (dashed curves) and the corresponding loudness

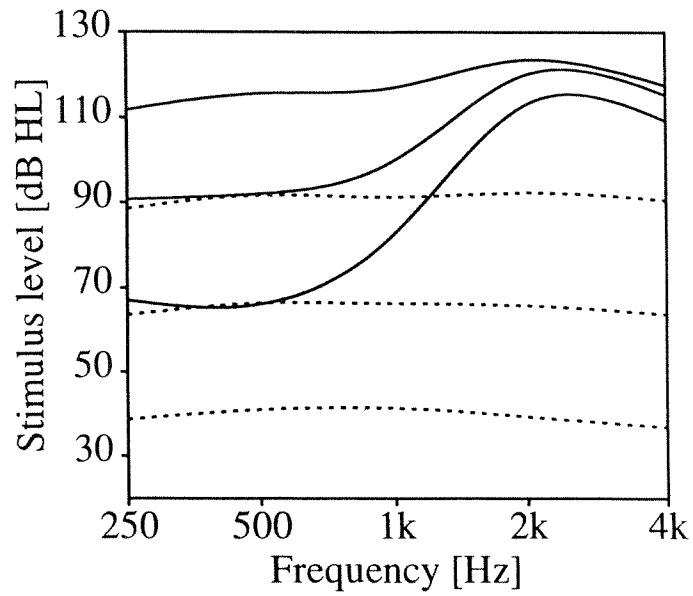

Figure 3.
Equal loudness category contours for subjects with normal hearing (- - - - -) and for one subject with sensorineural hearing impairment (--------). The three curves denote the level of a third-octave-filtered noise required to produce the loudness impression "very soft," "comfortable," and "very loud" as a function of frequency.

impression of the individual impaired listener (solid curves). This amplification is composed out of the (static) linear frequency shaping part (which does not depend on the input level) and the (dynamic) nonlinear compression part. The linear frequency shaping transforms the loudness sensation "comfortably loud" from the impaired listener into the corresponding sensation of the normal listener (i.e., it compensates for the level difference between the intermediate dashed and the intermediate solid curve in **Figure 3**). The nonlinear compression characteristic summarizes all input-level dependent deviations from this static amplification. For example, if the input level in a certain band equates the level belonging to the impression "comfortably loud," then the whole amplification is already provided by the linear frequency shaping part. Therefore, the dynamic compression part would set the amplification value to one. If the input level is higher, this value will decrease, whereas it will increase if the input level is lower. The attack and release time (i.e., decay of the impulse response to 1/e) were both set to 7 ms for all frequency bands and were not adjusted individually.

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993

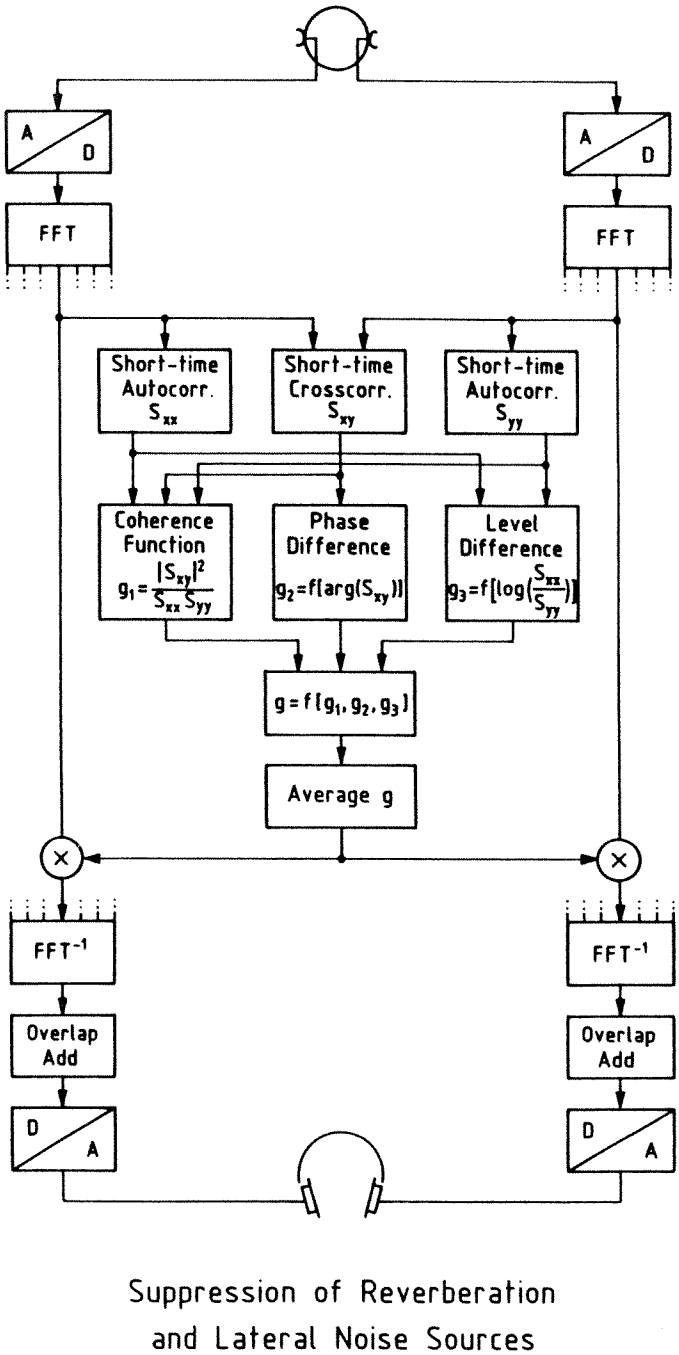

Figure 4.
Block diagram of the algorithm for suppressing reverberation and lateral incident sound sources.

The algorithm for suppressing lateral noise sources and reverberation evaluates averaged interaural time and intensity differences to detect lateral incident sound components. It further evaluates the interaural coherence to detect reverberation processes in the input signals. Frequency bands showing desired values of these interaural parameters (i.e., interaural time and intensity differences close to the desired "reference" values and interaural coherence close to 1) are passed through unchanged, whereas frequency bands with undesired values are attenuated. The lateral noise suppression part of the algorithm is a modification of the algorithm described by Kollmeier and Peissig (13) where instantaneous interaural phase and intensity differences were evaluated. In reverberant situations, however, these instantaneous values within each frequency band do not provide much information about the angle of incidence of a sound source located outside of the reverberation radius. In addition, the normal binaural system is capable of localizing sound sources even in extremely reverberant situations by, for example, evaluating the first wave front and detecting interaural time and level differences of the envelopes. Therefore, the current algorithm evaluates the phase of the short-term cross-correlation and the ratio of the short-term autocorrelation between each pair of frequency bands that are related to the phase and level differences of the input signals' envelopes, respectively. Thus, they should provide more reliable information about the angle of incidence in a reverberant room than the instantaneous interaural phase and intensity differences.

A block diagram of the algorithm is given in **Figure 4**. As above, the incoming signal is segmented, windowed, padded with zeros, Fourier-transformed and back-transformed after processing in the frequency domain. Within each frequency band, the short-term auto- and cross-correlation is computed for the left and the right stereo channel with an exponential weighting window as follows: If X and Y denote the complex output signals of the bandpass filters at the right and left stereo channel, respectively, n denotes the index of the time, and α denotes a coefficient between 0 and 1, we can write:

$$S_{xx}(n) = (1-\alpha)|X(n)|^2 + \alpha S_{xx}(n-1)$$

$$S_{yy}(n) = (1-\alpha)|Y(n)|^2 + \alpha S_{yy}(n-1)$$

$$S_{xy}(n) = (1-\alpha)X(n)Y^*(n) + \alpha S_{xy}(n-1)$$

From the values S_{xx}, S_{yy}, and S_{xy}, the interaural phase and level differences of the signal envelope and the interaural coherence are computed in each frequency band. The respective functions f_1 and f_2 are used to calculate weighting factors g_2 and g_3

from these values. The shape of f_1 and f_2 determines both the range of incident angles for attenuation as well as the maximum attenuation within this region. They are obtained by measurements and may be optimized interactively later on. The weighting factor g_1 is directly given by the short-term coherence. By combining the weighting factors g_1, g_2, and g_3, the performance of the algorithm can be changed to suppress either reverberation or lateral sound sources or to perform a combination of both. In order to suppress processing artifacts, the final weighting factors g are averaged over adjacent frequency channels. If the processing parameters are adjusted properly, the algorithm yields very natural-sounding output signals and performs a satisfactory suppression of reverberation and lateral incident sounds.

Subjects

Six subjects with sensorineural hearing impairment, aged between 25 and 89 years with different degrees of high frequency hearing loss, participated in this study. All subjects were clinically examined to rule out a middle-ear dysfunction and to classify the hearing loss to be of cochlear origin with a positive recruitment phenomenon. The audiometric thresholds at 500 Hz and 4 kHz are given in **Table 1**. In addition, the binaural speech intelligibility threshold is provided, that is, the signal-to-noise ratio for 50 percent correct performance in a German monosyllable rhyme test in speech-simulating, continuous noise (16). For a prescription of the dynamic compression algorithm, a loudness scaling method was performed with third-octave-bandpass-filtered noise. The subject's task was to associate each stimulus with a subjective loudness category ("very soft," "soft," "comfortable," "loud," "very loud") and to further subdivide each category into 10 subcategories. This procedure yields a loudness scale between 0 and 50 partitioning units (17,18). The residual dynamic range (i.e., the difference in level between the loudness categories "very loud" and "very soft") is also included in **Table 1** for each audiometric frequency and both ears.

Assessment Methods

To assess the subjective quality of different versions of the hearing aid algorithms, recorded materials from different acoustic situations were presented to the subjects with the respective processing condition. All materials were either dummy-head recorded using the "Göttingen" dummy-head or using stereophonic miniature microphones inserted in the outer ear-canal of a human listener. The subjects were allowed to listen to a combination of acoustic situation and processing scheme for as long as they desired. They were asked to assess the subjective transmission quality within a scale of five

Table 1.
Audiometric data and residual dynamic range derived from the loudness scaling experiment (in parentheses) for six impaired listeners. The binaural speech intelligibility threshold in noise obtained with a rhyme test and the individual sentence intelligibility scores for the evaluation of the multiband compression are also included.*

			Hearing Loss (Residual Dynamic Range) Right (dB)		Left (dB)		Rhyme Test Threshold (dB)	Sentence Scores** (% correct)		
Subject	Age	Sex	500 Hz	4 kHz	500 Hz	4 kHz		unp.	lin.	comp.
JJ	25	M	55 (40)	105 (10)	60 (40)	105 (8)	3.0	35	38	24
JS	71	F	45 (35)	*** (15)	20 (50)	25 (50)	6.1	56	48	56
WH	68	M	45 (50)	75 (20)	50 (40)	80 (15)	2.0	36	48	0
RP	52	F	30 (40)	55 (20)	30 (20)	40 (20)	3.0	26	48	0
HS	89	M	55 (10)	*** (10)	40 (40)	60 (20)	2.6	23	36	29
WD	72	F	40 (40)	65 (40)	35 (50)	65 (30)	3.5	26	35	54

*For normal listeners, the average residual dynamic range is 50 dB and the speech intelligibility threshold in noise is −5 dB.
**Test conditions: unp. = unprocessed; lin. = linear frequency shaping without compression; comp. = linear frequency shaping with compression.
***No threshold measurable.

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993

categories ("bad," "poor," "reasonable," "good," "excellent").

Speech intelligibility was measured for a subset of the acoustical situations mentioned above using an open German sentence test recorded on compact disc (19). The subject's task was to repeat the whole sentence, and the number of correctly repeated words was scored. A complete test consisted of 10 short sentences. For intelligibility measurements with the dynamic compression algorithm, a dummy-head recording of cafeteria noise was used as background noise, which was added to a dummy-head recording of the speech material alone at a fixed signal-to-noise ratio. For assessing the noise reduction and dereverberation algorithm, a dummy-head recording of the speech signal and the interfering noise was performed in a reverberant room with a reverberation time of 2 to 3 sec. The desired signals (i.e., running speech for quality judgments and test sentences for speech intelligibility test) were radiated with a loudspeaker directly in front of the dummy-head at a distance of 1.5 m. The interfering noise was running speech radiated 30° from the left of the midline. The speech level was always adjusted to match the most comfortable listening level for each individual subject.

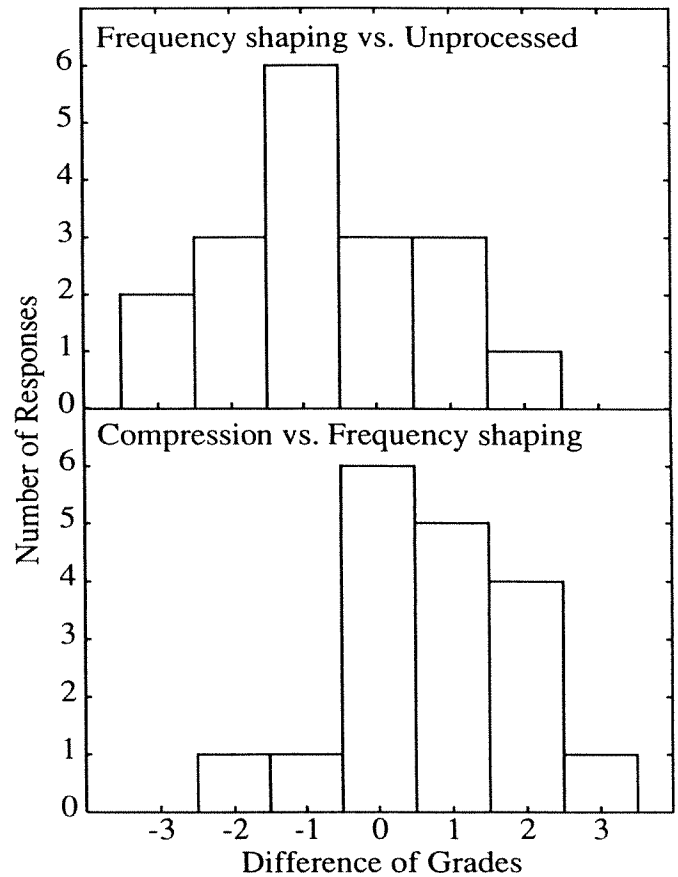

Figure 5.
Quality judgments of different versions of the compression algorithm for six impaired listeners and three different simulated acoustical situations. The upper panel gives the difference in grades between the condition with static linear frequency shaping and no processing. The lower panel gives the difference in grades between static linear frequency shaping plus compression versus shaping alone. Grades ranged from 1 ("bad") to 5 ("excellent").

RESULTS

Dynamic Compression Algorithm

For assessing the subjective quality of the dynamic compression algorithm, three dummy-head recordings of typical acoustical conditions were used: a sample of traffic noise, a loud doorbell presented in soft background noise, and a sample out of a string quartet by Schubert. All listening samples were recorded with stereophonic inserted ear-level microphones in real situations and were presented unprocessed, processed with linear frequency shaping alone, and with linear frequency shaping including compression. The sound samples were presented to the subjects with a Sennheiser HD 25 headphone. At the beginning of each session, an overall level adjustment of up to 10 dB was applied to match the average presentation level to the most comfortable listening level.

Figure 5 shows the differences in subjectively assessed transmission quality (expressed as grades ranging from 1, "very poor," to 5, "excellent") between the different processing conditions for all subjects and all three simulated acoustical situations. The upper panel of **Figure 5** gives the score difference between linear frequency shaping and the unprocessed version. On the average, the unprocessed version is preferred. However, since the range of scores varies considerably, no significant advantage or disadvantage of linear frequency shaping versus unprocessed speech can be derived from these score differences. The subjects attributed their preference for the unprocessed condition to not being accustomed to high frequencies with their own hearing aid. Specifically, the loud doorbell caused

uncomfortably loud sensations in the processed version, while the background noise was not audible at all. This effect was less prominent for the unprocessed version. The lower panel of **Figure 5** gives the difference in grades between linear frequency shaping, including compression versus linear frequency shaping. Obviously, the additional compression is judged to be positive due to the limitation of annoying acoustical components at high frequencies. This observation is quite unexpected for normal listeners who perceive a deterioration of transmission quality and an increase of processing artifacts caused by rapid dynamic compression. However, these artifacts appear to be inaudible for impaired listeners.

For measuring speech intelligibility, each subject was tested with two lists of ten sentences in each of the different processing conditions using cafeteria background noise. The average scores for each processing condition and each subject are included in **Table 1**. The difference in speech intelligibility score between the processed version with linear frequency shaping and the unprocessed version is given in the upper panel of **Figure 6**. On the average, intelligibility increases for linear frequency shaping. This effect is quite contrary to the assessed subjective preference of the unprocessed condition (*see* upper panel in **Figure 5**). However, the effect is rather small, since the interfering noise has approximately the same long-term spectrum as the speech signal. The lower panel of **Figure 6** gives the differences in speech intelligibility between the dynamic compression with linear frequency shaping versus linear frequency shaping alone. With few exceptions, intelligibility is increased by the addition of the dynamic compressor. These exceptions are caused by an erroneous fitting of the compressor characteristic for one subject; the loudness scaling yielded nearly the same level for the loudness categories "comfortable" and "very loud." Thus, the algorithm performs a clipping in all frequency channels, which nearly completely suppresses speech in the presence of an interfering noise and causes a drastic decrease in speech intelligibility.

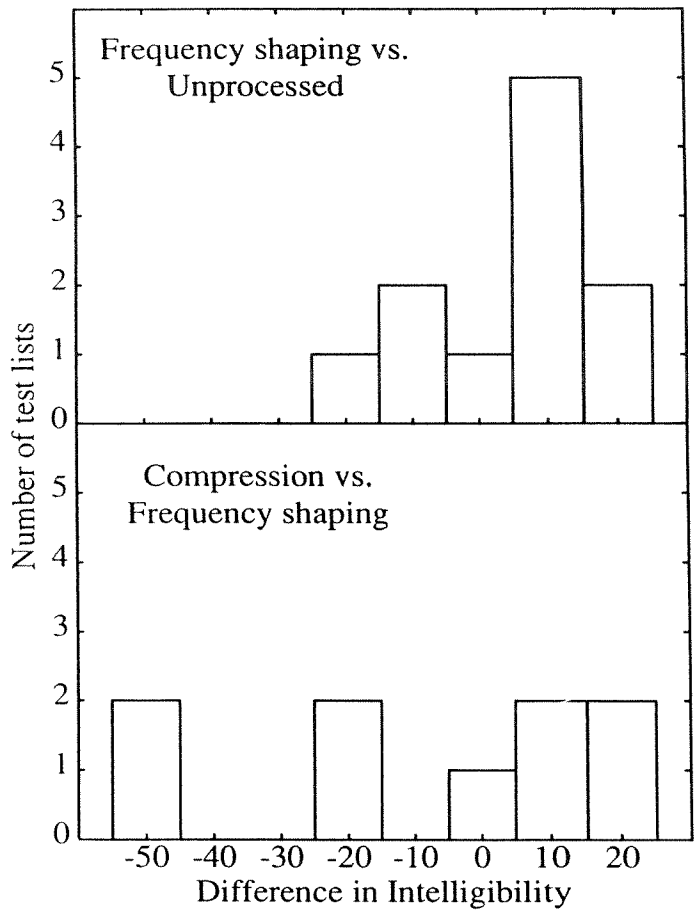

Figure 6.
Difference in speech intelligibility for static linear frequency shaping versus no processing (upper panel) and shaping plus compression versus shaping alone (lower panel). The score for each test list for each subject is counted for the three different processing conditions.

Noise and Reverberation Suppression

To evaluate the performance of the algorithm to suppress lateral noise sources and reverberation, an acoustic situation was simulated by dummy-head recordings in a reverberant room employing one target speaker and one interfering speaker (see above). The signal-to-noise ratio was individually adjusted for each subject within a range of −5 dB to +2 dB in order to obtain a speech intelligibility of approximately 50 percent for the binaural unprocessed condition. **Figure 7** gives the difference in subjective assessment of the transmission quality between the dereverberation algorithm and the unprocessed condition (upper panel) and between the combination of dereverberation and directional filter as compared with the unprocessed condition. Note that linear frequency shaping without dynamic compression was provided in all conditions, including the reference situation. For the dereverberation algorithm, five out of six subjects graded the quality of the processed signal as better than the unproc-

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993

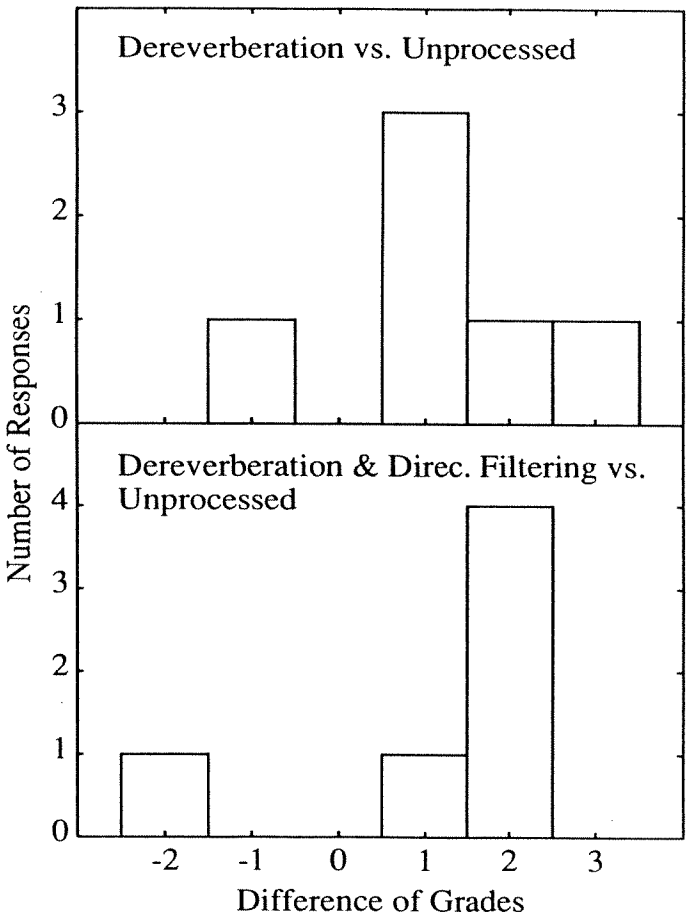

Figure 7.
Quality judgments of different versions of the interference suppression algorithm for six impaired listeners. The upper panel gives the difference in grades between the condition of suppression of reverberation and no processing. The lower panel gives the difference in grades for the combination of dereverberation and the suppression of lateral noise sources (i.e., directional filtering) versus no processing. Linear frequency shaping is always provided. Grades ranged from 1 ("bad") to 5 ("excellent").

essed material by at least one point. After the addition of the directional filter, four of six subjects reported an improvement of two grades as compared with the unprocessed version. Only one subject (JJ) reported better quality of the unprocessed version as compared with the dereverberation algorithm with and without additional directional filtering. This subject was the most severely impaired subject tested, and exhibited a very limited dynamic range (*see* **Table 1**). Apparently, the spectral changes introduced by the algorithms caused the speech signal to move out of this limited range.

Figure 8 gives the results of the speech intelligibility tests as the percentage of correctly repeated words. The first two bars for each subject give the results for the unprocessed, linear frequency shaped material, presented monaurally (first bar) or binaurally (second bar). Subject HS was only tested binaurally. Three out of five subjects exhibit a binaural gain in intelligibility compared with the monaural, unprocessed version. The binaural system of these subjects obviously manages to suppress parts of the interference caused by reverberation and interfering speech. However, subjects RP and JS exhibit a decrease in intelligibility if speech is also presented on the "worse" ear, indicating that the distorted internal representation of the input signals provided by this ear causes a "binaural confusion" rather than a binaural enhancement effect.

The third and fourth bar in **Figure 8** denote the intelligibility score for the dereverberation algorithm where the output signal is presented monaurally or binaurally to the subject, respectively. Compared with the linear shaped, unprocessed material (fourth bar versus second bar), a gain in speech intelligibility is obtained only for subject HS. This finding is consistent with a remark by Allen, et al. (6) that dereverberation algorithms tend to increase speech quality but not to improve speech intelligibility. However, after adding the directional filter, all subjects (except subject JJ) achieved a higher intelligibility for the monaural presentation than for the unprocessed version (fifth bar versus first bar). For the binaural presentation, however, no unambiguous conclusion can be drawn (cf. sixth bar versus second bar): three subjects (WH, WD, and JS) exhibited only a small change in intelligibility which is not significant. Only two subjects (RP and HS) obtained a significant gain in speech intelligibility of 25 percent with the combination of dereverberation and directional filtering.

The overall results from our subjects with various degrees of hearing impairment imply that the benefit obtainable for each individual listener from the preprocessing strategies described here depends on the hearing loss of the individual, the residual dynamic range in the high frequency region, and the signal-to-noise ratio of the test situation. Specifically, the two subjects with the smallest residual dynamic range at 4 kHz (subjects JJ and WH) exhibited the least benefit from the suppression of lateral noise sources and reverberation. This

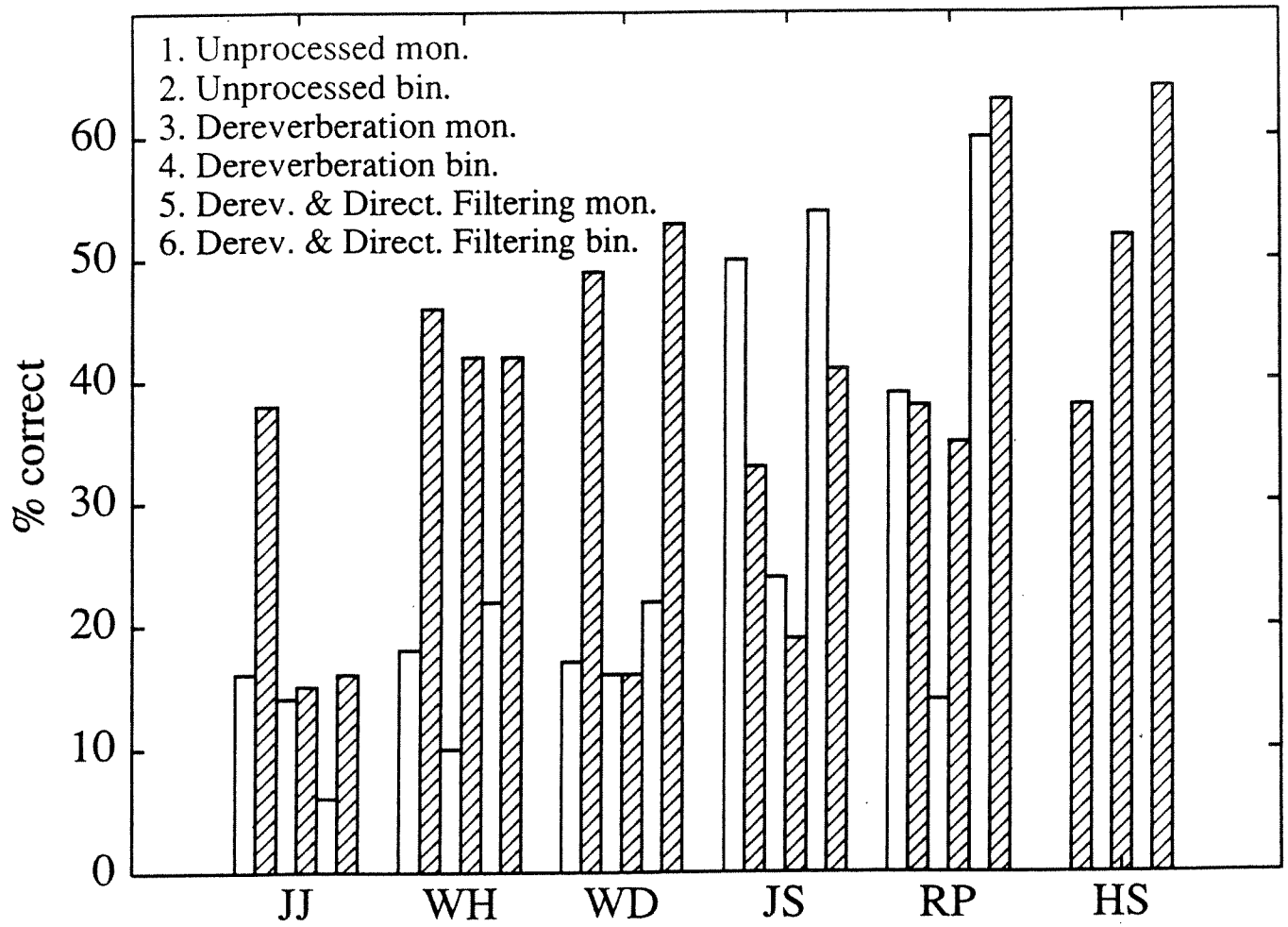

Figure 8.
Speech intelligibility results for different versions of the interference suppression algorithm for six impaired listeners. For each subject, scores were obtained for listening monaurally with the respective "better" ear and for listening binaurally (hatched). Three processing conditions were employed that all incorporated linear frequency shaping: *a*) unprocessed (columns 1 and 2); *b*) suppression of reverberation (columns 3 and 4); and, *c*) suppression of reverberation including suppression of lateral noise sources (columns 5 and 6).

effect might be due to the processing artifacts caused by suddenly switching on and off different frequency bands. They might be more distracting and disturbing if the remaining dynamic range is small. The subjects with the largest residual dynamic range at 4 kHz (WD and JS) were tested with the smallest signal-to-noise ratio of −2 dB. Their comparatively small gain in intelligibility provided by the algorithm might be explained by the unfavorable test condition, because the performance of the noise suppression algorithm decreases if the signal-to-noise ratio is decreased to values close to the speech reception threshold in noise of the normal listener.

DISCUSSION

Implementation of the Algorithms

The real-time implementation of the digital hearing aid algorithms proved to be very helpful in the developing and testing phase, where a number of processing parameters could interactively be adjusted in order to minimize the processing artifacts. For the dynamic compression algorithm, for example, musical tones and a perceivable roughness of the output signal occur if small time constants and no interactions between adjacent bands are involved. In addition, a small dynamic range of the output signal can only be achieved at the cost of

Journal of Rehabilitation Research and Development Vol. 30 No. 1 1993

deteriorating the transmission quality for normal listeners. Fortunately, impaired listeners do not necessarily perceive these alterations as a degradation of speech quality.

The real-time implementation also enabled interactive changes of processing parameters while fitting the algorithms to the requirements of the individual patient. Although the parameters of the compression algorithm were primarily prescribed by the loudness scaling results, adjustments of the overall level of up to 10 dB were required to adjust the output level of the algorithm to the most comfortable listening level of the individual subjects. This difference between prescribed and perceived loudness is due primarily to the loudness summation in realistic broadband signals (such as speech) which is not accounted for by the original fitting method based on third-octave-band loudness scaling values. In our algorithm, only a rough estimate of broadband loudness summation is provided by accounting for upward spread of masking and downward spread of masking. Ideally, more precise ways of estimating the overall loudness for a broadband signal from its spectral contributions should be incorporated. Although quite accurate models of loudness perception have been developed on the basis of relational scales, such as the sone-scale (20), a quantitative model based on categorical loudness perception has yet not been developed (18).

A considerable disadvantage of the real-time system described here is the specialized software that had to be written for each of the signal processors and for the host processor. Although the flexibility and portability of the software was increased by programming the general structure in a high-level language (C language) and programming only time-critical parts in assembly language, the software is still processor-dependent and a migration toward more powerful DSP chips might be difficult. A further disadvantage of distributing the signal processing tasks over three DSP chips is the considerable delay between the input signal and the output signal, which amounted to approximately 50 ms in our case. This delay results from the transfer of blocks between the AD/DA converters and the three signal processors and from the overlap-add technique, which operates on successive time frames. Therefore, the use of the current system as a master hearing aid is limited, since the delay between auditory and visual input might already deteriorate the ability of the patients to use lip reading to aid their perception of speech.

Dynamic Compression Algorithm

One important feature of the implemented compression algorithm is the separate adjustment of the static, linear frequency shaping and the nonlinear dynamic compression. While the former is performed with the maximum frequency resolution of approximately 60 Hz, the latter is performed at a much broader frequency resolution that corresponds to the critical bandwidth of the ear. In addition, the effective frequency resolution for the nonlinear compression can be altered by using different slope values when accounting for upward and downward spread of masking. If these slopes are assumed to be very flat, all frequency channels are synchronized and a broadband compression will effectively result. The values used in our algorithm reflect approximate values for normal listeners in psychoacoustical experiments.

By assessing separately the effect of linear frequency shaping and dynamic compression, it could be demonstrated that linear frequency shaping was subjectively judged to deteriorate speech quality, although speech intelligibility in noise increased. The negative assessment is primarily due to the subjects being unaccustomed to a high gain at high frequencies in hearing aids. Therefore, additional compression is subjectively judged to improve the speech quality. In addition, speech intelligibility is not deteriorated by the additional compression if the processing parameters are carefully selected. These results are in agreement with studies that multiband dynamic compression does not significantly improve speech intelligibility (4,21,22), but are not consistent with Plomp's notion (3) that dynamic compression has a negative effect on speech intelligibility. However, the time constants employed here were relatively large and the cross-channel interaction provided comparatively smooth transfer functions. Hence, only a small detrimental effect of dynamic compression on speech intelligibility would have been expected on the basis of Plomp's arguments. Therefore, our data cannot be used to argue against Plomp's conclusions that small time constants and a large number of independent channels should not be employed for hearing aids.

Noise and Reverberation Suppression

The algorithm for suppressing lateral noise sources and reverberation by exploiting binaural cues appears to operate quite efficiently even under adverse acoustical conditions (i.e., a reverberant environment). However, a trade-off exists between the potential of the algorithm to suppress interferences and its potential to preserve the quality of the transmitted speech (i.e., the absence of artifacts). High attenuation values of lateral sound sources imply large temporal and spectral fluctuations of the effective transfer function which inevitably produce processing artifacts. Hence, a realistic compromise between both specifications under different acoustical conditions has to be found empirically. This can be performed only if an interactive change of the processing parameters is possible, as in the real-time implementation described here.

Another important point is the performance of the algorithm as a function of the signal-to-noise ratio of the input signal: for high and intermediate signal-to-noise ratios, the algorithm operates quite well and yields virtually no artifacts. For low signal-to-noise ratios, however, the artifacts increase and no benefit is obtained from the algorithm as compared with the unprocessed situation, even for normal listeners. Therefore, the patients with moderate hearing loss who were tested at low signal-to-noise ratios obtained only a small benefit from the algorithm. However, patients with more severe hearing losses did profit from the algorithm at more favorable signal-to-noise ratios. In addition, it should be noted that the primary goal of the algorithms would be to enhance speech under conditions where normal listeners would not have difficulties understanding speech while impaired listeners would. In these situations, the signal-to-noise ratio is comparatively high and the algorithm would therefore be beneficial.

In conclusion, the algorithms presented here that are intended to be used in a "true binaural" hearing aid appear to have a large potential for aiding persons with hearing impairment. Specifically, the use of binaural information for suppressing reverberation and interfering noise appears promising. In addition, the real-time implementation of the algorithms is a salient tool for developing, testing, and assessing these algorithms. It is also a first step toward implementing these algorithms in future "intelligent" digital hearing aids.

ACKNOWLEDGMENT

The authors want to express their thanks and appreciation to the subjects who participated in this studies. Technical assistance by E. Rohrmoser, G. Kirschmann-Schröder, L. Martens, T. Hindermann and K. Werder is gratefully acknowledged.

REFERENCES

1. Working Group on Communication Aids for the Hearing-Impaired. Speech perception aids for hearing-impaired people: Current status and needed research. J Acoust Soc Am 1991;90:637-85.
2. Bustamante DK, Braida LD. Multiband compression limiting for hearing-impaired listeners. J Rehabil Res Dev 1987;24(4):149-60.
3. Plomp R. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation transfer function. J Acoust Soc Am 1988;83:2322-7.
4. Villchur E. Multiband compression processing for profound deafness. J Rehabil Res Dev 1987;24:135-48.
5. Durlach NI, Thompson CL, Colburn HS. Binaural interaction in impaired listeners. A review of past research. Audiology 1981;20:181-211.
6. Allen JB, Berkley DA, Blauert J. Multimicrophone signal-processing technique to remove room reverberation from speech signals. J Acoust Soc Am 1977;62:912-15.
7. Bodden M. Bewertung der störsprecherunterdrückung mit einem cocktail-party-prozessor. In: Fortschritte der akustik—DAGA '92. DPG-Kongreß-GmbH, Bad Honnef, 649-52.
8. Gaik W, Lindemann W. Ein digitales richtungsfilter, basierend auf kunstkopfsignalen. In: Fortschritte der akustik—DAGA '86. DPG-Kongreß-GmbH, Bad Honnef, 1986: 721-24.
9. Koch R. Störgeräusch-unterdrückung für hörhilfen—ein adaptiver cocktail-party-prozessor. In: Fortschritte der akustik—DAGA '90. DPG-Kongreß-GmbH, Bad Honnef, 1990: 1019-22.
10. Kollmeier B. Me methodik, modellierung und verbesserung der verständlichkeit von sprache. Göttingen, Germany: Habilitationsschrift, Universität Göttingen, 1990.
11. Peterson PM, Wei SM, Rabinowitz WM, Zurek PM. Robustness of an adaptive beamforming method for hearing aids. Acta Otolaryngol Suppl 1990;469:85-90.
12. Strube HW. Separation of several speakers recorded by two microphones (cocktail-party-processing). Sig Proc 1981;3:355-64.
13. Kollmeier B, Peissig J, Hohmann V. Binaural noise-reduction hearing aid scheme with real-time processing in the frequency domain. Scand Audiol Suppl. In press.
14. Allen JB. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. IEEE Trans Acoust Speech Sig Process 1977;25:235-8.

15. Zwicker E, Terhardt E. Analytical expressions for critical band rate and critical bandwidth as a function of frequency. J Acoust Soc Am 1980;68:1523-5.
16. v.Wallenberg EL, Kollmeier B. Sprachverständlichkeitsmessungen für die audiologie mit einem reimtest in deutscher sprache: erstellung und evaluation von testlisten. Audiol Akustik 1989;28:50-65.
17. Hellbrück J, Moser LM. Hörgeräte-audiometrie: ein computerunterstütztes psychologisches verfahren zur hörgeräteanpassung. Psychol Beiträge 1985;27:494-508.
18. Heller O. Oriented category scaling of loudness and speech audiometric validation. In: Schick A, et al., editors. Contributions to psychological acoustics. Oldenburg, Germany: Bibliothek und Informationssystem der Universität Oldenburg, ISBN3-8142-0400X, 1991.
19. Niemeyer W. Sprachaudiometrie mit sätzen I: grundlagen und testmaterial einer diagnostik des gesamtsprachverständnisses. HNO(Berl) 1967;15:335.
20. Zwicker E. Procedure for calculating loudness of temporally variable sounds. J Acoust Soc Am 1977;62:675-82.
21. Moore BCJ, Glasberg BR. A comparison of four methods of implementing automatic gain control (AGC) in hearing aids. Br J Audiol 1988;22:93-104.
22. Kollmeier B. Speech enhancement by filtering in the loudness domain. Acta Otolaryngol Suppl 1990; 469:207-14.