Department of Veterans Affairs

Journal of Rehabilitation Research and Development Vol. 29 No. 2, 1992
Pages 39–56

Spectral feature enhancement for people with sensorineural hearing impairment: Effects on speech intelligibility and quality

Michael A. Stone and Brian C.J. Moore
Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB England

Abstract—People with sensorineural hearing loss often have difficulty understanding speech in background noise at speech-to-noise ratios (0 to +6 dB) for which normally hearing people have little difficulty. Spectral analysis of speech in noise at these ratios typically shows that the major spectral prominences in the speech (formants) are well represented, but the spectral valleys between the formants are filled with noise. Hearing impaired people have a reduced ability to pick out the spectral prominences, and are more affected by the noise filling in the valleys, partly because of their reduced frequency selectivity. This paper describes a 16-channel bandpass filter bank, implemented in analog electronics, that attempts to enhance spectral features of speech in noise to improve intelligibility for the hearing impaired. Each channel generates an 'activity function' that is proportional to the magnitude of the signal envelope in that channel, averaged over a short period of time. A positively weighted activity function from the nth channel is combined with negatively weighted functions from channels $n-2$, $n-1$, $n+1$, and $n+2$, giving a correction signal used to control the gain of the bandpass signal in the nth channel. Recombining the bandpass signals results in an enhancement of spectral features of the speech in noise. Two different experiments are described here, one using the activity function as described, and the other using a nonlinear transform of the activity function. In both experiments, several different weighting patterns were used in calculating the correction signal. The intelligibility of speech in noise processed by the system was measured for subjects with moderate sensorineural hearing loss. In both experiments, no improvement in intelligibility was found. However, subjective ratings of the stimuli used in Experiment 2 indicated that some subjects judged the processed stimuli to have both higher quality and higher intelligibility than unprocessed stimuli.

Key words: *bandpass filters, hearing-impaired, sensorineural hearing loss, spectral analysis, speech intelligibility, speech-to-noise ratios.*

INTRODUCTION

People with moderate sensorineural hearing impairment often complain of difficulty in understanding speech in noise. They can understand speech reasonably well in one-to-one conversation in a quiet room, but they have great difficulty when there is background noise or reverberation, or when more than one person is talking. This difficulty appears to be related to a variety of abnormalities in the perception of sound (1) and it persists even when the speech is amplified sufficiently (by a hearing aid) to be well above the threshold for detection (1,2).

One well-documented abnormality of perception associated with sensorineural hearing loss is reduced frequency selectivity. Frequency selectivity refers to the ability of the ear to resolve a complex sound into its frequency components. This ability is often characterized by describing the ear as containing a bank of overlapping bandpass filters, known as the auditory filters (3). The characteristics of these filters for normally hearing people are reasonably well established (4,5,6,7,8). Sensorineural hearing loss, and particularly cochlear hearing loss, is associated with broader-than-normal auditory filters, that is,

Address all correspondence and requests for reprints to: Michael A. Stone, Department of Experimental Psychology, University of Cambridge, Downing St., Cambridge CB2 3EB, England.

DOI: 10.1682/JRRD.1992.04.0039

reduced frequency selectivity (9,10). This has two main consequences for speech perception. First, it means that the spectral features of speech sounds are less well resolved. For example, the spectral contrast (dB difference between peaks and valleys in the spectrum) required for vowels to be identified is greater for impaired than for normal listeners (11). Secondly, it means that background noise has a greater masking effect, since broader filters generally pass more background noise. Several studies have shown that the ability to understand speech in noise is correlated with measures of auditory filter bandwidth, although the effects of filter bandwidth are difficult to separate from the effects of a simple loss of sensitivity to weak sounds, since the two are highly correlated (12,13,14).

Several authors have described attempts to improve speech intelligibility for the hearing impaired by enhancement of spectral features (15,16,17). Boers (15) processed a set of sentences to increase the level differences between peaks and valleys in the spectrum. Noise was added *after* the processing, and the effects of the processing were assessed by measuring the speech-to-noise ratio required for 50 percent of the words to be understood. Overall, the processing reduced intelligibility, although two impaired listeners did show a slight improvement with the processed signals. Of course, this kind of processing would not be feasible with naturally occurring signals; with these, the speech would already be contaminated with noise, and the processing would have to operate on the speech-plus-noise.

Summerfield, *et al.* (17) synthesized 'whispered' speech sounds, and investigated the effect of narrowing the bandwidths of the formants (spectral resonances) used in synthesis. Narrowing these bandwidths led to both sharper spectral peaks and greater peak-to-valley ratios. However, it had only small effects on speech intelligibility; identification of consonants at the end of syllables tended to be slightly better when the formant bandwidths were half of their nominal normal values. Speech intelligibility in noise was not tested.

Simpson, *et al.* (16) described a method of digital signal processing of speech in noise in order to increase differences in level between peaks and valleys in the spectrum. Before spectral enhancement, the spectra were smoothed to eliminate minor peaks and ripples, using smoothing filters based on the properties of the auditory filters in normal ears. The enhancement was also done on a frequency scale related to the frequency resolution of normal ears (5). The intelligibility of the speech in noise was measured using subjects with moderate sensorineural hearing loss. The results showed small but reasonably consistent improvements in speech intelligibility for the processed speech. Although this form of processing appears promising, it is not easy to implement in a wearable hearing aid, at least with current technology; the processing used by Simpson, *et al.* (16) ran at about two hundred times real time on a reasonably fast laboratory computer (Masscomp 5400 with floating-point accelerator).

In this paper we describe a method of processing speech in noise to enhance spectral features that is similar in concept to the processing used by Simpson, *et al.* (16), but which uses analog processes running in real time. This type of processing would be easier to implement in a wearable hearing aid in the near future. However, practical constraints limited the number of filters used in the processing to 16, much less than the 128 used by Simpson, *et al.*

It is useful to make a distinction between two types of rationales for performing spectral enhancement. One rationale is based on a consideration of spectrograms of speech in noise at the sorts of speech-to-noise ratios (below +6 dB) where impaired people have problems. These spectrograms usually give a fairly clear representation of the major spectral prominences (formants) in the speech, but the spectral 'valleys' between the formants are filled with noise. It appears that normally hearing people are hardly affected by the noise in the valleys, but hearing impaired people are badly affected, probably because of their broader auditory filters. Hence, processing which reduces the level of noise between the formants may reduce the masking effect and may improve speech intelligibility. A possible mechanism for this improvement is as follows: the perceived frequency of a given formant may be partly determined by the time pattern at the outputs of the auditory filters tuned close to the formant frequency (18,19). Background noise disturbs this time pattern, which may lead to reduced accuracy in determining the formant frequency. Suppressing the noise between formants may partly 'clean up' the time pattern, restoring accuracy.

The second rationale for spectral enhancement is that it can, within limits, produce in the impaired ear a representation of the spectral details of sounds closer to that found in a normal ear with unprocessed sounds. The recognition of speech sounds requires a determination of their spectral shape. Spectral shapes can be described as being composed of spectral modulations of different rates; the spectral envelope can be decomposed into a series of sinusoidal spectral modulation components. For the ear, a logarithmic, rather than linear, frequency scale is more applicable. Spectral modulations on such a scale can be defined in units of ripples/octave (20). The representation of a

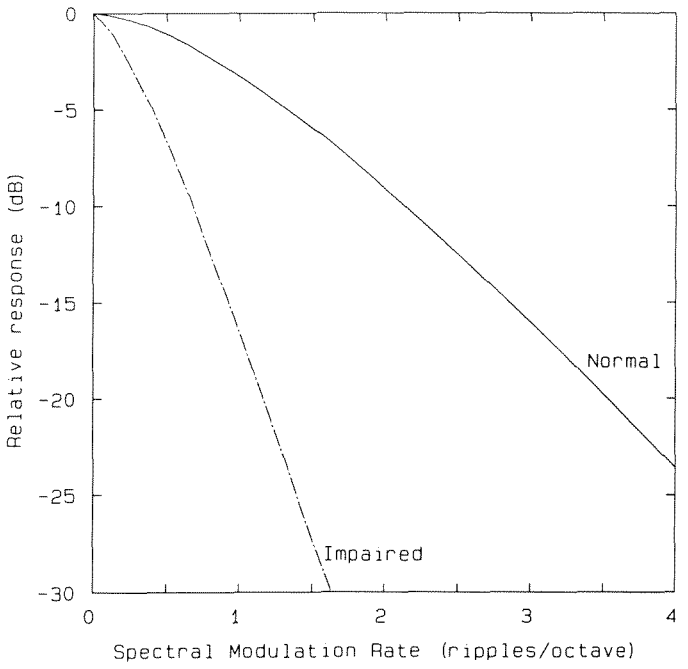

Figure 1.
Simulation of the ability of a normal and an impaired ear to resolve spectral modulation as a function of modulation rate. Solid curve: normal ear. Dash-dotted curve: impaired ear with bandwidths three times normal.

spectral shape in the auditory system is called the excitation pattern. The excitation pattern of a given sound may be defined as the magnitude of the output of the auditory filters as a function of center frequency, in response to that sound (5,6). This representation of the spectrum has the effect of attenuating spectral modulations at high rates, resulting in a smoothing of the spectrum. This is illustrated in **Figure 1**, which shows the relative response to spectral modulations of different rates at the output of a simulated array of auditory filters (5). For a normal ear (upper curve), rates of 2 ripples/octave are attenuated by about 9 dB, and rates of 4 ripples/octave are attenuated by about 23 dB. For an ear with a moderate sensorineural hearing loss, the auditory filters typically have bandwidths about three times greater than normal (9). For such an ear (lower curve), rates of 2 ripples/octave are attenuated by about 40 dB and even rates as low as 1 ripple/octave are attenuated by about 16 dB.

In principle, by prefiltering a spectrum in the spectral-modulation domain, boosting modulations at high rates, it is possible to make the excitation pattern in an impaired ear become more like the excitation pattern evoked by the unfiltered spectrum in a normal ear. This process is referred to commonly in the signal-processing literature as deconvolution, but, in order to be performed accurately, it requires an accurate knowledge of the degradation suffered by the signal in passing through the impaired ear. Even in theory it is not always possible to perform perfect deconvolution. In practice, there is a limit in the extent to which this can be done, since for high rates, above about 2 ripples/octave, the amount of boost needed becomes so large as to be impractical. Thus, while it is not possible to process signals to create a completely normal excitation pattern in an impaired ear, it is possible to make the excitation pattern more like that in a normal ear, particularly for low rates (below about 2 ripples/octave). Van Veen and Houtgast (20) have shown that ripples at rates around 2 ripples/octave are particularly important for judgments of vowel quality (timbre). Thus, preprocessing of spectra may make it possible to restore the most important aspects of the excitation pattern to a more normal form.

At present, there are no strong reasons for preferring one rationale for spectral enhancement over the other. In practice, there is a fair degree of overlap between the two approaches, especially when consideration is taken of other effects of hearing impairment, as will be addressed below. The schemes described here were intended both to suppress noise in the spectral valleys between formants and to enhance ripples in the spectrum at higher rates. However, the parameters of the processing were chosen by consideration of the effects of the processing in the spectral-modulation domain.

There are two complications to the approaches outlined above which are not addressed by the provision of simple spectral enhancement so that 'peaks' stand out more than 'valleys.'

1) *Upward Spread of Masking.* In hearing-impaired people, the auditory filters are often not symmetric on a linear frequency scale, as found in normal subjects at low to moderate sound levels (6,8). For the former, the low-frequency 'skirts' of the filters often have shallower slopes than the high-frequency skirts. Consequently, higher frequencies tend to be masked by lower frequencies more than lower frequencies are masked by higher ones. This also happens for normal subjects at high sound levels, a phenomenon known as the 'upward spread of masking.' Enhancement of signals at lower frequencies may then produce an increase of masking of higher frequencies, especially since the low-frequency skirt of the auditory filter becomes progressively shallower with increasing signal level. Asymmetric enhancement along the frequency axis may therefore be beneficial. In signal processing terms, this calls for an enhancement filter with a nonlinear phase response.

Journal of Rehabilitation Research and Development Vol. 29 No. 2 Spring 1992

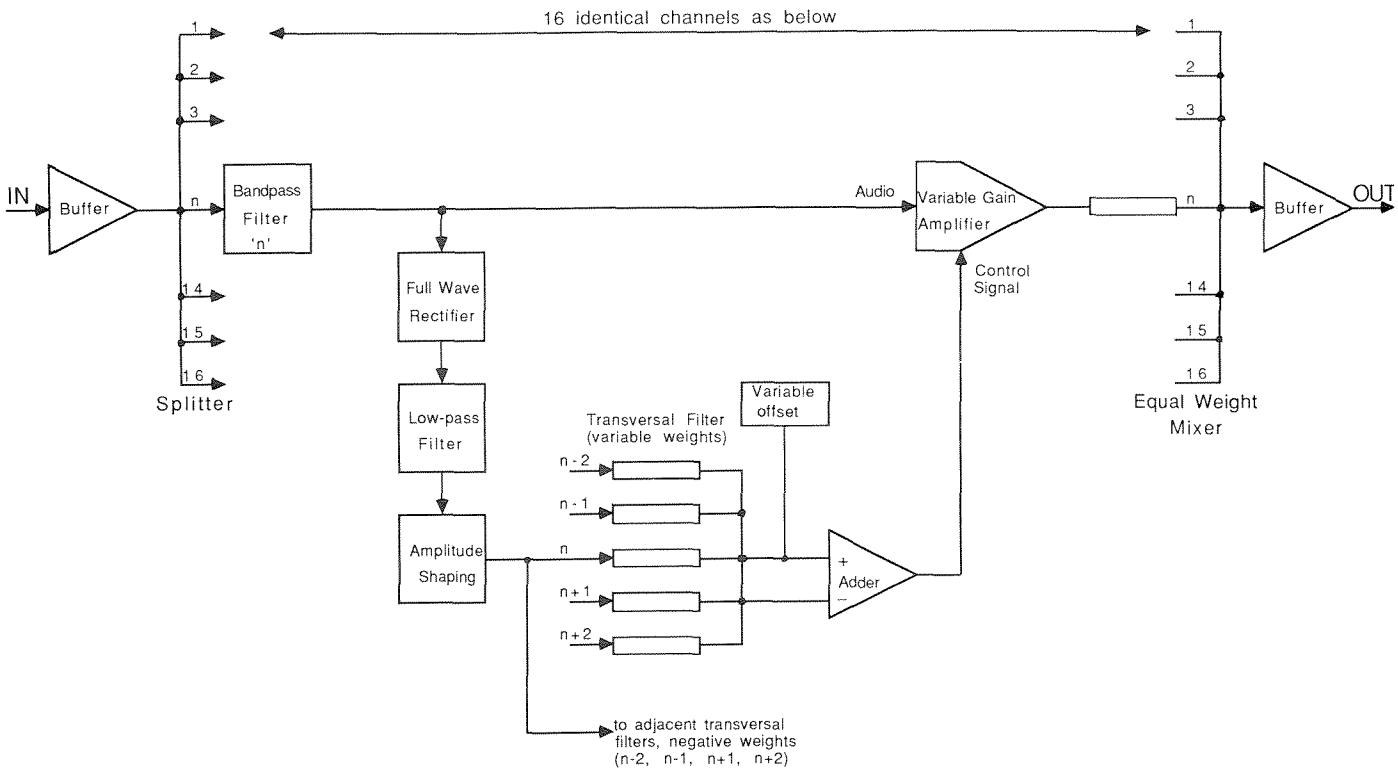

Figure 2.
A schematic diagram of the components of the system used to process speech in noise. Only one channel is shown in detail between the splitter and the mixer.

2) *Loudness Recruitment.* Subjects with sensorineural hearing loss have elevated thresholds for detecting sounds, but the level at which sounds become uncomfortably loud is often similar to that found for normal subjects. Therefore, the range of usable sound levels is reduced. The abnormally rapid growth of loudness with increasing sound level is called loudness recruitment. Enhancement schemes should therefore limit the increase in level of spectral components. To achieve a similar level of contrast enhancement to that which would occur if loudness recruitment was ignored, the level of the spectral components in the spectral valleys must undergo an extra amount of reduction.

In both experiments, enhancement schemes were chosen that took these factors into account, and an attempt was made to explore their relative importance.

METHOD

Equipment

The equipment configuration is shown schematically in **Figure 2**. The incoming signal is buffered and distributed in parallel to 16 bandpass filters. Since each of the 16 channels is similar, only one is shown in the schematic. Each bandpass filter comprises a 6th order low-pass section followed by a 6th order high-pass section. Each filter is designed using a Chebyshev 0.5-dB ripple polynomial, giving an initial rate of attenuation of 60 dB/octave, averaged over the octave beyond the −3 dB frequency. The −0.5 dB points of each bandpass filter are at 0.98 and 1.02 times the nominal center frequency. The corresponding ratios for the −3 dB points are 0.95 and 1.05. Such high-Q filters are difficult to realize simply as operational amplifiers configured as bandpass filters due to the constraints imposed on gain and slew rate. So, each filter section is implemented by the cascading of three active voltage-controlled, voltage-source (VCVS), 2-pole blocks. The center frequencies for the bandpass filters are spaced at quarter-octave intervals, from 263 Hz to 3,531 Hz, a total of 3.75 octaves. When the outputs from all filter sections are recombined, the filter phase function causes some ripple in the passband due to frequency-dependent cancellation or addition. The filters were designed so that there was a slight peak at the center frequency of each bandpass section. This arose from the addition of components from the two filters on either side, where the filter responses crossed over at around −13 dB relative to the response at the center. The overall response

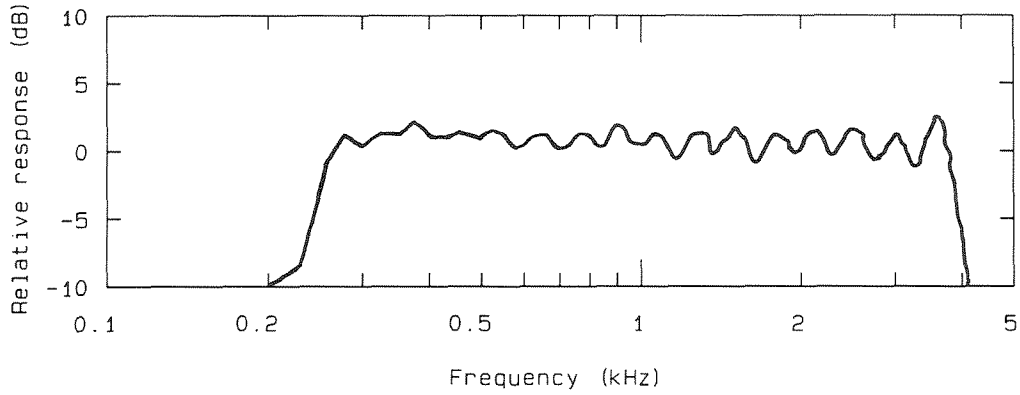

Figure 3. A plot of the system transfer characteristic on a logarithmic frequency axis.

of the combined filters was ±2 dB over the entire range of 255 to 3,870 Hz, and −3 dB at 250 and 3,900 Hz. A plot of the frequency response for the combined output of all 16 channels is shown in **Figure 3**.

The bandwidths of the filters were chosen to be reasonably close to the bandwidths of the auditory filters found in a normal ear (4,5). However, the slopes of the filters were less than those of the auditory filters, which typically have slopes of 80 to 200 dB/octave. The ear may be regarded as effectively having a very large number of filters with overlapping center frequencies. Each point on the basilar membrane within the cochlea acts like a filter tuned to a particular center frequency, and the center frequency changes continuously and progressively with position along the basilar membrane. Practical constraints limited the number of filters in our filter bank. Ideally, we would have liked to have spaced them at, say, 0.25 times the filter bandwidth. However, to make the equipment of reasonable size and cost, the filters were spaced roughly 'edge to edge.' This leads to filters equally spaced on a logarithmic frequency axis. Such spacing has two advantages. First, it ensures that the passband ripple is kept constant across frequency. Secondly, it corresponds approximately to the way that center frequency is mapped to place on the basilar membrane (6).

Each bandpass signal proceeded into a board with its own variable gain amplifier, while additionally, in a side chain, it was full-wave rectified and smoothed with two single-pole low-pass R-C filters. The −3 dB point of the first stage was 1.33 octaves below the −3 dB point of the second stage. This smoothing function attempted to mimic the temporal resolution of the ear (21,22,23). The two corner frequencies had values of 13 and 32 Hz for the lowest center frequency and they increased in quarter-octave steps, scaling with the filter center frequency, for center frequencies below 1 kHz, but were fixed at 40 and 100 Hz for frequencies above 1 kHz. The resultant smoothed signal, the activity signal, gave a measure of the envelope of the bandpass signal. The rise time of the activity signal in response to an abrupt step in amplitude, measured between the 10 percent and 90 percent points, ranged from 30 ms at the lowest center frequency to 9.5 ms for center frequencies above 1 kHz. The activity signal was amplified linearly in the first experiment, but was passed through a logarithmic amplifier in the second experiment. This optional action was performed by the amplitude-shaping network shown in **Figure 2**.

The activity signal for a given channel was weighted and combined with activity signals from up to two channels on either side of that filter. The weightings were such that the central contribution, from channel (n), was always positive, while the contributions from the side channels, (n−2), (n−1), (n+1), and (n+2), were always negative. This weighting over several channels therefore has the appearance of a transversal finite impulse response (FIR) filter. This filter will be referred to as the enhancement filter. However, instead of operating in the conventional domain of linear frequency, the filter operated in the domain of spectral modulation along the dimension of ripples/octave, mentioned earlier. This was the dimension along which we wished to enhance spectral features. The combined and weighted activity signals that later implemented gain control could optionally have an offset signal added, which could be used to modify the properties of the enhancement being performed. If the gain control signal is regarded as an AC waveform, the offset acted as a DC level shift.

The weighted activity signal was then returned to the variable gain amplifier where it modified the level of the bandpass signal. Finally, the outputs from the 16 variable-gain boards were recombined in a linear mixer and buffered either to drive headphones or to give a line-level output for recording.

Journal of Rehabilitation Research and Development Vol. 29 No. 2 Spring 1992

EXPERIMENT 1

Circuit details

The equipment outlined above was used to perform linear enhancement. The rationale behind the use of a linear processing scheme is to model the spectral smearing as a process that happens at the level of the mechanical motion of the basilar membrane. The weighted combined signal acted to vary the gain of a CA3080 transconductance amplifier. The gain of this integrated circuit, in the circuit configuration used, was linearly proportional to the applied control signal and the variable offset shown in **Figure 2**. To ensure a high signal-to-noise ratio at the output of the processing electronics, the bandpass signals were subject to time-invariant emphasis and de-emphasis as they passed their respective variable-gain board; the emphasis rose at 6 dB/octave above 500 Hz, with a corresponding inverse at the output of each board.

Computer programs were written to calculate: 1) the Fourier transform of the enhancement FIR filter weightings (in the spectral modulation domain); 2) the degree of spectral asymmetry in the enhancement filter response; and, 3) the maximum enhancement available and the balance as to how this increase in spectral contrast was achieved, either by enhancement of spectral peaks, suppression of spectral valleys, or a combination of both. If the peaks are enhanced too much, this can lead to a marked increase in the dynamic range of the speech, which can cause problems for the hearing impaired, since the dynamic range over which they can hear is usually smaller than normal: this is the constraint imposed by recruitment. Enhancement filter tap weightings were adjusted to compensate as far as possible for the spectral blurring caused by an auditory filter broadened symmetrically by a factor of 3. At the same time, the offset, mentioned earlier, was also adjusted to modify the response obtained from the third simulation program. Since only a 5-tap FIR filter was used, it was not possible to compensate exactly for the spectral blurring produced by the hypothetical broadened auditory filters (exact compensation is, in any case, not possible, for reasons discussed in the Introduction). Also, it should be remembered that, although most hearing-impaired subjects have broader than normal auditory filters, the exact shapes of the filters can vary considerably from one subject to another (9,10). Therefore, in these experiments, the compensation for spectral blurring should be regarded as broadly appropriate for the class of subjects tested, rather than being exactly suited to any individual.

The DC offset voltage added to the gain control voltage was set to limit the maximum increase in level in any individual channel to +2.5 dB. This value was chosen to enable some peak-enhancement without excessive expansion of dynamic range. All other enhancement was achieved by suppression in adjacent channels.

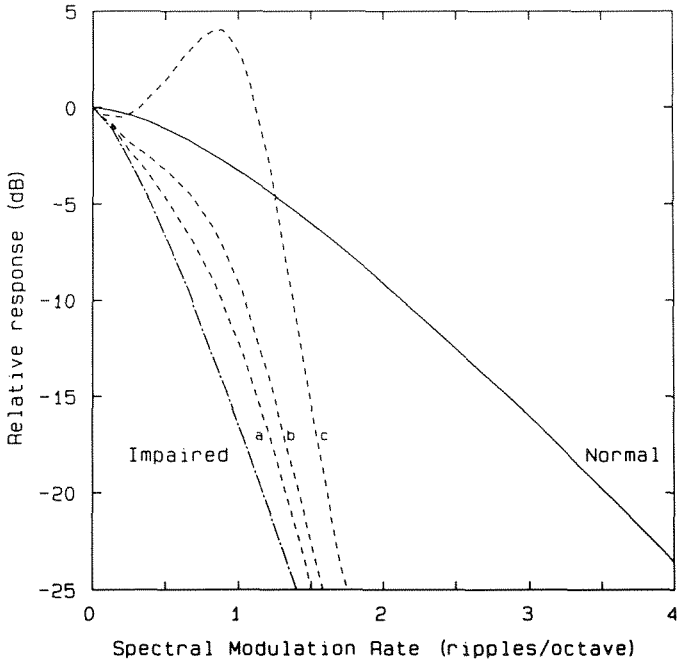

Figure 4A.

Figure 4A-C.
The effect on the transmission of spectral modulation of the processing schemes assessed in Experiment 1. The solid and dash-dotted curves show responses of simulated normal and impaired ears without processing, while the dashed curves, a, b, and c, represent the effect of the processing schemes on the 'effective' response of an impaired ear for signals at various levels relative to the peak level. Each panel represents a different processing scheme.

The relative effects of the processing schemes employed on the loss in spectral detail due to auditory filters broadened by a factor of 3 are illustrated in **Figure 4A, Figure 4B**, and **Figure 4C**. Each panel shows results for a different processing scheme. The solid curve in each of the three panels shows the relative response to spectral ripples of different rates for a simulated normal ear. The dash-dotted curve shows the corresponding response for a simulated impaired ear with auditory filters broadened by a factor of 3. Curves a, b, and c show the responses of the simulated impaired ear to processed signals, with signal level increasing from a to c. We will refer to these curves as showing the 'effective response' of the impaired ear. They show the extent to which each scheme is effective in restoring the response of the impaired ear to normal. A detailed description of each processing scheme is given below.

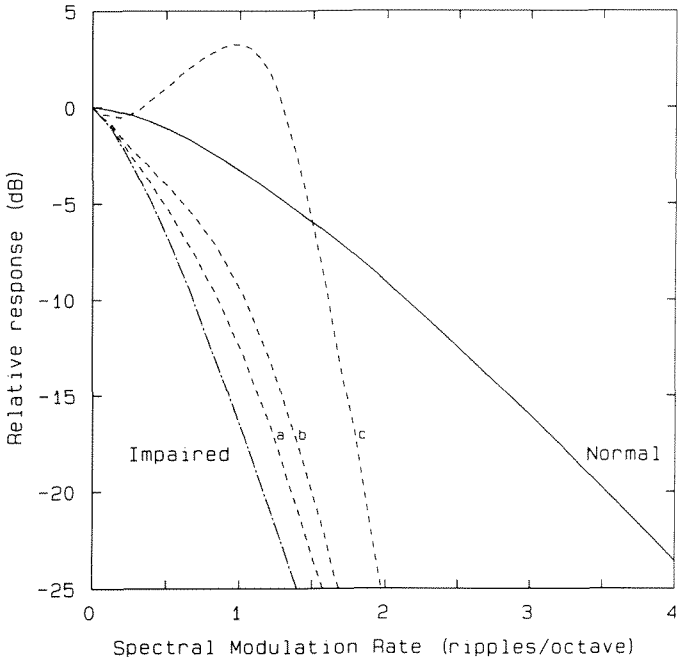

Figure 4B.

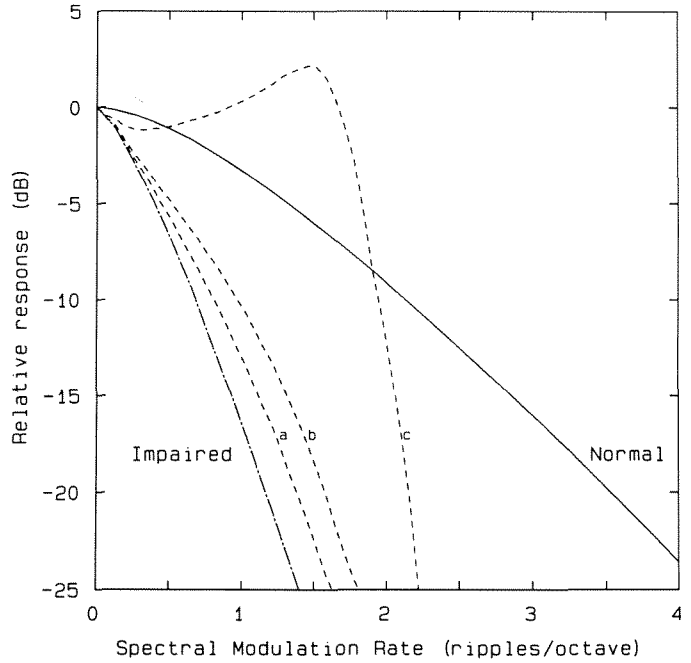

Figure 4C.

The curves a–c show the enhancement occurring when the peak amplitude of the spectrally modulated signal ranges across channels from 0.316 (curve a), to 0.5 (curve b), or to 1 (curve c) relative to the peak amplitude that a channel can carry. These ranges are equivalent to peak values of −10, −6, or 0 dB. Notice that the processing has little effect for signal levels below this range. This is to be expected from the use of weightings of a linear representation of the activity signal. Also, the variable offset, which was used to limit the maximum growth of the signal in each channel, imposed extra constraints, limiting the effectiveness of the linear weighting.

Four different processing schemes were used:

(i) The signal was passed through the filter bank and recombined, with no enhancement performed; this is the control condition, referred to as 0, 0, 0, 0, 0.

(ii) Tap weightings were in the ratio −0.5, −0.3, +1, −0.3, −0.5, for the taps (n−2), (n−1), n, (n+1), (n+2), respectively. The activity signal so derived was used to control the gain in the nth channel. This scheme attempted to compensate for blurring over the low spectral detail range of 0 to 1.25 ripples/octave (see **Figure 4A**). It can be argued that impaired subjects may have learned to process only coarse spectral features, because of long experience with their reduced frequency selectivity. Given that only a brief training period was possible with our laboratory-based system, boosting spectral modulations at higher rates might have increased the energy of the signal without adding any extra usable information. This scheme may be regarded as enhancing spectral modulations in the range where subjects already know how to make use of the spectral information.

(iii) Tap weightings were in the ratio −1, −1, +1, −0.33, 0, using the same nomenclature as in (ii) above. This gave an asymmetric filter, designed to enhance spectral modulations up to rates of 1.5 ripples/octave, similar to the range enhanced in (ii) above, but with an asymmetry (see **Figure 4B**). The asymmetry in the filter impulse response meant that the filter had different effects for positive-going and negative-going edges in the spectrum. This was intended to take into account the finding that, in hearing-impaired subjects, the upward spread of masking from low frequencies to high is generally greater than the downward spread from high frequencies to low. For a positive-going edge, the first channel after the edge (i.e., the first channel at the higher level) is enhanced; this potentially gives more downward spread of masking, but this is not usually a problem. On the other hand, for a negative-going edge, the last channel before the edge (again the one with the high level) is not enhanced, avoiding an increase in the upward spread of masking at the expense of additional suppression of the low-level channel immediately following the edge. In

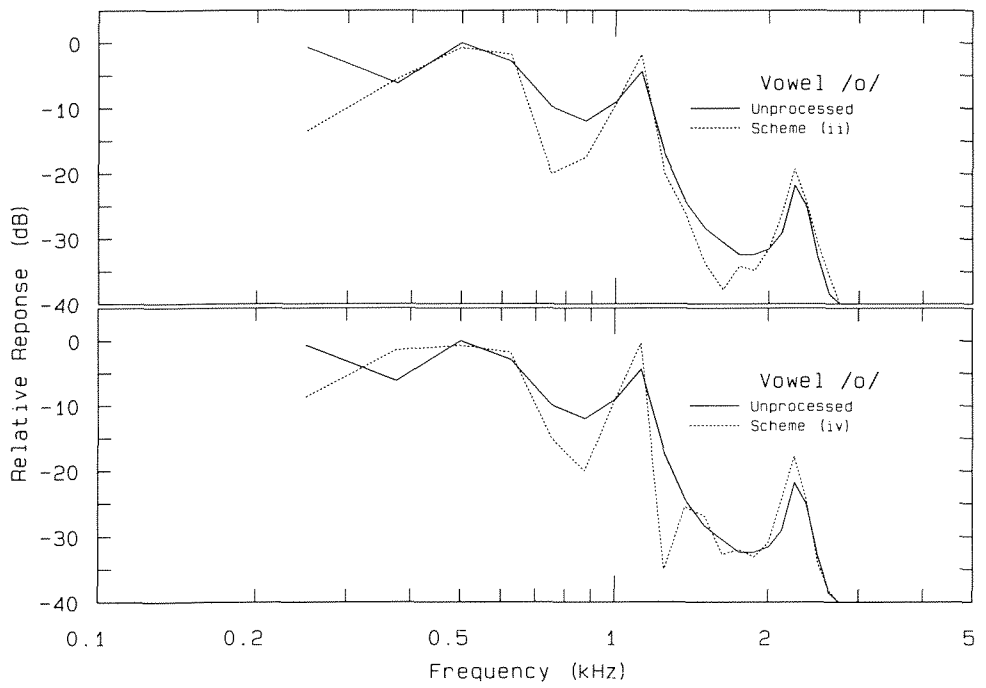

Figure 5A.

Figure 5A,B.
Spectral envelopes of the synthetic vowels /o/ (panel A) and /I/ (panel B) either unprocessed (solid line) or processed using scheme (ii) or (iv) (dotted lines). No background noise was present. The spectral envelopes were obtained by sampling the spectrum at harmonic frequencies of the 125-Hz fundamental.

practice, this low-level channel would be masked, even if no enhancement had taken place. What we were trying to do was to limit the masking spread so that low-level information carried in channels at some distance higher in frequency than the edge was potentially usable.

(iv) Tap weightings were in the ratio 0, −1, +1, −1, 0. This scheme attempted to compensate for spectral blurring over the wider range of zero to 1.9 ripples/octave (see **Figure 4C**). It was not possible to correct for blurring beyond 1.9 ripples/octave with this equipment.

The channels at either end of the frequency range which were not spanned by the full 3- or 5-tap aperture had their negative terms increased in magnitude to prevent over-enhancement (i.e., the sum of coefficients across all active taps was constant).

The effects of processing schemes (ii) and (iv) on the spectra of complex signals are illustrated in **Figure 5A** and **Figure 5B**. The figure shows the spectral envelopes of two synthetic vowels, without (solid lines) and with (dotted lines) processing. The vowels were presented in quiet at levels close to the maximum permissible. For the vowel sound /o/ (as in 'show'; **Figure 5A**), the first, second, and third formant frequencies are quite widely spaced, and both processing schemes increase the peak-to-valley ratios in the spectra. For the vowel /I/ (as in 'hid'), the second and third formants are more closely spaced, and the energy in the valleys between these two formants is reduced more by scheme (iv) than by scheme (ii), due to the greater effect of scheme (iv) at higher spectral ripple rates.

Test methods

Speech intelligibility in noise was measured using 12 sentence lists selected from the BKB lists (24). One additional sentence list was used in a warm-up session, where presentation levels and other variables could be adjusted to suit the subject. Each sentence list consists of 16 sentences with either 3 or 4 key words, making a total of 50 key words per sentence list. The sentence lists were recorded on compact cassette with a continuous background random noise whose spectrum was the same as the long-term average spectrum of the speech, averaged over all of the sentence lists. These materials were identical to those used in the control condition of Simpson, *et al.* (16). The RMS level of the background noise, relative to the RMS level of the speech was −3 dB. The background noise was continuous throughout each list. There was an interval of 10 seconds between each 2-second sentence during which the subject was asked to report what words had been heard. Subjects were encouraged to respond, even if they were uncertain. For each setting of the filter coefficients detailed above, the master cassettes were replayed from an Aiwa F660 cassette deck through the processing electronics and recorded onto compact cassettes using a Pioneer CT-F950 cassette deck. Noise reduction circuits of the Dolby type were not used in the recordings in case the noise reduc-

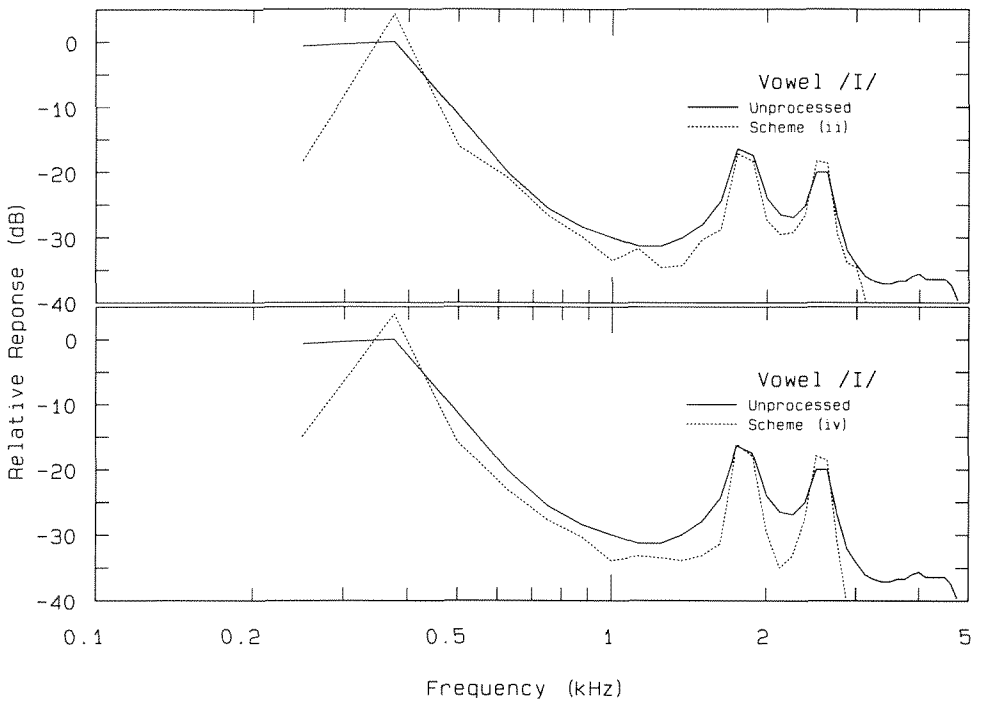

Figure 5B.

tion circuitry interfered with the processing already performed. Although this does imply an increased level of 'hiss' in the replay, the tape noise was mainly high frequency, that is, in and above the upper two octaves of interest (25), and it was not audible to the hearing-impaired subjects used in these experiments.

Five subjects (JS, AS, GW, JW, IH) were tested at two different presentation levels. Their audiograms and highest comfortable levels (HCLs) for both ears are given in **Table 1**. All subjects had been diagnosed as having hearing loss of cochlear origin. They were tested wearing their normal hearing aid(s). Three subjects used single (monaural) aids which were linear [i.e., which had no automatic gain control (AGC) circuits]. These aids were worn in the 'better' ear, which was the left ear for all subjects. One subject (JS) used two linear aids, and one subject (JW) used two Qualitone Custom Genesis aids which have AGC. Subjects were tested in a double-walled sound attenuating room whose walls were lined with 10cm-thick foam. The subject was seated 1.3 m from a loudspeaker (Monitor Audio MA4), facing toward the loudspeaker. Each subject set the volume of his/her aid(s) to that which would be used in a domestic listening environment; this volume setting was kept constant throughout all of the testing for that subject. The background noise was presented at two different levels: 44 and 64 dBA. Although the lower level is rather quiet, all subjects reported that they could hear the background noise. At this stage, the concern was that the subjects may have set their aids to listening levels where significant amounts of distortion would be generated; the lower levels ensured that this was not the case. The peak speech levels were approximately 12–14 dB higher than this. As reported later in this paper, we later tested subjects without their aids to eliminate the problem of distortion.

The different processing schemes were tested in a different order for each subject. Each subject was tested using all 12 sentence lists. In a given test session, each of the four processing schemes was tested with one sentence list. The processing schemes were then repeated in the same order using a second block of four sentence lists, and so on for a third. Thus, for each subject, three lists were used for each processing scheme.

Results of Experiment 1

The mean scores (number of key words correct out of 150) for each subject, each level, and each processing scheme are presented in **Table 2**. It is clear that, while the scores differed across subjects and levels, the scores for the different processing schemes were similar for each subject and level. An analysis of variance (ANOVA) was conducted with factors of subject, level, and processing scheme. There were highly significant differences between subjects [$F(4,80)=42.7$, $p<0.001$], and between levels [$F(1,80)=74.1$, $p<0.001$], performance being worse at the lower level. There was no significant difference between processing schemes.

Journal of Rehabilitation Research and Development Vol. 29 No. 2 Spring 1992

Table 1.
Absolute thresholds and *Highest Comfortable Levels (HCLs)* in dB HL of the hearing-impaired subjects.

Subject	Ear	Frequency in kHz .125	.250	.500	1.00	2.00	4.00	8.00
JS	L	65	70	55	45	35	50	60
		>70	*>90*	*110*	*105*	*95*	*100*	*>100*
	R	30	30	35	45	40	40	50
		>70	*>90*	*105*	*110*	*105*	*100*	*>100*
AS	L	30	25	30	60	60	75	80
		>70	*>90*	*110*	*100*	*95*	*110*	*>100*
	R	15	10	10	60	55	65	80
		>70	*>90*	*110*	*110*	*115*	*>120*	*>100*
GW	L	40	45	50	50	50	55	55
		>70	*>90*	*100*	*100*	*100*	*110*	*>100*
	R	45	50	50	55	50	55	55
		>70	*>90*	*95*	*100*	*95*	*100*	*>100*
JW	L	50	60	55	55	60	65	75
		>70	*85*	*90*	*95*	*100*	*110*	*>100*
	R	50	50	50	55	60	70	70
		>70	*85*	*90*	*95*	*100*	*110*	*100*
IH	L	35	35	35	45	45	65	85
		>70	*>90*	*100*	*100*	*90*	*90*	*100*
	R	60	65	50	55	70	90	100
		>70	*>90*	*100*	*95*	*95*	*110*	*>100*
JC	L	30	35	30	35	40	55	55
		>70	*>90*	*105*	*100*	*105*	*>120*	*100*
	R	25	30	35	35	50	65	60
		>70	*90*	*100*	*100*	*110*	*115*	*100*
CC	L	25	30	20	35	55	95	80
		>70	*80*	*90*	*95*	*95*	*120*	*>100*
	R	30	35	30	40	60	85	95
		>70	*85*	*95*	*100*	*95*	*115*	*>100*
AD	L	20	20	30	55	50	55	90
		65	*70*	*75*	*85*	*75*	*75*	*100*
	R	30	25	25	50	55	60	85
		65	*70*	*75*	*85*	*75*	*75*	*95*
JL	L	5	5	20	20	35	55	80
		85	*85*	*85*	*85*	*80*	*80*	*90*
	R	0	0	10	15	30	65	80
		85	*80*	*85*	*85*	*85*	*85*	*90*

HCLs are printed in italics and slightly offset for ease of reading.

Table 1. *(Continued)*
Absolute thresholds and *Highest Comfortable Levels (HCLs)* in dB HL of the hearing-impaired subjects.

Subject	Ear	Frequency in kHz .125	.250	.500	1.00	2.00	4.00	8.00
BP	L	35	35	40	55	70	85	>100
		65	*70*	*70*	*75*	*95*	*>100*	*>100*
	R	30	35	45	55	65	75	>100
		70	*70*	*75*	*75*	*85*	*>100*	*>100*
ED	L	15	25	25	35	35	30	65
		60	*65*	*70*	*70*	*70*	*75*	*80*
	R	40	40	40	45	50	70	75
		65	*65*	*70*	*75*	*70*	*80*	*90*
LM	L	30	25	35	40	50	50	65
		65	*70*	*70*	*75*	*75*	*80*	*90*
	R	30	30	35	40	45	40	60
		70	*70*	*75*	*75*	*75*	*80*	*90*

Discussion of Experiment 1

The results are disappointing in that there was no significant effect of any of the processing schemes on speech intelligibility in noise. One possible reason is illustrated in **Figure 4**. Spectral enhancement only occurred in a given spectral region when the level in the channel in that region approached the maximum permissible level for that channel. Spectral enhancement was negligible when the level was more than 10 dB below the maximum channel level. Aware of this, we had increased the weighting of each channel activity signal by 1.5 dB for each additional quarter octave above 500 Hz: this was an attempt to match the generation of the activity signals to the average spectrum of speech. However, even with this weighting of the activ-

Table 2.
Results of Experiment 1. Intelligibility scores for each subject and each processing condition.

Processing Condition						Subject Scores (max 150)						Mean Score %
Low Presentation Level						**JS**	**AS**	**GW**	**JW**	**IH**		
(i)	0,	0,	0,	0,	0:	116	56	80	66	87	:	54.0
(ii)	−0.5,	−0.3,	+1,	−0.3,	−0.5:	114	55	82	90	86	:	57.0
(iii)	−1,	−1,	+1,	−0.33,	0:	104	50	78	55	90	:	50.2
(iv)	0,	−1,	+1,	−1,	0:	127	65	79	72	100	:	59.0
High Presentation Level						**JS**	**AS**	**GW**	**JW**	**IH**		
(i)	0,	0,	0,	0,	0:	124	74	110	106	114	:	70.4
(ii)	−0.5,	−0.3,	+1,	−0.3,	−0.5:	136	86	119	100	111	:	73.6
(iii)	−1,	−1,	+1,	−0.33,	0:	128	81	117	112	105	:	72.4
(iv)	0,	−1,	+1,	−1,	0:	125	62	113	101	103	:	67.2

Weightings used in the enhancement filter for each condition are indicated on the left.

Journal of Rehabilitation Research and Development Vol. 29 No. 2 Spring 1992

ity signal, only relatively high-level spectral features were enhanced. One might argue that the processing schemes were enhancing spectral features that would have already been clearly audible, while weaker spectral features, which would have been harder to hear, were not enhanced.

The limited dynamic range over which the enhancement occurred was a consequence of the use of a linear activity signal. In the next experiment, a logarithmic transform of the activity signal was used to increase the dynamic range over which enhancement occurred.

EXPERIMENT 2

The equipment was modified to include a stage that passed each channel activity signal through a logarithmic amplifier operating over a dynamic range in excess of 60 dB. The logarithmic amplifier was placed after the 2-pole R-C low-pass filter that had smoothed the envelope of the signal in that channel. The same method of implementing a transversal FIR filter (operating in the domain of ripples/octave) was used as in Experiment 1. The gain control signal at the output of the FIR filter was 'anti-logged' (exponentiated), using an MC3340 integrated circuit. In the semiconductor data books this is known as a voltage-controlled amplifier, and its transfer function, shown in **Figure 6**, is distinctly nonlinear compared to the CA3080 used in Experiment 1. It offers a gain control range of +13 to −80 dB, where gain is approximately proportional to the exponential of the control voltage (with due consideration being taken for polarity inversion and offset of the control signal). However, the deviation from the ideal anti-logging characteristic is such that, for enhancement (positive control voltage), the exponent is less than unity (i.e., compressive), and for suppression (negative control voltage), it is greater than unity (i.e., expansive). This transfer characteristic is therefore suitable for preventing large increases in level in any channel that may arise as a result of the processing. The increase in gain with increasing control voltage rolls off to a soft limit, with no step change in signal level or distortion if this limit is exceeded. Therefore, no artifacts are heard as this limit is reached. However, there is some feedthrough of the control voltage, but this is totally low frequency (below 200 Hz), compared to the signal being controlled, and is easy to remove by filtering. Throughout this experiment, the maximum increase in level in any channel was set to +6 dB. This figure was chosen as a compromise to offer a degree of peak enhancement, without leading to excessive increases in the dynamic range of speech.

The variable gain amplifier board was modified so that the emphasis/de-emphasis circuits used in Experiment 1

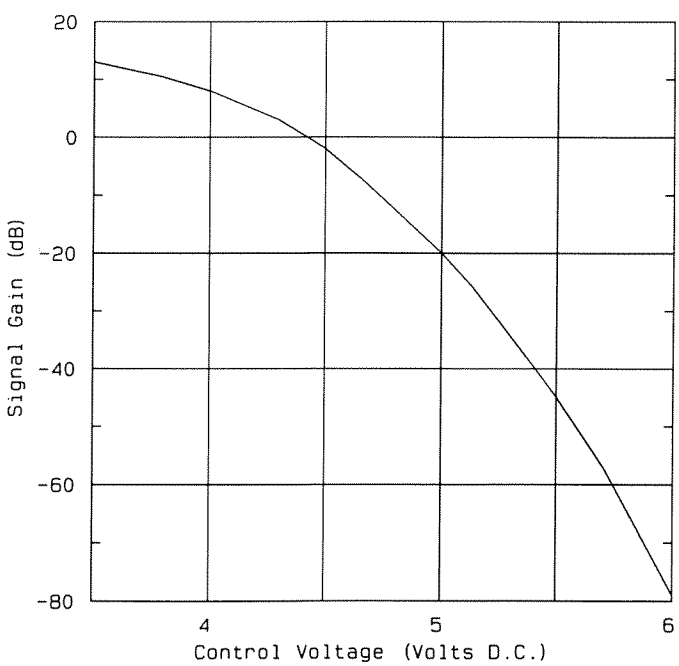

Figure 6.
The transfer characteristic of the MC3340 voltage-controlled amplifier used as the variable gain amplifier in Experiment 2.

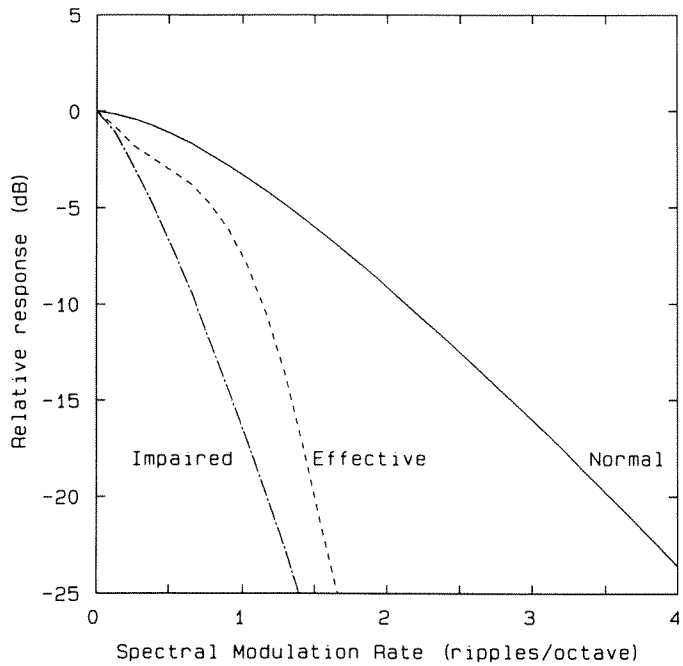

Figure 7A.

Figure 7A-C.
As **Figure 4**, but showing the effect of the processing schemes assessed in Experiment 2. For these schemes, the response is almost independent of signal level.

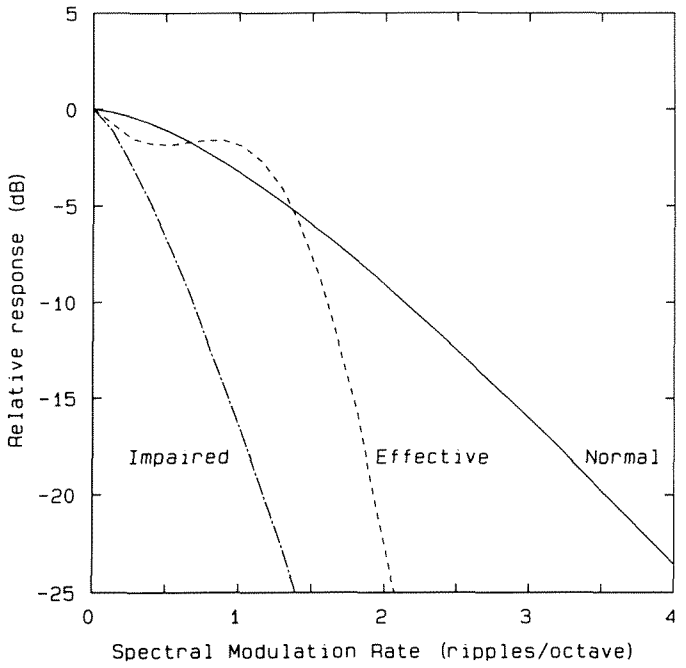

Figure 7B.

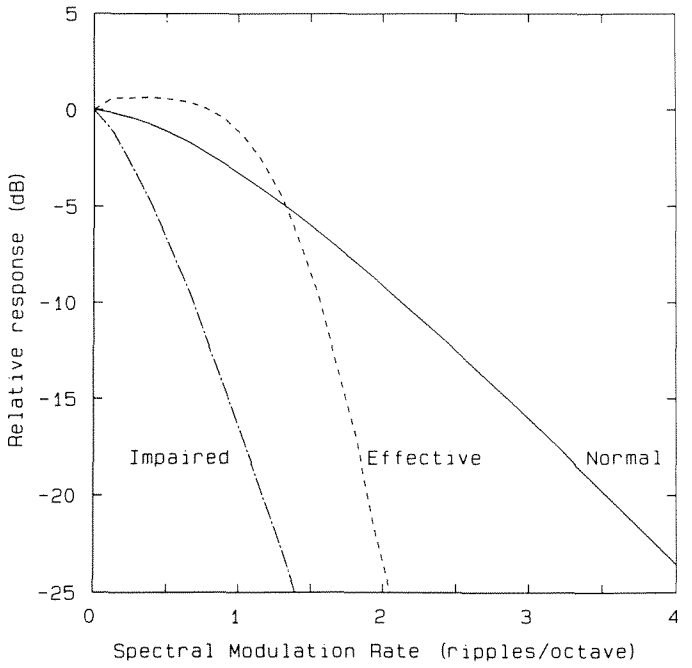

Figure 7C.

were removed; these circuits were no longer essential since the board had been rewired to handle larger signals.

Using rewritten simulation programs, as developed for Experiment 1, five different enhancement schemes were selected for assessment. Their effects on the resolution of spectral ripples at different rates are shown in **Figure 7A, Figure 7B,** and **Figure 7C**. As in **Figure 4**, the solid curve shows the relative response to spectral ripples of different rates for a simulated normal ear. The dash-dotted curve shows the corresponding response for a simulated impaired ear with auditory filters broadened by a factor of 3. The short-dashed curves show the response of the impaired ear to signals preprocessed by the processing schemes (ii) to (iv), described below. Note that, in contrast to the schemes used in Experiment 1, the enhancement of spectral resolution produced by these schemes is almost independent of overall level.

The enhancement schemes used were:

(i) Signal separated by the filter bank and then recombined with no enhancement. This is the control condition (identical to condition (i) of Experiment 1).

(ii) Tap weightings in the ratio −0.4, −0.1, +1, −0.1, −0.4, using the same nomenclature as for Experiment 1. This symmetric filter attempts to restore spectral resolution over the range 0 to 1.0 ripples/octave, and is similar to condition (ii) of Experiment 1. The effective response is shown in **Figure 7A**. This filter does not provide full correction for any spectral modulation rate, but this undercorrection helps to prevent an effect sometimes noted in Experiment 1 where some individual speech sounds seemed too prominent.

(iii) Tap weightings of 0, −0.5, +1, −0.5, 0. This symmetric filter attempts to restore spectral resolution over the range 0 to 1.5 ripples/octave, and is similar to condition (iv) of Experiment 1. The effective response is shown in **Figure 7B**.

(iv) Tap weightings of 0, −0.75, +1, −0.25, 0. This represents an asymmetric filter, with correction attempted over the range 0 to 1.6 ripples/octave. The effective response is shown in **Figure 7C**.

(v) Tap weightings of 0, −0.25, +1, −0.75, 0. This filter has an impulse response the mirror image of that in (iv) above, but its effective response is the same as that in (iv).

It was anticipated that filters (iv) and (v) would have different effects on the upward spread of masking. The filter in (iv) should *decrease* the spread, while the filter in (v) should *increase* the spread, relative to the symmetric filter used in condition (iii) of this experiment. The rationale behind this is similar to that given earlier in the description of condition (iii) of Experiment 1.

The sum over all taps was always zero for all enhancement schemes. The channels at either end of the frequency range which were not spanned by the full 3- or 5-tap aperture had their enhancement terms decreased in magnitude by an amount equal to that of the 'missing' tap. Con-

Journal of Rehabilitation Research and Development Vol. 29 No. 2 Spring 1992

sequently, for a flat spectrum input to the system, a flat spectrum was present at the output, independent of the input signal level.

Test methods

Using the same source material as in Experiment 1, 11 sentence lists in noise (speech/noise ratio +3 dB) were replayed from compact cassette, processed by the electronics and recorded using a Sony DTC-1000ES digital audio tape (R-DAT) recorder, which uses 16-bit linear coding at a 48-kHz sampling rate. One list, list 1, was used for practice, and ten (3, 5, 6, 8, 9, 10, 14, 18, 19, and 21) for testing proper. Ten subjects (AS, IH, JC, JS, GW, AD, JL, BP, ED, LM), took part in the tests; their audiograms and HCLs are given in **Table 1**. They were tested without using their own hearing aids. All subjects except JS normally used linear hearing aids; JS had aids with protective compression limiting, which would not have been activated by the stimuli used. In order to compensate for the lack of aids, which usually give a high-frequency emphasis, the off-tape signals passed through a spectrum-shaping network that rolled off at 12 dB/octave below 200 Hz, was 'flat' from 200 to 400 Hz, and rose smoothly to +2 dB at 600 Hz and +15 dB at 4 kHz. This form of spectral shaping is similar to that commonly used in commercial hearing aids.*

Subjects were allowed to adjust the level of the replayed noise to that which they found comfortable for everyday conversation in a domestic environment. The level of the background noise was set to 70, 67, 66, 65, 61, 59, 63, 67, 64, and 62 dBA, for AS, IH, JC, JS, GW, AD, JL, BP, ED, and LM, respectively. The sound levels were measured with a meter close to their ears, and later checked by measuring the sound field where their head had been. The five processing conditions were evaluated in an order that was counterbalanced across subjects using a Latin square design (26): with ten subjects it was possible to obtain a score from two subjects at each position of the Latin square. Two consecutive lists were used for each condition, giving a maximum score of 100 correct key words.

Results of Experiment 2

Table 3 shows individual and mean scores for the five processing schemes. An ANOVA showed that there were significant differences between subjects [$F(9,36)=79.7$, $p<0.001$], and between processing schemes [$F(4,36)=6.42$, $p=0.007$]. The analysis package used (GENSTAT) gave estimates of the standard errors of the differences between pairs of mean scores for the different processing schemes. These standard errors were used to assess the significance of the differences between means (via *t*-tests with degrees of freedom corresponding to the residual variance in the ANOVA). Schemes (iii) and (v) did not differ from each other, but both gave significantly lower scores ($p<0.05$) than schemes (i) and (ii). Scheme (iv) did not differ significantly from any of the other schemes. In other words, the processing never improved performance, and sometimes degraded it.

Discussion of Experiment 2

Again, the results do not show that any improvement has been achieved by the processing. Indeed, some of the processing conditions produced a significant worsening of performance. This is both disappointing and somewhat surprising, since the processing used in Experiment 2 was rather similar in principle to the digital processing used by Simpson, *et al.* (16), and they found small but consistent improvements in intelligibility for processed speech. Several differences can be identified between our processing and theirs:

1. We used only 16 filters on a logarithmic frequency scale, whereas they effectively used 128 on a linear frequency scale; however, they did reduce this resolution (especially at high frequencies) by converting this fine-grained spectral representation to an excitation pattern based on auditory filtering. It is possible that, in our system, using a smaller number of filters, a formant in speech could fall on the boundary between two filters, where the system response is slightly lower (**Figure 3**), and would not be enhanced as effectively as a formant lying at the center of a filter. Their larger number of filters ensured that there was always a filter centered close to the frequency of any formant. In addition, it enabled them to achieve a smoother overall frequency response.
2. Their processing preserved the short-term phase spectrum of the speech, whereas ours produced some distortion of the phase spectrum. The role of the phase spectrum in speech intelligibility is somewhat unclear, but there is evidence that it can carry speech information (27).
3. Our processing limited the increase in level in any channel produced by the enhancement to 6 dB,

* In pilot trials for this experiment, separated in time from the main trials, to minimize learning effects, we had run five subjects (AS, IH, JC, JS, and GW), with and without their hearing aids (but with frequency compensation in the latter case). Scores for identifying sentences in noise were significantly higher without their hearing aids (*t*-test, $p<0.01$). This may have been due either to distortion or to inadequate frequency responses in their own hearing aids.

Table 3.
Results of Experiment 2. Intelligibility scores for each subject and each processing condition.

Processing Condition	Subject Scores (max 100)											Mean Score %
	JS	AS	GW	IH	JC	AD	JL	BP	ED	LM		
(i)	78	75	82	71	88	65	87	57	55	58	:	71.6
(ii)	67	87	82	68	81	59	83	61	56	60	:	70.4
(iii)	68	62	79	74	74	36	86	32	58	49	:	61.8
(iv)	65	82	71	65	84	58	84	41	61	46	:	65.7
(v)	67	74	74	76	83	43	49	39	46	42	:	63.3

whereas the processing of Simpson, *et al.* did not have such a limit.

It remains unclear whether these factors are sufficient to account for the difference between the two studies.

SUBJECTIVE IMPRESSIONS

Changes resulting from the processing

The subjective impression of the processing, to a normal ear, is that the speech appears to stand out more clearly against the background noise. However, some words appear to have lost some of the dynamics of their attack (i.e., hard attacks are softened, and long vowels appear slightly modulated). This appears to be an effect of the processing itself, rather than of our use of a background noise. This noise often caused the low amplitude initial speech sounds to be inaudible (to a normal ear) even in the unprocessed condition. The hearing-impaired subjects also often reported that the processed speech appeared to stand out more clearly from the background noise. This raises the possibility that there may be benefits of the processing that are not revealed by the intelligibility tests. It is possible that, when listening conditions are difficult, subjects may expend more effort in order to maintain performance (28). This will tend to reduce differences in intelligibility produced by the processing. Indeed, it is a common finding that processing of speech in noise can improve subjective clarity without altering measured intelligibility (29).

To test these ideas, two tests were performed where the hearing-impaired subjects made pair-wise subjective comparisons between sentences processed by conditions (i), (ii), (iii), and (iv) of Experiment 2, rating them for sound quality in one set of tests and intelligibility in the other set. Condition (v) was left out since this had identical processing to condition (iv), except for the reversed phase response of the enhancement filter. For each of the six possible comparisons of processing condition, ten pairs of sentences were compared. Within each pair the same sentence was used, the sentences differing only in the way they were processed. Sixty different sentences were used, taken from BKB lists 3, 5, 6, and 8. They were edited so that each sentence in each condition was preceded by 1.5 seconds of its masking noise and followed by 0.5 seconds of the same noise. There was no silent interval between the noises for the two sentences in a pair. Thus, the two sentences were separated by 2 seconds of masking noise. Following the end of the noise for the second sentence, there were 5 seconds of silence during which the subject indicated which sentence had the higher quality or intelligibility. Five of the sentence pairs were presented as condition X followed by condition Y, while the other five were presented as condition Y followed by condition X. The order of presentation of the sentence pairs was randomized for both comparison of processing condition and order of presentation within each individual test. All editing was done digitally using a Masscomp 5400 computer system. Final stimuli were recorded on digital audio tape (Sony DTC 1000ES).

The same experimental conditions were used as in Experiment 2, with ten subjects listening without their hearing aids and at the same presentation level, again with high-frequency pre-emphasis applied to compensate for the lack of a hearing aid. In the first test, subjects were asked to indicate which sentence in each pair had the higher sound quality in terms of pleasantness and freedom from distortion and noise. In the second test, they were asked to indicate which sentence in each pair they felt was more intelligible.

Journal of Rehabilitation Research and Development Vol. 29 No. 2 Spring 1992

Table 4.
Ranking of the processing conditions based on paired comparisons.

Preference by:	QUALITY				INTELLIGIBILITY			
Condition:	i	ii	iii	iv	i	ii	iii	v
Subject								
JS	4 (10)	2 (17)	3 (15)	1 (18)	4 (8)	2 (17)	3 (16)	1 (19)
AS	4 (10)	2 (18)	3 (11)	1 (21)	1 (18)	2 (16)	3 (14)	4 (12)
GW	4 (6)	3 (14)	2 (19)	1 (21)	4 (8)	1 (20)	2 (18)	3 (14)
JW	3 (13)	1 (18)	4 (11)	1 (18)	3 (15)	2 (18)	4 (8)	1 (19)
IH	2=(15)	4 (11)	1 (19)	2=(15)	4 (11)	1=(18)	1=(18)	3 (13)
JC	1 (17)	2 (15)	3=(14)	3=(14)	4 (11)	3 (14)	1 (19)	2 (16)
CC	1=(17)	3 (15)	4 (11)	1=(17)	1 (19)	2 (15)	3=(13)	3=(13)
AD	4 (10)	1 (19)	2 (18)	3 (13)	4 (11)	3 (14)	2 (15)	1 (20)
BP	1 (23)	2 (18)	4 (9)	3 (10)	1 (20)	4 (11)	3 (13)	2 (16)
LM	1 (19)	3 (14)	2 (18)	4 (9)	3 (14)	4 (13)	2 (16)	1 (17)
Totals (max 300)								
	4 (140)	1 (159)	3 (145)	2 (156)	4 (135)	2 (156)	3 (150)	1 (159)

Judgments were based either on quality (left half) or intelligibility (right half).

The results of the rank orderings are shown in **Table 4** as a position from 1 (top) to 4 (bottom) followed by a score in brackets. The position is followed by an '=' sign if the score ties with another processing condition. The scores indicate the number of times that a subject preferred that condition over any of the other three conditions. For a comparison of 60 sentences and four processing conditions, each processing condition appeared 30 times; hence, this is the maximum score possible for an individual subject.

For both quality and intelligibility, the control condition received the lowest overall ratings. The conditions receiving the highest overall scores for quality and intelligibility were (ii) and (iv). Five subjects out of ten rated condition (iv) highest for quality, and four subjects out of ten rated it highest for intelligibility. Recall that this was the condition that was intended to be most effective in combating the effects of the upward spread of masking. However, considerable individual differences were also apparent. Four out of ten subjects in the quality rating, and two out of ten in the intelligibility rating showed a preference for no processing. Because of these large individual differences, the overall effect of processing condition was not significant (chi-squared test).

It is noteworthy that the judgments of quality and intelligibility were sometimes quite different. For example, subject LM rated condition (iv) lowest for quality, but highest for intelligibility; subject JC rated condition (i) highest for quality, but lowest for intelligibility.

DISCUSSION AND CONCLUSIONS

Overall, the results of these experiments must be considered disappointing. None of the processing schemes evaluated in Experiments 1 and 2 produced a significant improvement in the intelligibility of speech in noise. However, speech processed using schemes (ii) and (iv) of Experiment 2 was judged to be higher in both quality and intelligibility than unprocessed speech [condition (i)] by some subjects. The discrepancy between the intelligibility measurements and the subjective evaluations may occur because the information content of the speech in noise is more important for the former than for the latter. The processing schemes may have the effect of reducing background noise and improving subjective quality, while at the same time they degrade certain speech cues.

Substantial individual differences were found in both the objective intelligibility tests and the subjective evaluations. Further work is needed to establish whether there is a sub-class of those with sensorineural hearing loss who might benefit from processing of the type described here,

and to clarify what the characteristics of this population are.

Our processing method, based on analog electronics, had the advantage over earlier processing schemes, such as that used by Simpson, *et al.* (16), that it ran in real time. This made it much easier to generate test materials. Also, the system had the potential advantage of being possible to implement in a wearable form. However, our system was still laboratory-based. Thus, the subjects did not have an opportunity to practice in the different processing conditions, or to experience the processing under everyday conditions. It seems quite plausible that the full benefit of a given processing scheme will only be realized when the subjects have had extensive experience using it. A processing scheme may make available cues that were not previously available to a hearing-impaired subject, and the subject may need to learn to make use of those cues. In addition, some speech cues may be transformed into an unfamiliar form that needs to be relearned. Substantial learning effects are sometimes found even for very simple processing, such as linear amplification with frequency response shaping. Learning effects for complex processing schemes, such as those used with cochlear implants, can extend over months or even years. Our failure to demonstrate advantages of the spectral enhancement may have been due as much to lack of familiarity and practice as to limitations in the processing itself.

This creates a dilemma. It takes considerable time, effort, and money to produce a wearable aid incorporating the type of processing described here. It would hardly seem worthwhile to produce a wearable aid without promising results from laboratory (nonwearable) prototypes. Yet, nonwearable prototypes may be unsuitable for demonstrating the benefits of the processing, especially when the effects are subtle.

ACKNOWLEDGMENTS

We thank Tom Baer, Brian Glasberg, Michael Shailer and two anonymous reviewers for their helpful comments on an earlier version of this paper. This work was supported by the Medical Research Council (UK).

REFERENCES

1. Plomp R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. J Acoust Soc Am 1978;63:533-49.
2. Plomp R. A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. J Speech Hear Res 1987; 29:146-54.
3. Fletcher H. Auditory patterns. Rev Mod Physics 1940;12:47-65.
4. Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. Hear Res 1990;47:103-38.
5. Moore BCJ, Glasberg BR. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. J Acoust Soc Am 1983;74:750-53.
6. Moore BCJ, Glasberg BR. Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. Hear Res 1987;28:209-25.
7. Patterson RD. Auditory filter shapes derived with noise stimuli. J Acoust Soc Am 1976;59:640-54.
8. Patterson RD, Moore BCJ. Auditory filters and excitation patterns as representations of frequency resolution. In: Moore BCJ editor. Frequency selectivity in hearing. London: Academic Press, 1986: 123-77.
9. Glasberg BR, Moore BCJ. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. J Acoust Soc Am 1986;79:1020-33.
10. Tyler RS. Frequency resolution in hearing-impaired listeners. In: Moore BCJ editor. Frequency selectivity in hearing. London: Academic Press, 1986: 309-63.
11. Leek MR, Dorman MF, Summerfield AQ. Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners. J Acoust Soc Am 1987;81:148-54.
12. Dreschler WA, Plomp R. Relations between psychophysical data and speech perception for hearing impaired subjects. J Acoust Soc Am 1980;68:1608-15.
13. Dreschler WA, Plomp R. Relations between psychophysical data and speech perception for hearing impaired subjects. II. J Acoust Soc Am 1985;78:1261-70.
14. Glasberg BR, Moore BCJ. Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. Scand Audiol Suppl 1989;32:1-25.
15. Boers PM. Formant enhancement of speech for listeners with sensorineural hearing loss. IPO Annual Progress Report, 1980;15:21-8.
16. Simpson AM, Moore BCJ, Glasberg BR. Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners. In: Hearing impairment and signal-processing hearing aids. Smoorenburg GF editor. Acta Otolaryngol Suppl 1990;469:101-7.
17. Summerfield AQ, Foster J, Tyler R, Bailey PJ. Influences of formant narrowing and auditory frequency selectivity on identification of place of articulation in stop consonants. Speech Commun 1985;4:213-29.
18. Rosen S, Fourcin AJ. Frequency selectivity and the perception of speech. In: Moore BCJ editor. Frequency selectivity in hearing. London: Academic Press, 1986: 373-487.
19. Young ED, Sachs MB. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J Acoust Soc Am 1979;66:1381-403.
20. Van Veen TM, Houtgast T. Spectral sharpness and vowel dissimilarity. J Acoust Soc Am 1983;77:628-34.
21. Moore BCJ, Glasberg BR, Plack CJ, Biswas AK. The shape of the ear's temporal window. J Acoust Soc Am 1988;83:1102-16.
22. Plack CJ, Moore BCJ. Temporal window shape as a function of frequency and level. J Acoust Soc Am 1990;87:2178-87.
23. Viemeister NF. Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am 1979;49:78-82.

24. Bench J, Bamford J. Speech-hearing tests and the spoken language of hearing-impaired children. London: Academic Press, 1979.
25. Berkovitz R, Gundry K. Dolby B-type noise reduction system. Audio, September 1973.
26. Snedecor GW, Cochran WG. Statistical methods. 6th ed. Ames, IA: Iowa State University Press, 1967.
27. Schroeder MR, Strube HW. Flat-spectrum speech. J Acoust Soc Am 1986;79:1580-83.
28. Gatehouse S, Gordon J. Response times to speech stimuli as measures of benefit from amplification. Brit J Aud 1990;24:63-8.
29. Lim JS. Speech enhancement. Englewood Cliffs, NJ: Prentice Hall, 1983.