

Published in final edited form as:

J Acoust Soc Am. 2008 February ; 123(2): 1140–1153. doi:10.1121/1.2824018.

Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech

Kathryn Hopkins^{a)}, Brian C. J. Moore, and Michael A. Stone

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kingdom

Abstract

Speech reception thresholds (SRTs) were measured with a competing talker background for signals processed to contain variable amounts of temporal fine structure (TFS) information, using nine normal-hearing and nine hearing-impaired subjects. Signals (speech and background talker) were bandpass filtered into channels. Channel signals for channel numbers above a “cut-off channel” (CO) were vocoded to remove TFS information, while channel signals for channel numbers of CO and below were left unprocessed. Signals from all channels were combined. As a group, hearing-impaired subjects benefited less than normal-hearing subjects from the additional TFS information that was available as CO increased. The amount of benefit varied between hearing-impaired individuals, with some showing no improvement in SRT and one showing an improvement similar to that for normal-hearing subjects. The reduced ability to take advantage of TFS information in speech may partially explain why subjects with cochlear hearing loss get less benefit from listening in a fluctuating background than normal-hearing subjects. TFS information may be important in identifying the temporal “dips” in such a background.

Keywords

43.71.Ky; 43.66.Sr; 43.71.Es [MW]

I. INTRODUCTION

Information in speech is redundant. For normal-hearing subjects, this means that the signal is robust to corruption, and that speech remains intelligible under adverse listening conditions, such as in high levels of background noise. In the normal auditory system, a complex sound like speech is filtered into frequency channels on the basilar membrane. The signal at a given place can be considered as a time-varying envelope superimposed on the more rapid fluctuations of a carrier (temporal fine structure, TFS) whose rate depends partly on the center frequency and bandwidth of the channel. The relative envelope magnitude across channels conveys information about the spectral shape of the signal and changes in the relative envelope magnitude indicate how the short-term spectrum changes over time. The TFS carries information both about the fundamental frequency (F0) of the sound (when it is periodic) and about its short-term spectrum. For example, if at a particular time there is a formant centered at frequency f_x (and hence one or more relatively intense stimulus components near f_x), then channels centered close to f_x will show a TFS synchronized to f_x , and this will be reflected in the patterns of phase locking in those channels (Young and Sachs, 1979). In the mammalian auditory system, phase locking tends to break down for

frequencies above 4–5 kHz (Palmer and Russell, 1986), so it is generally assumed that TFS information is not used for frequencies above that limit. The role of TFS in speech perception for frequencies below 5 kHz remains somewhat unclear.

Many studies have assessed the relative importance of TFS and envelope information for speech intelligibility, for normal-hearing subjects. “Vocoder” processing has been used to remove TFS information from speech, so allowing speech intelligibility based on envelope and spectral cues to be measured (Dudley, 1939; Van Tasell *et al.*, 1987; Shannon *et al.*, 1995). A speech signal is filtered into a number of channels (N), and the envelope of each channel signal is used to modulate a carrier signal, typically a noise (for a noise vocoder) or a sine wave with a frequency equal to the channel center frequency (for a tone vocoder). The modulated signal for each channel is filtered to restrict the bandwidth to the original channel bandwidth and the modulated signals from each channel are then combined. For a single talker, provided that N is sufficiently large, the resulting signal is highly intelligible to both normal-hearing and hearing-impaired subjects (Shannon *et al.*, 1995; Turner *et al.*, 1995; Baskent, 2006; Lorenzi *et al.*, 2006b). However, if the original signal includes both a target talker and a background sound, intelligibility is greatly reduced, even for normal-hearing subjects (Dorman *et al.*, 1998; Fu *et al.*, 1998; Qin and Oxenham, 2003; Stone and Moore, 2003), leading to the suggestion that TFS information may be important for separation of a talker and background into separate auditory streams (Friesen *et al.*, 2001).

As well as removing TFS information, vocoder processing also “smears” spectral information, an effect that is greatest when N is small. If the analysis filters are of a similar width to the auditory filters, however, the spectral information that is available to the auditory system is only slightly reduced compared to normal, though TFS information is still absent.

Another method for assessing the roles of different types of temporal information in speech is to attempt to remove envelope information but to leave TFS information (partially) intact. This was first attempted by infinite peak clipping of a wideband speech signal (Licklider and Pollack, 1948), and later by using the Hilbert transform (Bracewell, 1986) to separate envelope and TFS information in each of a number of frequency channels (Smith *et al.*, 2002). The TFS in each channel is preserved, but the envelope cues are removed, and the channel signals are then combined. Effectively, the processing in each channel behaves like a very fast compressor with an infinite compression ratio. For brevity, we will refer to this signal-processing method as “TFS processing.” At first sight, this method may appear to remove temporal envelope information and leave TFS intact. However, because envelope and TFS information are correlated, the envelope can be partially re-introduced by filtering in the peripheral auditory system, especially when the channels used in the processing have large bandwidths (Ghitza, 2001). This problem is reduced if the signal is split into many narrow channels before removal of the envelope information, although some envelope cues may remain.

Gilbert and Lorenzi (2006) investigated the extent to which these recovered cues could be used to identify vowel-consonant-vowel (VCV) nonsense syllables. They subjected nonsense syllables to TFS processing, and then passed the resulting signals through an array of filters that simulated filtering in the normal peripheral auditory system. The envelopes of the outputs of these filters were extracted and used to modulate tones at the center frequencies of the filters. This is similar to tone-vocoder processing. The modulated tones were summed and presented to normal-hearing subjects, who were asked to identify the consonant that was presented. When the TFS processing used a small number of broad channels, subjects could identify the consonant accurately from the recovered envelope information. However, when the number of channels used in the TFS processing was eight

or more, subjects scored close to chance. The authors concluded that if a signal is filtered into a sufficiently large number of channels before removing envelope cues, any recovered envelope cues are insufficient for intelligibility of VCVs. VCV syllables that are TFS processed with a large number of analysis channels are reasonably intelligible to normally hearing listeners, after some training (Lorenzi *et al.*, 2006b), which suggests that TFS cues alone can convey useful speech information.

Results from several studies have led to the suggestion that the ability to use TFS information is adversely affected by cochlear hearing loss. Much of this work has investigated the discrimination of synthetic complex sounds by hearing-impaired subjects (Lacher-Fougère and Demany, 1998; 2005; Moore and Skrodzka, 2002; Moore and Moore, 2003; Moore *et al.*, 2006; Hopkins and Moore, 2007). For example, Hopkins and Moore (2007) tested the ability of normal-hearing and hearing-impaired subjects to discriminate a harmonic complex tone from a frequency-shifted tone, in which all components were shifted up by the same amount in Hz (de Boer, 1956). The frequency-shifted tone had very similar temporal envelope and spectral envelope characteristics to the harmonic tone, but a different TFS. All tones were passed through a fixed bandpass filter, to reduce excitation-pattern cues. When the filter was centered on the 11th component, so that the components within the passband were unresolved, subjects with moderate cochlear hearing loss performed poorly, while normal-hearing subjects could do the task well. Hopkins and Moore concluded that moderate cochlear hearing loss usually led to a reduced ability to use TFS information.

The reason for this is not clear. One possibility is that the precision of phase locking is reduced by cochlear hearing loss. One study found that phase locking was reduced in animals with induced hearing loss (Woolf *et al.*, 1981), but another study found normal phase locking in such animals (Harrison and Evans, 1979). It is unclear whether the types of pathologies that cause cochlear hearing loss in humans lead to reduced phase locking. Another possible reason for a reduced ability to use TFS information is that TFS information could be decoded by cross correlation of the outputs of two points on the basilar membrane (Loeb *et al.*, 1983; Shamma, 1985). A deficit in this process, produced by a change in the traveling wave on the basilar membrane, would impair the ability to use TFS information even if phase locking were normal. The broader auditory filters typically associated with cochlear hearing loss (Liberman and Kiang, 1978; Glasberg and Moore, 1986) could also lead to a reduced ability to use TFS information. The TFS at the output of these broader filters in response to a complex sound will have more rapid fluctuations and be more complex than normal. Such outputs may be uninterpretable by the central auditory system (Sek and Moore, 1995; Moore and Sek, 1996).

A reduced ability to use TFS information could explain some of the perceptual problems of hearing-impaired subjects (Lorenzi *et al.*, 2006a). TFS information may be important when listening in background noise, especially when the background is temporally modulated, as is often the case when listening in “real life,” for example, when more than one person is speaking. Normal-hearing subjects show better speech intelligibility (or lower speech reception thresholds, SRTs) when listening in a fluctuating background than when listening in a steady background (Festen and Plomp, 1990; Baer and Moore, 1994; Peters *et al.*, 1998; Füllgrabe *et al.*, 2006), an effect which is sometimes called “masking release.” Hearing-impaired subjects show a much smaller masking release, and it has been suggested that this may be because they are poorer at “listening in the dips” of a fluctuating masker than normal-hearing subjects (Duquesnoy and Plomp, 1983; Peters *et al.*, 1998; Lorenzi *et al.*, 2006b). Reduced audibility may account for some of the reduction in masking release measured for hearing-impaired subjects (Bacon *et al.*, 1998), although the effect persists even when audibility is restored (Peters *et al.*, 1998; Lorenzi *et al.*, 2006a). TFS information may be important in “dip listening” tasks, as it could be used to identify points in the

stimulus when the level of the target is high relative to the level of the masker; if the target and masker do not differ in their TFS, or no TFS information is available, dip listening may be ineffective.

Some studies have investigated the ability of hearing-impaired subjects to use TFS information in speech. Buss *et al.* (2004) showed that there was a correlation between temporal processing as assessed with psychoacoustic tasks and the ability of hearing-impaired subjects to recognize words in quiet. Lorenzi *et al.* (2006b) attempted to measure the ability of young and elderly hearing-impaired subjects to use TFS information in speech more directly. They applied 16-channel TFS processing to VCV nonsense syllables and asked subjects to identify the consonant in each syllable. According to Gilbert and Lorenzi (2006), this number of channels should be sufficient to prevent the use of recovered envelope cues. Hearing-impaired subjects performed poorly at this task, while normal-hearing subjects scored around 90% correct after some training. Lorenzi *et al.* interpreted this result as indicating that the hearing-impaired subjects had a very limited ability to use the TFS information in the speech, whereas it was usable by normal-hearing subjects. Lorenzi *et al.* (2006b) also measured masking release for the young hearing-impaired subjects when listening to unprocessed speech in steady and modulated noise. The amount of masking release was highly correlated with the score obtained for speech in quiet that had been subjected to TFS processing. This result is consistent with the argument made earlier, that the ability to use TFS is important for listening in the dips of a background sound.

A potential problem with the use of TFS-processed signals is that during gaps in the speech in a particular processing channel, low-level recording noise is amplified to the same level as the speech information. This is because the process is equivalent to multi-channel compression with an infinite compression ratio; whatever the original envelope amplitude in a given channel, the output envelope amplitude is constant. Channels with no speech information at a particular time are filled with distracting background sound. As a result, TFS-processed speech sounds harsh and very noisy. This may pose a particular problem to hearing-impaired subjects who, because of their broadened auditory filters, would suffer more from masking between channels. The problem becomes worse as the signal is split into more channels, as this results in more across-channel masking. Also, hearing-impaired listeners would be poorer at recovering any envelope cues that may still be available, again as a result of their broadened auditory filters. This could account for some of the difference in performance between normal-hearing and hearing-impaired subjects when listening to TFS speech. Here, a different approach was used to assess the use of TFS information by normal-hearing and hearing-impaired subjects. Rather than creating a signal that contains speech information only in its TFS, performance was measured as a function of the number of channels containing TFS information; the other channels were noise or tone vocoded, so that they conveyed only envelope information.

II. RATIONALE

Hopkins and Moore (2007) found that subjects with moderate cochlear hearing loss could make little use of TFS information to discriminate complex tones. If similar subjects were completely unable to use TFS information in speech, they would be expected to perform as well when listening to speech that had been vocoded to remove TFS information as when listening to unprocessed speech, provided that N was sufficiently large that the frequency selectivity of the processing was similar to or better than that of the peripheral auditory system of the subject, thus avoiding significant loss of spectral information. However, Baskent (2006) found that hearing-impaired subjects performed better in a phoneme identification task when the syllables were unprocessed than when they were processed with

a 32-channel noise-band vocoder. The disparity might arise because hearing-impaired subjects may be able to use TFS information at low carrier frequencies, but may be unable to use it at high frequencies. Hopkins and Moore (2007) showed that hearing-impaired subjects had a greatly reduced ability to discriminate the TFS of complex tones with unresolved components when all components were above 900 Hz, but they did not investigate sensitivity to TFS for lower frequencies. It is possible that subjects with moderate cochlear hearing loss are able to use TFS information below 900 Hz, which could explain why they performed better in the unprocessed condition than in the 32-channel vocoded condition in the study of Baskent (2006). If subjects with moderate cochlear hearing loss can use TFS information only at low carrier frequencies, progressively replacing vocoded information with unprocessed information, starting at low frequencies, should improve performance only up to a cut-off frequency above which TFS information cannot be used. This hypothesis was tested here.

SRTs corresponding to 50% correct keyword identification were measured for signals that were unprocessed for channels up to and including cut-off channel number (CO) and were vocoded for higher-frequency channels. The value of CO, which determined the amount of TFS information available in the signal, was varied from 0 to 32. A competing-talker background was used, because, as described earlier, TFS information may be particularly important for listening in backgrounds that have temporal “dips.”

III. METHOD

A. Subjects

Nine normal-hearing subjects and nine hearing-impaired subjects took part in the experiment. The normal-hearing subjects were aged between 18 and 27 years and had audiometric thresholds of 15 dB hearing level (HL) or less at octave frequencies between 250 and 8000 Hz. The audiograms of the test ears of the hearing-impaired subjects are shown in Fig. 1 (subjects HI 1 to HI 9) and the age of each subject is shown in parentheses. All hearing-impaired subjects had airborne gaps of 15 dB or less, and normal tympanograms, suggesting that their hearing loss was cochlear in origin. Hearing-impaired subjects were tested with the “TEN HL” test, which indicated no cochlear dead region for any subject (Moore *et al.*, 2004).

B. Speech material

Subjects were asked to repeat sentences presented in a competing talker background. The background began 500 ms before the target sentence, and continued after the target sentence had finished for about 700 ms (the exact value depended on the length of the target sentence). Each sentence list was added to a randomly chosen portion of a passage of continuous prose spoken by a competing talker. Long gaps between sentences and pauses for breath were removed from the background passage by hand editing. The same passage was used in both training and testing sessions. Both the target sentences and the competing talker passage had the same long-term spectral shape; for frequencies up to 500 Hz, the spectrum level was roughly constant, and for frequencies above 500 Hz the spectrum level fell by 9 dB per octave. For the training session, IEEE sentences were used (Rothauser *et al.*, 1969). For the testing session, sentences were taken from the adaptive sentence list (ASL) corpus (MacLeod and Summerfield, 1990). Both target and competing talkers were male speakers of British English. The target talker had a fundamental frequency (F0) range of about 130–200 Hz, and the competing talker had a larger F0 range of about 130–280 Hz. The target and background speech were added together at the appropriate signal-to-background ratio (SBR) before processing.

C. Processing and equipment

Speech signals were split into 32 channels with center frequencies spanning the range 100–10,000 Hz, with an array of linear-phase, finite-impulse-response (FIR) filters. The filters had a variable order so that the transition bands of each filter had similar slopes when plotted on a logarithmic frequency scale. Each filter was designed to have a response of -6 dB at the frequencies at which its response intersected with the responses of the two adjacent filters. Channel edges were regularly spaced on an equivalent-rectangular-bandwidth (ERB_N) number scale and each channel was 1 ERB_N wide (Glasberg and Moore, 1990). This filtering was designed to simulate the frequency selectivity of the normal auditory system, so that the processing preserved nearly all of the spectral information available in the original signal. The signals from each channel were time aligned to compensate for the time delays introduced by the bandpass filtering. Stimuli were processed with nine values of CO (0, 4, 8, 12, 16, 20, 24, 28, 32). Channels with channel numbers up to and including CO were not processed further. Channels with channel numbers above CO were vocoded. The signals from these channels were half-wave rectified and these rectified signals were used to modulate white noises.¹ Each modulated noise was subsequently filtered with the initial analysis filters and shaped to have the same spectral shape as the long-term spectrum of the original target speech from that channel. Consequently, envelope fluctuations with frequencies greater than half of the channel bandwidth were attenuated. After processing, the signals from the vocoded and unprocessed channels were added together. All signals were generated with a high-quality 16 bit PC soundcard (Lynx One) at a sampling rate of 22,050 Hz, passed through a Mackie 1202-VLZ mixing desk and presented to the subject monaurally via Sennheiser HD580 headphones. Subjects were seated in a double-walled sound-attenuating chamber.

D. Procedure

Microphones were placed in both the chamber and control room to allow communication between the experimenter and subject, although the control room microphone was only routed to the chamber headphones in the gaps between stimulus presentations. Target speech was presented to the normal-hearing subjects at a constant rms level of 65 dB sound pressure level (SPL), which was equivalent to a spectrum level of 36.6 dB (re 20 μ Pa) between 100 and 500 Hz (see Sec. III B for a description of the spectral shape). The level of the competing talker was varied to give the appropriate SBR, except when the SBR was less than -16 dB. Below this SBR, the level of the competing talker was not increased further, but instead the level of the target speech was reduced, to prevent the combined signal becoming uncomfortably loud. In practice, this was not necessary for any of the hearing-impaired subjects.

Previous studies have shown that audibility can account for some of the difference in performance between normal-hearing and hearing-impaired subjects listening in a temporally modulated background if stimuli are presented at the same level to both groups of subjects (Bacon *et al.*, 1998; George *et al.*, 2006). To reduce such effects, gains were applied to the combined target and background signal as prescribed by the CAMEQ hearing aid fitting method, according to the audiometric thresholds of each subject (Moore *et al.*, 1998). Gains were specified at audiometric frequencies between 250 and 6000 Hz. The CAMEQ gains are designed to ensure speech audibility between these frequencies. Relatively more gain is prescribed for higher frequencies and this compensates for the

¹Normally, the rectification would be followed by lowpass filtering, or the Hilbert transform would be used to extract the envelope. The omission of this stage in our processing meant that the modulator contained high-frequency components related to the TFS of the signal. However, these high-frequency components resulted in sidebands that were removed by the subsequent bandpass filtering. Listening tests and physical measurements confirmed that the processing used here gave results that were almost identical to those obtained when the Hilbert transform was used to extract the envelope.

increased upward spread of masking that is expected at higher overall levels, so helping to avoid the “rollover” effect on speech intelligibility as overall level increases (Fletcher, 1953; Studebaker *et al.*, 1999). The CAMEQ gains were applied to the processed signals using a linear-phase FIR filter with 443 taps.

E. Audibility calculations

To check that the target speech would be audible for the hearing-impaired subjects after the CAMEQ gains were applied, excitation patterns were calculated for a signal that had the same long-term average spectrum as the target speech signal used for each subject. The spectrum for each subject was obtained by determining the long-term average spectrum of the speech with an overall level of 65 dB SPL and adding the CAMEQ gains at each frequency (with interpolation of gains for frequencies between the values specified by CAMEQ). Mean excitation levels between 100 and 8000 Hz were calculated for each subject using a model similar to that proposed by Moore and Glasberg (2004), but updated to incorporate the middle ear transfer function proposed by Glasberg and Moore (2006). Excitation levels are calculated relative to the excitation evoked by a 1000 Hz tone presented in free field with frontal incidence at a level of 0 dB SPL. The model allows the audiometric thresholds (in dB HL) of the individual subject to be entered. Default values were assumed for the proportion of the hearing loss attributed to outer hair cell and inner hair cell dysfunction. The model also gave estimates of the excitation level at threshold as a function of frequency for each subject. Figure 2 shows the excitation level at threshold and the mean excitation evoked by the (amplified) speech signal for each subject. The excitation level for the target speech was well above threshold excitation level except at very high or very low frequencies for some subjects. For most subjects, and for frequencies between 500 and 5000 Hz (the frequency range that is most important for speech intelligibility), the excitation level of the target speech was more than 15 dB above the excitation level at threshold, meaning that the entire dynamic range of the speech would have been audible (ANSI, 1997).

F. Training

Previous studies using vocoded speech material have shown large learning effects (Stone and Moore, 2003; 2004; Davis *et al.*, 2005), so a training period was included before testing. Training lasted approximately 1 h, and took place separately to the testing session. First, subjects were played two passages of connected discourse to familiarize them with the task of listening in a competing talker background and to introduce the vocoder-processed speech. The first passage was unprocessed, and the second was vocoded across all 32 channels. The level of the competing talker was initially low, but was increased gradually throughout the passages. Subjects were instructed to listen to the target talker for as long as possible. The hearing-impaired subjects found this difficult, and so were given transcripts of the target passages to follow, which made the task easier.

For the next phase of training, IEEE sentences were presented at a fixed SBR. Six lists were presented, each made up of ten sentences. The sentences were processed with different values of CO and an SBR was selected by the experimenter to yield scores of approximately 70% correct. Subjects were required to repeat each sentence and the number of correctly identified key words was recorded. When subjects did not repeat the sentence perfectly, they were told the correct answer, and the sentence was repeated.

Finally, subjects were given an opportunity to practice the task used in the testing session. Four word lists similar to those in the ASL corpus were used. The same procedure was used as for the testing session, as described below.

G. Testing

Two consecutively presented ASL sentence lists were used for each condition and the order of presentation of conditions was counterbalanced across subjects. The SBR of the target and competing talker was varied adaptively. If a subject identified two or more keywords correctly in a sentence, the next sentence was presented with a SBR that was k dB lower, and if the subject identified fewer than two keywords correctly, the next sentence was presented with a SBR that was k dB higher. Before the third turnpoint was reached, k was equal to 4 dB; subsequently it was equal to 2 dB. The first sentence in each list was initially presented at an adverse SBR, at which the subject was expected to identify no keywords correctly. If the subject scored fewer than two keywords correctly, this sentence was repeated at an SBR that was 4 dB higher until at least two keywords were correctly identified. Subsequent sentences in each list were presented once only. For each sentence list, the total number of keywords presented at each SBR was recorded, as well as the number of keywords that were identified correctly for each SBR. The first sentence in each list was not included in these totals, as subjects could have heard this sentence more than once.

H. Analysis

For each SBR, the total keywords presented and keywords correct were summed for the two sentences lists that were presented for each condition. These values were used to perform a probit analysis (Finney, 1971), from which the SRT corresponding to the SBR required for 50% correct identification was estimated for each subject and each condition. In some cases, because of the scatter in the data, the probit analysis failed to fit the data and gave a slope of the psychometric function that was not significantly different from zero. This happened for at least one of the conditions for five of the normal-hearing subjects, but for only one of the hearing-impaired subjects (HI 8). For these cases, the SRT was estimated by plotting the proportion of correctly identified words against the SBR at which the words were presented. A line was drawn by eye to best fit the data points, and this line was used to exclude points from the probit analysis that did not fit the general trend. The probit analysis was then redone. In one case (NH 5, CO=20), after this procedure the probit analysis still did not give a psychometric function with a slope significantly different from zero, so the SRT for this case was treated as a missing data point for the remaining analysis.

An analysis of variance (ANOVA) was performed on all of the data from the normal-hearing and hearing-impaired subjects, with a within-subjects factor of CO and a between-subjects factor of subject type (normal hearing or hearing impaired).

IV. RESULTS

Figure 3 shows the mean data for both normal-hearing and hearing-impaired subjects. Mean SRTs are plotted for each value of CO/ N . The hearing-impaired subjects performed more poorly than the normal-hearing subjects in all conditions, but the difference in performance varied with CO; for larger values of CO, the difference in performance between normal-hearing and hearing-impaired subjects was greater. The main effects of subject type and CO were significant [$F(1,8)=95.2$, $p<0.001$ and $F(8,128)=53.9$, $p<0.001$, respectively], and there was also a significant interaction between subject type and CO [$F(8,128)=12.2$, $p<0.001$].

CO had a greater effect on performance for the normal-hearing subjects than for the hearing-impaired subjects. For example, both subject groups performed better when speech was completely unprocessed (CO=32) than when it was completely vocoded (CO=0), but the difference in performance between these conditions was much greater for the normal-

hearing than for the hearing-impaired subjects (mean differences were 15.8 and 4.9 dB, respectively). *Post hoc* Fisher's least-significant-difference (LSD) tests were used to determine whether the SRTs measured with different values of CO were significantly different from each other within each subject group. Tables I and II show the differences between mean scores for each value of CO for the normal-hearing subjects and hearing-impaired subjects, respectively. Values greater than the least significant difference are shown in bold.

Figure 4 shows results for individual hearing-impaired subjects. The mean results for the normal-hearing subjects are shown in the bottom-right panel for comparison. Between-subject variability in overall performance was larger for the hearing-impaired than for the normal-hearing subjects. The pattern of results across conditions also varied more between hearing-impaired subjects. The benefit gained from the additional TFS information that was present when CO was large varied, with some hearing-impaired subjects benefiting little, if at all (for example, HI 4 and HI 5) and others benefiting almost as much as the normal-hearing subjects (HI 8).

V. DISCUSSION

Normal-hearing subjects appear to benefit more than hearing-impaired subjects from the replacement of vocoded speech information with unprocessed speech information. This is consistent with the idea that the hearing-impaired subjects had a reduced ability to use TFS information, which is consistent with previously published results (Lorenzi *et al.*, 2006b; Moore *et al.*, 2006; Hopkins and Moore, 2007). Both groups did, however, improve as CO increased, though the amount of benefit from the additional TFS information varied across hearing-impaired subjects. This may reflect different abilities to use TFS information among hearing-impaired subjects with broadly similar audiometric thresholds, which could account for the weak correlation between audiometric thresholds and the ability to understand speech in noise previously reported for hearing-impaired subjects (Festen and Plomp, 1983; Glasberg and Moore, 1989). Other studies have also reported large individual differences in performance between hearing-impaired subjects when tasks require the use of TFS information (Buss *et al.*, 2004; Moore *et al.*, 2006).

One possible concern is that the mean age of the normal-hearing subjects was much less than the mean age of the hearing-impaired subjects (21.9 and 56.8 years, respectively), so the reduced benefit from the additional TFS might have been due to age rather than to hearing loss *per se*. Some previous studies have been interpreted as indicating that older subjects with near-normal audiometric thresholds have temporal processing deficits (Pichora-Fuller, 2003; Pichora-Fuller *et al.*, 2006). However, other studies have tested young and elderly normal-hearing subjects listening to target speech in a temporally modulated background similar to that used here, and found relatively small differences in performance between the two groups (Takahashi and Bacon, 1992; Peters *et al.*, 1998; Dubno *et al.*, 2002; Lorenzi *et al.*, 2006a). The differences were much smaller than the difference in performance seen here between the hearing-impaired and normal-hearing subjects in the unprocessed condition (CO=32). It is possible that some of the reduced ability to use TFS information seen for the hearing-impaired subjects in the current study could be attributed to their age, rather than their hearing loss. However, the pattern of results for the two young hearing-impaired subjects tested here (HI 2, 23 years and HI 6, 26 years) did not differ markedly from the pattern of the mean data for the hearing-impaired subjects, suggesting that hearing loss, rather than age was the important factor contributing to the reduced ability to use TFS information. The benefit from the addition of TFS was quantified as the difference between the SRT for CO=0 and CO=32. The correlation between benefit and age

for the hearing-impaired subjects was not significant ($r=-0.26$, $p=0.50$). Again, this suggests that age was not the important factor determining the benefit from added TFS information.

Phase locking in the normal auditory system is widely believed to break down for frequencies above 4000–5000 Hz (Palmer and Russell, 1986; Moore, 2003). If this is true, TFS information above 4000–5000 Hz should be unusable, and so no improvement in performance would be expected when TFS information was added in the higher-frequency channels. Consistent with this, the Fisher LSD tests showed no significant difference in performance for the normal-hearing subjects for CO values from 24 to 32; CO=24 corresponds to a cut-off frequency of 4102 Hz. For the hearing-impaired subjects, performance appears to plateau at a lower value of CO. With one exception, LSD tests revealed no significant difference in performance for values of CO from 16 to 32; CO=16 corresponds to a frequency of 1605 Hz. This is consistent with the idea that hearing-impaired subjects may be able to use TFS information at low frequencies, but are unable to use higher-frequency TFS information, even for frequencies where phase locking is believed to be robust in the normal auditory system.

A possible concern when comparing hearing-impaired and normal-hearing subjects is that differences in performance may be explained by differences in audibility. For our experiment, this explanation seems unlikely. Figure 2 shows that the entire dynamic range of speech would have been audible for most of the hearing-impaired subjects for frequencies between 500 and 5000 Hz. Audibility was compromised at very low and high frequencies for some subjects, but this reduced audibility is unlikely to have affected speech intelligibility and cannot account for the large differences between hearing-impaired and normal-hearing subjects.

The improvement in SRT as CO increased has been interpreted so far as reflecting an ability to use TFS information, but this is not the only interpretation of these results. Another possibility is connected with the idea that a noise-band vocoder may introduce distracting or masking low-frequency modulations into the signal. Whitmal *et al.* (2007) found that normal-hearing subjects scored better when tested with a tone vocoder than with a noise vocoder (see also Dorman *et al.*, 1997). They suggested that modulations introduced by the noise carrier may have caused a reduction in speech intelligibility. The modulation spectrum of a bandpass filtered noise is triangular (Schwartz, 1970), with more modulation energy at low frequencies. This means that the modulations introduced by the noise carrier are dominated by modulation frequencies that are similar to those thought to be important in understanding speech (Drullman *et al.*, 1994a; 1994b; Shannon *et al.*, 1995). When the number of analysis channels is large (so the channel widths are small), this is an even greater problem, as higher-frequency modulations are removed when the channel signals are filtered after vocoder processing, leaving the signal even more dominated by low-frequency noise modulations. It is possible that the normal-hearing subjects and the hearing-impaired subjects who showed greater improvement as CO increased did not benefit from the additional TFS information, but performed better because the spurious modulations introduced by the noise carrier were reduced, as the proportion of the signal that was vocoded was reduced.

Another factor that might have influenced the change in performance with increasing CO is connected with the effect of the processing on the representation of high-rate envelope fluctuations. Rosen (1992) suggested that modulations between 50 and 500 Hz are important in providing information on voice periodicity, and this voice periodicity information is important for listening in a competing talker background (Brokx and Nootboom, 1982; Assmann and Summerfield, 1990). In experiment one, the speakers were male, with F0s varying between 130 and 280 Hz. The processing used 32 channels, which were equally

spaced on an ERB_N -number scale; each channel was 1 ERB_N wide. This was intended to simulate the frequency selectivity of the normal auditory system. The highest modulation rate that can be carried by a channel is determined by the bandwidth of the channel. The filtering that was used subsequent to modulation of the channel carriers would have attenuated the sidebands produced by the modulation, hence reducing the modulation depth for high rates. As a result, voice periodicity information would have been partially removed from channels tuned to lower center frequencies.

This restriction of periodicity information would not have any important effects for normally hearing subjects, because the filters used in the processing had comparable widths to the “normal” auditory filter. However, hearing-impaired subjects generally have broader auditory filters than normal-hearing subjects, so, for unprocessed speech, higher-frequency modulation sidebands would be attenuated less by the peripheral auditory system. Consequently, post-processing filtering of the vocoded signal into 1 ERB_N -wide channels could reduce the periodicity information available to the hearing-impaired subjects, and this could be a reason for their worse performance with $CO=0$ than with $CO=32$.

These possible explanations for the improvement in performance with increasing CO found for the normal-hearing subjects and some of the hearing-impaired subjects were investigated in experiment two.

VI. EXPERIMENT TWO

A. Rationale

Experiment two was broadly similar to experiment one, but a tone vocoder was used rather than a noise vocoder. The carrier signals were sine waves with frequencies equal to the channel center frequencies. No random modulations were introduced by the carrier signals, unlike for the noise vocoder used in experiment one.

Previous work has suggested that subjects with moderate cochlear hearing loss have auditory filters that are between two and four times as broad as those for normal-hearing subjects (Glasberg and Moore, 1986; Moore, 2007). In experiment two, the signal was divided into either 8 or 16 channels before processing, rather than 32, so that each channel was wider, and more comparable to the auditory filters of the hearing-impaired subjects (channels were 4 or 2 ERB_N wide rather than 1, as previously). This avoided the possible loss of modulation at F_0 rates. A consequence of splitting the signal into fewer channels before vocoder processing is that more spectral detail from the original signal is lost. If the filters used in the processing are broader than those in the peripheral auditory system, as they would be for normal-hearing subjects, this in itself may lead to poorer performance. To check the effect of decreasing N , and to allow comparison with the results of experiment one, a condition was run with $N=32$ and $CO=0$, so that stimuli were fully tone vocoded, but with the same value of N as for experiment one.

B. Method

1. Subjects—Five of the normal-hearing and seven of the hearing-impaired subjects who took part in experiment one also took part in experiment two. Four normal-hearing and two hearing-impaired subjects were newly recruited. Recruitment criteria were the same as for experiment one. The audiograms and ages of all of the hearing-impaired subjects used in both experiments are shown in Fig. 1. HI 10 and HI 11 took part in experiment two only, HI 4 and HI 6 took part in experiment one only, and the remaining subjects took part in both experiments.

2. Speech material—ASL lists were used for training, as most of the subjects had already heard these in the testing session for experiment one. For the testing session, Bench-Kowal-Bench (BKB) sentence material was used, which is similar in style to the ASL sentence material (Bench and Bamford, 1979).

3. Processing—Sentences were processed in a similar way as for experiment one, but a tone vocoder was used rather than a noise vocoder, and the signal was split into 8 or 16 channels before processing rather than 32 (so channels were 4 or 2 ERB_N wide rather than 1 ERB_N wide). Signals were split into channels, and the envelope of each channel was extracted as before. Sine waves with frequencies equal to the center frequency of each channel were used as carrier signals rather than noise bands, and these sine waves were modulated with the envelope of the original channel signals. As before, processed channel signals were filtered to remove sidebands that were introduced as a result of the processing, so limiting the frequency of modulation that could be carried in each channel.

4. Conditions and procedure—For $N=8$, values of CO were 2, 4, 6 and 8. For $N=16$, values of CO were 0, 4, 8 and 12 (note that the condition where $N=16$ and CO=16 is the same as $N=8$ and CO=8, so this condition was not retested). For five of the normal-hearing subjects, and five of the hearing-impaired subjects, an additional condition was tested, with $N=32$ and CO=0, but still using a tone vocoder. The procedure for training and testing sessions was the same as for experiment one, except for the differences in sentence material, as noted previously. Data were analyzed in the same way as for experiment one.

C. Results

The results are summarized in Fig. 5. Mean SRTs are plotted against CO/ N . A given value of CO/ N corresponds to a fixed frequency, as indicated at the top of the panels in Fig. 5. As in experiment one, the hearing-impaired subjects performed more poorly than the normal-hearing subjects for all conditions, and the difference in performance between the two groups was greatest when CO/ $N=1$. An ANOVA was performed with N and CO/ N as within-subject factors and subject type as a between-subject factor. The main effects and two-way interactions were all highly significant ($p < 0.001$) and the three-way interaction was also significant [$F(4,60)=3.33$, $p=0.02$].

Differences between mean results for different values of CO/ N are shown in Tables III and IV for the normal-hearing and hearing-impaired subjects, respectively. For the normal-hearing subjects, SRTs did not differ significantly for CO/ $N=0.75$ and 1 (for $N=8$) or for CO/ $N=0.5$, 0.75 and 1 (for $N=16$). Thus, performance reached a plateau for higher values of CO/ N , as found in experiment one. For the hearing-impaired subjects, SRTs reached a plateau at a lower value of CO/ N . SRTs did not differ significantly for CO/ $N=0.25$, 0.5, 0.75 and 1 (for $N=8$) or for CO/ $N=0.5$, 0.75 and 1 (for $N=16$). For $N=16$, the SRT for the normal-hearing subjects decreased by 16.0 dB as CO/ N was increased from 0 to 1. This is similar to, but slightly larger than the decrease found in experiment one for 32-channel processing. For $N=8$, the decrease was larger, at 22.0 dB, because the SRT was higher for $N=8$ than for $N=16$ when the signal was fully vocoded (CO/ $N=0$). For the hearing-impaired subjects, the decrease in SRT with increasing CO/ N was much smaller, 4.1 dB for $N=16$ and 6.4 dB for $N=8$. Thus, as found in experiment one, the benefit of progressively adding TFS information (by increasing CO/ N) was much smaller for the hearing-impaired than for the normal-hearing subjects, despite the use of a tone vocoder and a smaller number of channels in experiment two.

Fisher LSD tests revealed that normal-hearing subjects performed better for $N=16$ than for $N=8$, except when CO/ $N=0.75$ (see Table V). Hearing-impaired subjects performed better

for $N=16$ when $CO=0$, but not for higher values of CO . For $CO/N=0.25$, performance was significantly better for $N=8$ than $N=16$ for the hearing-impaired subjects, and when $CO/N=0.5$, there was no significant difference in performance for $N=8$ and 16 .

Three student's t tests were performed to assess whether there was a significant effect of number of channels ($N=8, 16$ or 32) when $CO=0$ (i.e., when the signal was completely vocoded), for those hearing-impaired and normal-hearing subjects who were tested in all conditions. A Bonferroni correction for multiple comparisons was applied. The normal-hearing subjects performed significantly better when $N=32$ than when $N=16$ ($p=0.01$), whereas the hearing-impaired subjects did not perform significantly differently for the two conditions ($p=0.12$). However, the SRT of the hearing-impaired subjects did increase when the value of N was decreased further to 8 , and the difference in SRT between 32 and 8 channels was significant ($p=0.05$).

Performance in the condition when $N=32$ and $CO=0$ was much better than for the same condition in experiment one, for both normal-hearing and hearing-impaired subjects (the mean SRTs were 7.8 and 4.7 dB lower, respectively, in experiment two).

D. Discussion

The pattern of results was similar for experiments one and two, for both the normal-hearing and hearing-impaired subjects. This suggests that neither the random amplitude fluctuations introduced by the noise vocoder nor the partial removal of high-rate envelope modulations by the relatively narrow filters used in experiment one entirely explain the (small) benefit of adding TFS information found for the hearing-impaired subjects in experiment one. Rather, the results are consistent with the idea that the improvement in SRT as CO increased resulted mainly from the use of TFS information, and the improvement was smaller for the hearing-impaired than for the normal-hearing subjects because the latter have a greatly reduced ability to use TFS information.

Performance was better for the tone vocoder (experiment two) than for the noise vocoder (experiment one) with matched N , but different sentence material was used for the two experiments, which makes the comparison difficult. Better performance has been reported for the ASL sentence lists used in experiment one than for the BKB lists used in experiment two (MacLeod and Summerfield, 1990). If the same sentence material had been used for the two experiments, an even larger difference might have been observed. Overall, the comparison of results for experiments one and two with $N=32$ and $CO=0$ is consistent with previous results (Dorman *et al.*, 1997; Whitmal *et al.*, 2007) and with the hypothesis that the random amplitude fluctuations introduced by the noise vocoder have a deleterious effect on performance.

Normal-hearing subjects benefited from the greater spectral information in the vocoded signal when $N=32$ than when $N=16$, whereas the hearing-impaired subjects did not. This is consistent with what would be expected from the greater auditory-filter bandwidths that are typically found for hearing-impaired subjects (Glasberg and Moore, 1986; Moore, 2007). The normal-hearing subjects, who were expected to have relatively sharp filters, benefited significantly from the greater spectral information provided by more channels, while the hearing-impaired subjects, who were expected to have relatively broad filters, benefited little, if at all. These findings are consistent with those of Baskent (2006), who found a similar plateau in performance as N increased above 16 for hearing-impaired subjects.

Previous work has concentrated on reduced frequency selectivity as an explanation for the supra-threshold deficits associated with moderate cochlear hearing loss. Reduced frequency selectivity means that hearing-impaired listeners are more susceptible to masking across

frequencies, and this partially explains why they perform poorly when listening in background sounds. The different patterns of performance for the normal-hearing and hearing-impaired subjects in the results presented here cannot be accounted for by differences in across-frequency masking. Similar amounts of masking would be expected in all of the conditions that were tested, so if deficits caused by cochlear hearing loss were only a result of across-frequency masking, a similar pattern of performance would have been expected for the normal-hearing and hearing-impaired subjects. Reduced spectral resolution may account for the differences in performance between the subject groups when $CO=0$, when no TFS information was available. Indeed, the fact that the hearing-impaired subjects were tested using higher overall sound levels than the normal-hearing subjects might have exacerbated this effect, since auditory filters tend to broaden at high levels (Glasberg and Moore, 1990). However, changes in auditory-filter bandwidth with level tend to be smaller for hearing-impaired than for normally hearing subjects (Moore, 2007), so the effect of level is unlikely to be large. For whatever reasons, speech intelligibility worsens at very high sound levels for both normal-hearing and hearing-impaired subjects (Summers and Cord, 2007), so the higher level used here for the hearing-impaired subjects may have contributed to their poorer overall performance. However, the increasing deficit as TFS information was added is unlikely to reflect this “rollover effect,” since the speech level was the same for all values of CO .

Another possible factor that may have influenced our results is that some of the hearing-impaired subjects may not have been able to make effective use of information conveyed by the higher-frequency components in the speech, even though those components would have been audible. In other words, the lack of benefit from adding TFS information may reflect a general lack of ability to use information from the higher-frequency components in speech. However, a reduced ability to use information from the higher-frequency (>2000 Hz) components in speech has mainly been found for subjects with hearing losses greater than about 60 dB (Ching *et al.*, 1998; Hogan and Turner, 1998; Vickers *et al.*, 2001). Hearing-impaired subjects with hearing losses less than 60 dB do seem to be able to make effective use of information from such high-frequency components (Skinner and Miller, 1983; Vickers *et al.*, 2001; Baer *et al.*, 2002). Several of our subjects had hearing losses of 60 dB or less for frequencies up to about 4 kHz, but they still failed to show a clear benefit as CO/N was increased above 0.5 (corresponding to a frequency of 1600 Hz). For example, HI 4 had audiometric thresholds of 55 dB or better for all frequencies up to 6000 Hz, but did not show any benefit of increasing CO/N .

Overall, it seems likely that the increasing deficit of the hearing-impaired subjects as CO/N was increased reflects a different ability to use TFS information between the two groups. It is possible that reduced frequency selectivity may contribute to a reduced ability to use TFS information. The outputs of broader auditory filters would have a more complex TFS than the outputs of narrower filters, as found in normal-hearing subjects. It is possible that such complex outputs may not be interpretable by the central auditory system. Deficits in phase locking would also be expected to reduce the ability to use TFS, as inaccuracies in phase locking would degrade information about TFS available to the central auditory system. Understanding the mechanism responsible for the observed deficit in the ability to use TFS would be an interesting topic for future research.

The individual differences in benefit from the addition of TFS information found here between hearing-impaired subjects may explain the relatively poor correlation between audiometric thresholds and speech intelligibility in noise (Festen and Plomp, 1983; Glasberg and Moore, 1989). For the subjects tested here, the amount of benefit gained from addition of TFS information [$(SRT \text{ for } CO=0) - (SRT \text{ for } CO=32)$] was not significantly correlated with the mean of audiometric thresholds at 250, 500, 1000, 2000 and 4000 Hz ($r=-0.04$,

$p=0.92$). The ability to use TFS information may be a factor affecting speech intelligibility that is not well predicted by traditional audiometry.

VII. CONCLUSIONS

Hearing-impaired subjects benefited less than normal-hearing subjects from TFS information added to a vocoded speech signal when listening in a competing talker background. The amount of benefit varied between subjects, with some not benefiting at all. The same general pattern of results was found regardless of whether a noise vocoder or a tone vocoder was used. It is argued that subjects with moderate cochlear hearing loss have a limited ability to use TFS information, especially for medium and high frequencies. This may explain some of the speech perception deficits found for such subjects, especially the reduced ability to take advantage of temporal dips in a competing background.

Acknowledgments

This work was supported by the MRC (UK). We thank Christian Lorenzi and one anonymous reviewer for helpful comments on an earlier version of this paper.

References

- ANSI. ANSI S3.5-1997, Methods for the Calculation of the Speech Intelligibility Index. American National Standards Institute; New York: 1997.
- Assmann PF, Summerfield AQ. Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *J. Acoust. Soc. Am.* 1990; 88:680–697. [PubMed: 2212292]
- Bacon SP, Opie JM, Montoya DY. The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J. Speech Lang. Hear. Res.* 1998; 41:549–563. [PubMed: 9638921]
- Baer T, Moore BCJ. Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J. Acoust. Soc. Am.* 1994; 95:2277–2280. [PubMed: 8201124]
- Baer T, Moore BCJ, Kluk K. Effects of lowpass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J. Acoust. Soc. Am.* 2002; 112:1133–1144. [PubMed: 12243160]
- Baskent D. Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels. *J. Acoust. Soc. Am.* 2006; 120:2908–2925. [PubMed: 17139748]
- Bench, J.; Bamford, J. *Speech-Hearing Tests and the Spoken Language of Hearing-Impaired Children.* Academic; London: 1979.
- Bracewell, RN. *The Fourier Transform and its Applications.* McGraw–Hill; New York: 1986. p. 267-272.
- Brokx JPL, Nootboom SG. Intonation and the perceptual separation of simultaneous voices. *J. Phonetics.* 1982; 10:23–36.
- Buss E, Hall JW III, Grose JH. Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear Hear.* 2004; 25:242–250. [PubMed: 15179115]
- Ching T, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *J. Acoust. Soc. Am.* 1998; 103:1128–1140. [PubMed: 9479766]
- Davis MH, Johnsrude IS, Hervais-Adelman A, Taylor K, McGettigan C. Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *J. Exp. Psychol. Gen.* 2005; 134:222–241. [PubMed: 15869347]
- de Boer E. Pitch of inharmonic signals. *Nature (London).* 1956; 178:535–536. [PubMed: 13358790]
- Dorman MF, Loizou PC, Rainey D. Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J. Acoust. Soc. Am.* 1997; 102:2403–2411. [PubMed: 9348698]

- Dorman MF, Loizou PC, Fitzke J, Tu Z. The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6–20 channels. *J. Acoust. Soc. Am.* 1998; 104:3583–3585. [PubMed: 9857516]
- Drullman R, Festen JM, Plomp R. Effect of reducing slow temporal modulations on speech reception. *J. Acoust. Soc. Am.* 1994a; 95:2670–2680. [PubMed: 8207140]
- Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am.* 1994b; 95:1053–1064. [PubMed: 8132899]
- Dubno JR, Horwitz AR, Ahlstrom JB. Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *J. Acoust. Soc. Am.* 2002; 111:2897–2907. [PubMed: 12083223]
- Dudley H. Remaking speech. *J. Acoust. Soc. Am.* 1939; 11:169–177.
- Duquesnoy AJ, Plomp R. The effect of a hearing-aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise. *J. Acoust. Soc. Am.* 1983; 73:2166–2173. [PubMed: 6875102]
- Festen JM, Plomp R. Relations between auditory functions in impaired hearing. *J. Acoust. Soc. Am.* 1983; 73:652–662. [PubMed: 6841805]
- Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 1990; 88:1725–1736. [PubMed: 2262629]
- Finney, DJ. Probit Analysis. Cambridge University Press; Cambridge: 1971.
- Fletcher, H. Speech and Hearing in Communication. Van Nostrand; New York: 1953.
- Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.* 2001; 110:1150–1163. [PubMed: 11519582]
- Fu Q-J, Shannon RV, Wang X. Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *J. Acoust. Soc. Am.* 1998; 104:3586–3596. [PubMed: 9857517]
- Füllgrabe C, Berthommier F, Lorenzi C. Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.* 2006; 211:74–84. [PubMed: 16289579]
- George EL, Festen JM, Houtgast T. Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 2006; 120:2295–2311. [PubMed: 17069325]
- Ghitza O. On the upper cutoff frequency of the auditory criticalband envelope detectors in the context of speech perception. *J. Acoust. Soc. Am.* 2001; 110:1628–1640. [PubMed: 11572372]
- Gilbert G, Lorenzi C. The ability of listeners to use recovered envelope cues from speech fine structure. *J. Acoust. Soc. Am.* 2006; 119:2438–2444. [PubMed: 16642856]
- Glasberg BR, Moore BC. Prediction of absolute thresholds and equal-loudness contours using a modified loudness model. *J. Acoust. Soc. Am.* 2006; 120:585–588. [PubMed: 16938942]
- Glasberg BR, Moore BCJ. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J. Acoust. Soc. Am.* 1986; 79:1020–1033. [PubMed: 3700857]
- Glasberg BR, Moore BCJ. Psychoacoustic abilities of subjects with unilateral and bilateral cochlear impairments and their relationship to the ability to understand speech. *Scand. Audiol. Suppl.* 1989; 32:1–25. [PubMed: 2711118]
- Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear. Res.* 1990; 47:103–138. [PubMed: 2228789]
- Harrison RV, Evans EF. Cochlear fiber responses in guinea pigs with well defined cochlear lesions. *Scand. Audiol. Suppl.* 1979; 9:83–92. [PubMed: 294692]
- Hogan CA, Turner CW. High-frequency audibility: Benefits for hearing-impaired listeners. *J. Acoust. Soc. Am.* 1998; 104:432–441. [PubMed: 9670535]
- Hopkins K, Moore BCJ. Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J. Acoust. Soc. Am.* 2007; 122:1055–1068. [PubMed: 17672653]
- Lacher-Fougère S, Demany L. Modulation detection by normal and hearing-impaired listeners. *Audiology.* 1998; 37:109–121. [PubMed: 9547924]

- Lacher-Fougère S, Demany L. Consequences of cochlear damage for the detection of interaural phase differences. *J. Acoust. Soc. Am.* 2005; 118:2519–2526. [PubMed: 16266172]
- Lieberman MC, Kiang NYS. Acoustic trauma in cats: Cochlear pathology and auditory-nerve activity. *Acta Oto-Laryngol., Suppl.* 1978; 358:1–63.
- Licklider JCR, Pollack I. Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech. *J. Acoust. Soc. Am.* 1948; 20:42–52.
- Loeb GE, White MW, Merzenich MM. Spatial cross correlation: A proposed mechanism for acoustic pitch perception. *Biol. Cybern.* 1983; 47:149–163. [PubMed: 6615914]
- Lorenzi C, Husson M, Ardoint M, Debrulle X. Speech masking release in listeners with flat hearing loss: Effects of masker fluctuation rate on identification scores and phonetic feature reception. *Int. J. Audiol.* 2006a; 45:487–495. [PubMed: 17005491]
- Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U.S.A.* 2006b; 103:18866–18869. [PubMed: 17116863]
- MacLeod A, Summerfield Q. A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *Br. J. Audiol.* 1990; 24:29–43. [PubMed: 2317599]
- Moore, BCJ. *An Introduction to the Psychology of Hearing*. 5th ed. Academic; San Diego: 2003.
- Moore, BCJ. *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*. 2nd ed. Wiley; Chichester: 2007.
- Moore BCJ, Glasberg BR. A revised model of loudness perception applied to cochlear hearing loss. *Hear. Res.* 2004; 188:70–88. [PubMed: 14759572]
- Moore BCJ, Moore GA. Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects. *Hear. Res.* 2003; 182:153–163. [PubMed: 12948610]
- Moore BCJ, Sek A. Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *J. Acoust. Soc. Am.* 1996; 100:2320–2331. [PubMed: 8865639]
- Moore BCJ, Skrodzka E. Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation. *J. Acoust. Soc. Am.* 2002; 111:327–335. [PubMed: 11833538]
- Moore BCJ, Alcántara JI, Glasberg BR. Development and evaluation of a procedure for fitting multi-channel compression hearing aids. *Br. J. Audiol.* 1998; 32:177–195. [PubMed: 9710336]
- Moore BCJ, Glasberg BR, Hopkins K. Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hear. Res.* 2006; 222:16–27. [PubMed: 17030477]
- Moore BCJ, Glasberg BR, Stone MA. New version of the TEN test with calibrations in dB HL. *Ear Hear.* 2004; 25:478–487. [PubMed: 15599194]
- Palmer AR, Russell IJ. Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner haircells. *Hear. Res.* 1986; 24:1–15. [PubMed: 3759671]
- Peters RW, Moore BCJ, Baer T. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.* 1998; 103:577–587. [PubMed: 9440343]
- Pichora-Fuller MK. Cognitive aging and auditory information processing. *Int. J. Audiol.* 2003; 42(Suppl. 2):2S26–32. [PubMed: 12918626]
- Pichora-Fuller MK, Schneider BA, Benson NJ, Hamstra SJ, Storzer E. Effect of age on detection of gaps in speech and nonspeech markers varying in duration and spectral symmetry. *J. Acoust. Soc. Am.* 2006; 119:1143–1155. [PubMed: 16521775]
- Qin MK, Oxenham AJ. Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *J. Acoust. Soc. Am.* 2003; 114:446–454. [PubMed: 12880055]
- Rosen S. Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. London, Ser. B.* 1992; 336:367–373. [PubMed: 1354376]

- Rothauser EH, Chapman WD, Guttman N, Nordby KS, Silbiger HR, Urbanek GE, Weinstock M. I.E.E.E. recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust.* 1969; 17:227–246.
- Schwartz, M. *Information Transmission, Modulation, and Noise*. McGraw–Hill; Kogakusha, Tokyo: 1970.
- Sek A, Moore BCJ. Frequency discrimination as a function of frequency, measured in several ways. *J. Acoust. Soc. Am.* 1995; 97:2479–2486. [PubMed: 7714264]
- Shamma SA. Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve. *J. Acoust. Soc. Am.* 1985; 78:1622–1632. [PubMed: 3840813]
- Shannon RV, Zeng F-G, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. *Science.* 1995; 270:303–304. [PubMed: 7569981]
- Skinner MW, Miller JD. Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. *Audiology.* 1983; 22:253–279. [PubMed: 6870667]
- Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory perception. *Nature (London).* 2002; 416:87–90. [PubMed: 11882898]
- Stone MA, Moore BCJ. Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *J. Acoust. Soc. Am.* 2003; 114:1023–1034. [PubMed: 12942981]
- Stone MA, Moore BCJ. Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task. *J. Acoust. Soc. Am.* 2004; 116:2311–2323. [PubMed: 15532662]
- Studebaker GA, Sherbecoe RL, McDaniel DM, Gwaltney CA. Monosyllabic word recognition at higher-than-normal speech and noise levels. *J. Acoust. Soc. Am.* 1999; 105:2431–2444. [PubMed: 10212424]
- Summers V, Cord MT. Intelligibility of speech in noise at high presentation levels: Effects of hearing loss and frequency region. *J. Acoust. Soc. Am.* 2007; 112:1130–1137. [PubMed: 17672659]
- Takahashi GA, Bacon SP. Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J. Speech Hear. Res.* 1992; 35:1410–1421. [PubMed: 1494284]
- Turner CW, Souza PE, Forget LN. Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* 1995; 97:2568–2576. [PubMed: 7714274]
- Van Tasell DJ, Soli SD, Kirby VM, Widin GP. Speech waveform envelope cues for consonant recognition. *J. Acoust. Soc. Am.* 1987; 82:1152–1161. [PubMed: 3680774]
- Vickers DA, Moore BCJ, Baer T. Effects of lowpass filtering on the intelligibility of speech in quiet for people with and with-out dead regions at high frequencies. *J. Acoust. Soc. Am.* 2001; 110:1164–1175. [PubMed: 11519583]
- Whitmal NA, Poissant SF, Freyman RL, Heifer KS. Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience. *J. Acoust. Soc. Am.* 2007; 122:2376–2388. [PubMed: 17902872]
- Woolf NK, Ryan AF, Bone RC. Neural phase-locking properties in the absence of outer hair cells. *Hear. Res.* 1981; 4:335–346. [PubMed: 7263519]
- 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; 66:1381–1403. [PubMed: 500976]

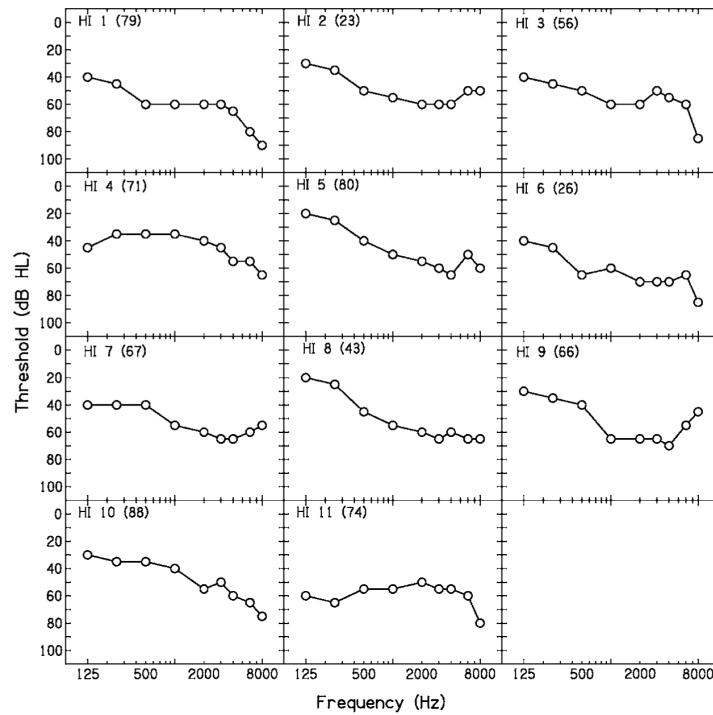


FIG. 1. Air conduction audiometric thresholds of the test ears of the hearing-impaired subjects for experiments one and two. The ages of the subjects (in years) are shown in parentheses.

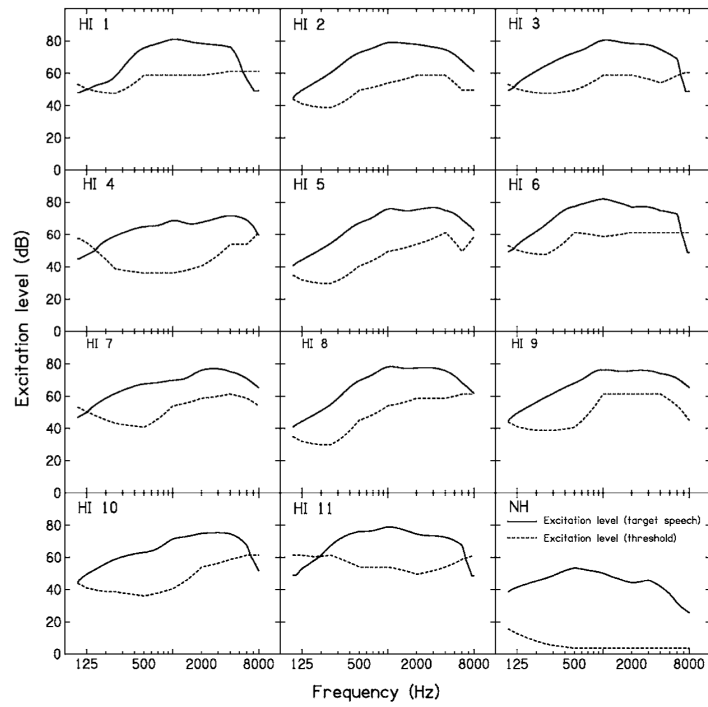


FIG. 2. Excitation levels of the target speech (solid lines) and excitation levels at threshold (dashed lines) for individual hearing-impaired subjects. Excitation levels for normal-hearing subjects are shown for comparison in the bottom-right panel.

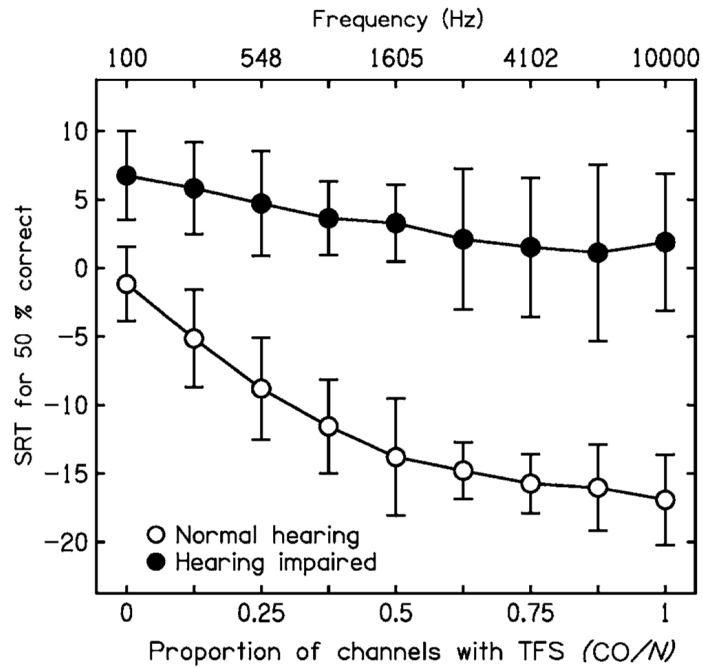


FIG. 3. Mean SRTs for normal-hearing and hearing-impaired subjects, plotted as a function of CO/N . The frequency corresponding to CO/N is shown along the top axis. Error bars show \pm one standard deviation across subjects.

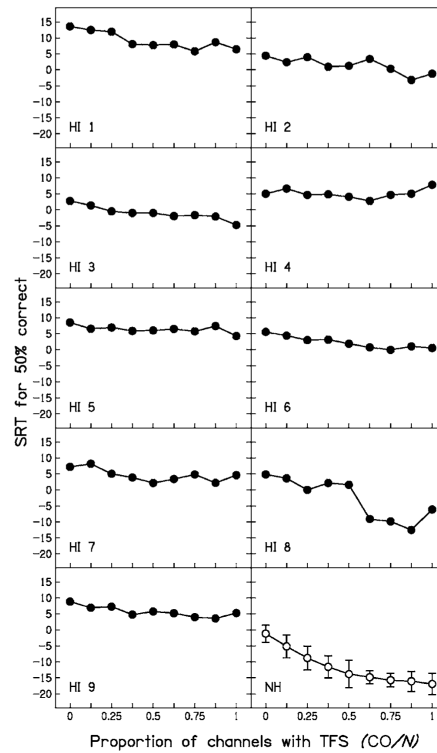


FIG. 4. Individual SRTs for the hearing-impaired subjects, plotted as a function of CO/N . Mean SRTs measured for the normal-hearing subjects are shown in the bottom-right panel for comparison.

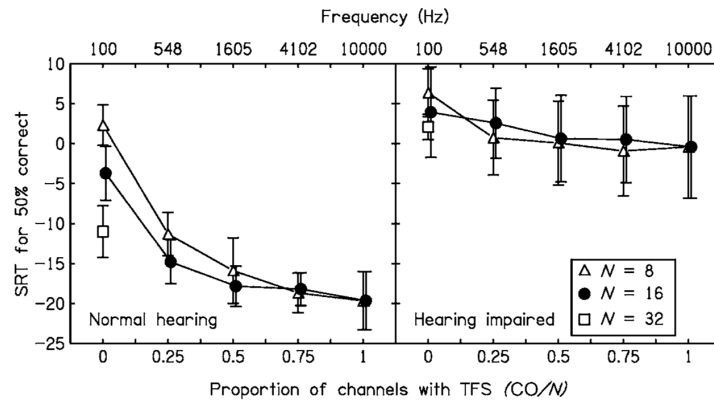


FIG. 5. Mean SRTs for normal-hearing subjects (left) and hearing-impaired subjects (right), plotted as a function of CO/N. Error bars show \pm one standard deviation across subjects.

TABLE I

Differences between mean SRTs measured with different values of CO for normal-hearing subjects in experiment one. Equivalent cut-off frequencies in Hz are also shown. The LSD calculated using Fisher's LSD procedure was 2.0. Differences equal to or above this value are shown in bold.

CO	0	4	8	12	16	20	24	28	32
Freq. (Hz)	100	277	548	965	1605	2590	4102	6427	10,000
0	100	0							
4	277	3.8	0						
8	548	7.4	3.7	0					
12	965	10.2	6.4	2.8	0				
16	1605	12.4	8.7	5.0	2.2	0			
20	2590	13.4	9.7	6.0	3.2	1.0	0		
24	4102	14.3	10.6	6.9	4.2	1.9	0.9	0	
28	6427	14.6	10.9	7.2	4.5	2.2	1.2	0.3	0
32	10,000	15.5	11.8	8.1	5.4	3.1	2.1	1.2	0.9
									0

TABLE II

As Table I, but for hearing-impaired subjects.

CO	0	4	8	12	16	20	24	28	32
Freq. (Hz)	100	277	548	965	1605	1605	4102	6427	10,000
0	100	0							
4	277	0.9	0						
8	548	2.1	1.1	0					
12	965	3.1	2.2	1.1	0				
16	1605	3.5	2.6	1.4	0.4	0			
20	2590	4.7	3.7	2.6	1.5	1.2	0		
24	4102	5.2	4.3	3.2	2.1	1.8	0.6	0	
28	6427	5.6	4.7	3.6	2.5	2.2	1.0	0.4	0
32	10,000	4.9	3.9	2.8	1.8	1.4	0.2	0.4	0.8
									0

TABLE III

As Table I, but for experiment two, which used two values of N . The LSD calculated using Fisher's LSD procedure was 2.3.

CO/N	N=8					N=16				
	0	0.25	0.5	0.75	1	0	0.25	0.5	0.75	1
Freq. (Hz)	100	548	1605	4102	10,000	100	548	1605	4102	10,000
0	100	0				0				
0.25	548	13.7	0			11.1	0			
0.5	1605	18.2	4.6	0		14.1	3.0	0		
0.75	4102	21.0	7.4	2.8	0	14.5	3.4	0.4	0	
1	10,000	22.0	8.3	3.8	1.0	0	15.2	4.1	1.1	0.7

TABLE IV

As Table III, but for hearing-impaired subjects.

CO/N	N=8								N=16							
	0	0.25	0.5	0.75	1	10,000	100	548	100	548	100	548	1605	4102	10,000	
Freq. (Hz)	100	548	1605	4102	10,000	100	548	1605	4102	10,000	100	548	1605	4102	10,000	
0	100	0									0					
0.25	548	5.3	0								1.4	0				
0.5	1605	5.9	0.6	0							3.2	1.8	0			
0.75	4102	7.1	1.8	1.2	0						3.6	2.0	0.1	0		
1	10,000	6.4	1.1	0.5	0.7	0					4.2	2.8	1.0	0.8	0	

TABLE V

Differences between mean SRTs measured for $N=8$ and $N=16$ for each value of CO/N, for normal-hearing and hearing-impaired subjects in experiment two. The LSD calculated using Fisher's LSD procedure was 1.7. Differences equal to or above this value are shown in bold.

CO/N	0	0.25	0.5	0.75	1
Freq. (Hz)	100	548	1605	4102	10,000
Normal hearing	6.0	3.5	2.0	0.5	0.7
Hearing impaired	2.3	1.7	0.4	1.5	0.1