

The 2008 Award of Merit Review

The Role of Temporal Fine Structure Processing in Pitch Perception, Masking, and Speech Perception for Normal-Hearing and Hearing-Impaired People

BRIAN C. J. MOORE

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

Received: 15 July 2008; Accepted: 23 September 2008; Online publication: 15 October 2008

ABSTRACT

Complex broadband sounds are decomposed by the auditory filters into a series of relatively narrowband signals, each of which can be considered as a slowly varying envelope (E) superimposed on a more rapid temporal fine structure (TFS). Both E and TFS information are represented in the timing of neural discharges, although TFS information as defined here depends on phase locking to individual cycles of the stimulus waveform. This paper reviews the role played by TFS in masking, pitch perception, and speech perception and concludes that cues derived from TFS play an important role for all three. TFS may be especially important for the ability to “listen in the dips” of fluctuating background sounds when detecting nonspeech and speech signals. Evidence is reviewed suggesting that cochlear hearing loss reduces the ability to use TFS cues. The perceptual consequences of this, and reasons why it may happen, are discussed.

Keywords: temporal processing, temporal fine structure, phase locking, cochlear hearing loss, speech perception, pitch perception

INTRODUCTION

When a complex broadband sound is analyzed in the cochlea of a normal ear, the result is a series of bandpass-filtered signals, each corresponding to one

position on the basilar membrane. This aspect of auditory analysis is often modeled (crudely) by short-term Fourier analysis, which expresses the signal in terms of the magnitude and phase of its spectral components. Traditionally, the spectral magnitudes have been regarded as of primary importance for perception, although under some conditions, the phases of the components play an important role (Moore 2002).

The bandpass signal at a specific place on the basilar membrane (or the signal produced by bandpass filtering to simulate the waveform at one place on the basilar membrane) can be analyzed using the Hilbert transform to create what is called the “analytic signal” (Bracewell 1986). The analytic signal can be thought of as a vector that rotates as a function of time; the length of the vector at any time represents the magnitude of the envelope of the signal at that time, and the rate of rotation represents the instantaneous frequency of the signal. In other words, the Hilbert transform can be used to decompose the time signal into its envelope (E; the relatively slow variations in amplitude over time) and temporal fine structure (TFS; the rapid oscillations with rate close to the center frequency of the band). This is illustrated in Figure 1, which shows the outputs of bandpass filters centered at 369, 1,499, and 4,803 Hz in response to the sound “en” in “sense”. Each filter was chosen to have a bandwidth of 1 ERB_N , where ERB_N stands for the equivalent rectangular bandwidth of the auditory filter as determined using young normally hearing listeners at moderate sound levels (Glasberg and Moore 1990; Moore 2003). The suffix N denotes normal hearing. The thick lines in Figure 1 show the Hilbert envelopes of the waveforms. Traditionally, the envelope has been regarded as the most

Correspondence to: Brian C. J. Moore · Department of Experimental Psychology · University of Cambridge · Downing Street, Cambridge, CB2 3EB, UK. email: bcjm@cam.ac.uk

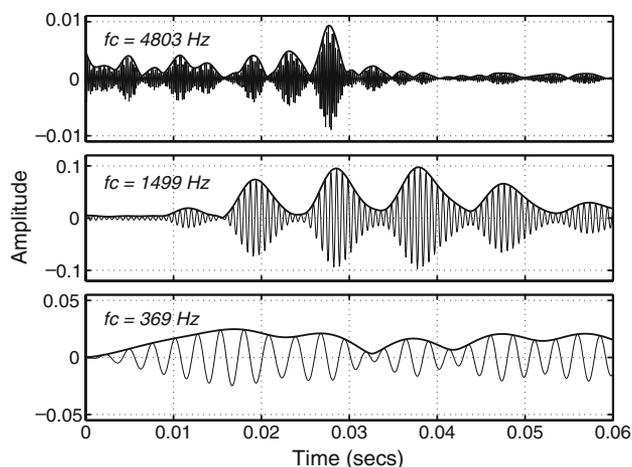


FIG. 1. Waveforms at the outputs of simulated normal auditory filters centered at 369, 1,499, and 4,803 Hz in response to the sound “en” in “sense”. The *thick lines* show the Hilbert envelopes of the waveforms.

important carrier of information, at least for speech signals. However, in this review, I argue for an important role of TFS information.

Both E and TFS information are represented in the timing of neural discharges, although TFS information depends on phase locking to individual cycles of the stimulus waveform (Young and Sachs 1979). In most mammals, phase locking weakens for frequencies above 4–5 kHz, although some useful phase locking information may persist for frequencies up to at least 10 kHz (Heinz et al. 2001). The upper limit of phase locking in humans is not known. Although TFS in the stimulus on the basilar membrane is present up to the highest audible frequencies, this paper is especially concerned with TFS information as represented in the patterns of phase locking in the auditory nerve. This information probably weakens at high frequencies, and so one way of exploring the use of TFS information is to examine changes in performance on various tasks as a function of frequency. Other ways will be described later in this paper.

THE ROLE OF TFS IN PITCH PERCEPTION

Evidence accrued over many years suggests that TFS plays a role in the perception of pitch for both pure and complex tones; for reviews, see Moore (2003) and Plack and Oxenham (2005). For steady pure tones, information from TFS seems to be necessary to account for the way that frequency discrimination varies with frequency (Heinz et al. 2001) and to account for the fact that frequency discrimination for very short tones is better than would be predicted from excitation-pattern (place) models, based on the broadening of the spectrum that occurs with decreas-

ing duration (Moore 1973). For steady complex tones, information from TFS may be important for coding the frequencies of individual resolved partials (Moore et al. 2006b) and also for coding the temporal structure of the waveform evoked on the basilar membrane by unresolved harmonics with rank below about 14 (Moore et al. 2006a; Moore and Moore 2003b). For complex tones containing only harmonics above the 14th, the pitch seems to be determined by E rather than by TFS cues (Moore and Moore 2003b) and the perceived pitch is relatively weak (Houtsma and Smurzynski 1990).

Information from TFS may also play a role in the detection of frequency modulation (FM) at low rates. Moore and Sek (1992, 1994) have shown that a place model based on excitation patterns can account for the detection of FM, or mixtures of FM and amplitude modulation (AM) when the FM rate is medium or high (10 Hz and above). However, when the FM rate is low (5 Hz or less), the model fails to predict the data (Moore and Sek 1995; Sek and Moore 1995). Moore and Sek proposed that, for FM rates below 5 Hz, FM is detected by virtue of the changes in phase locking to the carrier that occur over time. Note that the carrier frequency itself can be rather high (e.g., 4,000 Hz). Moore and Sek suggested that the mechanism for decoding the phase-locking information was “sluggish” and became less effective when the oscillations in frequency were rapid. Hence, it played little role for high modulation rates. This sluggishness may be similar to that observed for binaural processing of interaural phase differences or interaural correlation, which also depends on sensitivity to TFS (Blauert 1972; Grantham and Wightman 1978, 1979); however, one recent study has shown that rapid changes in interaural timing can be heard for certain complex stimuli (Siveke et al. 2008).

In the experiments of Moore and Sek on FM detection, subjects could always fall back on the use of place cues when TFS cues became less effective (at high modulation rates). However, this does not imply that TFS cues become completely unusable for modulation rates above 5 Hz. When the amount of modulation is well above the detection threshold, TFS cues may play a role for higher modulation rates, especially when place cues are of limited benefit.

MASKING AND THE ROLE OF TFS IN DIP LISTENING

It is often easier to detect a signal in a fluctuating background sound than in a steady background sound, especially when the frequency of the signal is different from the center frequency of the masker. This effect has usually been ascribed to the ability to

“listen in the dips” of the fluctuating background sound. I describe next an experiment that supports the idea that dip listening depends partly on the use of TFS information.

Moore and Glasberg (1987) measured the threshold for detecting a sinusoidal signal in a masker that consisted of either a single sinusoid or a pair of equal-amplitude sinusoids that produced beats at a rate depending on their frequency separation. The overall level of all maskers was 80 dB sound pressure level and the signal frequency was 1.8 times the masker center frequency, f_m , which was 250, 1,000, 3,000, or 5,275 Hz. For the beating masker, the beat rate was 4, 8, 16, 32, or 64 Hz. The mean results for three subjects are shown in Figure 2. For $f_m=1,000$ Hz, the threshold for the signal in the beating masker was considerably lower than the threshold in the steady masker. The difference (masking release) was largest (mean ≈ 25 dB) for the 4-Hz beat rate and decreased progressively as the beat rate was increased to 64 Hz (mean ≈ 10 dB). The pattern of results was similar for the masker centered at 250 Hz, although the masking release was smaller. However, for the highest masker center frequency, for which both the signal and the masker frequency fell in the range where phase locking is weak or absent, the masking release was smaller, at about 10 dB, and did not decrease markedly with increasing beat rate of the masker. The results for $f_m=3,000$ Hz were intermediate in form between those for $f_m=1,000$ and 5,275 Hz.

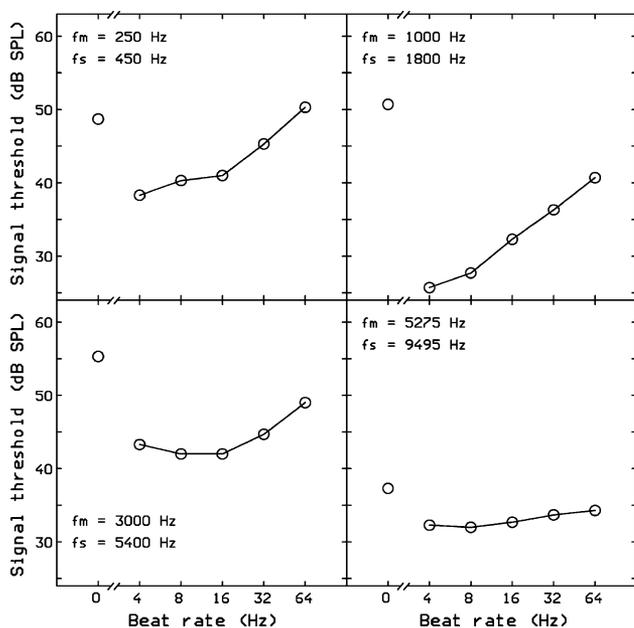


FIG. 2. Thresholds for detecting a sinusoidal signal in a masker consisting of a single sinusoid (beat rate=0) or a pair of sinusoids with beat rate as indicated. Data from Moore and Glasberg (1987).

These results are consistent with the idea that TFS provides a cue that allows effective dip listening when the masker and signal frequencies fall in the range where phase locking is relatively precise. The decrease in masking release with increasing masker beat rate is consistent with the idea that the mechanism that “decodes” TFS information is sluggish and is less effective when there are rapid changes in TFS. The amount of sluggishness may depend somewhat on center frequency, being greater for low frequencies; it may be that the period needs to be reasonably stable over a certain number of stimulus cycles for TFS to be extracted effectively. This could explain the reduced masking release for the masker centered at 250 Hz. When the masker and signal frequencies are too high to support precise phase locking, some masking release still occurs, but it is smaller and depends only slightly on the masker beat rate. Presumably, some other mechanism leads to masking release in this case, for example, comparison of short-term levels across frequency channels or a shift in the position of the excitation pattern as the masker envelope passes through a minimum; this mechanism appears to be only slightly sluggish. The results of other masking experiments have also been interpreted as indicating a role of TFS in dip listening (Schooneveldt and Moore 1987).

THE ROLE OF TFS IN SPEECH PERCEPTION

Several researchers have investigated the role of E and TFS cues in speech perception by processing speech sounds in such a way that they contain mainly E or TFS cues. This has been done using different forms of vocoders (Dudley 1939). The speech is filtered into a number of contiguous frequency bands. To preserve E cues, the envelope is extracted from the signal at the output of each band, and the envelope is used to modulate the amplitude of a noise band (noise vocoder) or a sinusoid (tone vocoder) centered at the frequency of the band from which the envelope was derived. The modulated carriers are then combined (usually after a second stage of filtering to restrict the spectrum of the modulated carriers to the original bandwidths). Speech processed in this way will be referred to as “E-speech”. Experiments using E-speech have shown that only a few bands are required to give good intelligibility for speech in quiet (Shannon et al. 1995), although more bands are required when background sounds are present (Qin and Oxenham 2003; Stone and Moore 2003); for a review, see Lorenzi and Moore (2008). Overall, the results suggest that E cues are sufficient to give good intelligibility for speech in quiet, but they are not sufficient when background sounds are present, especially when the background is

fluctuating. This may be the case because E cues alone are not sufficient to allow the perceptual segregation of mixtures of sounds, especially when information about the sounds is conveyed only by envelope fluctuations in different frequency bands.

Processing speech to remove E cues but preserve TFS cues is more difficult. In attempts to do this, each bandpass signal is divided by the envelope magnitude (usually the Hilbert envelope). As a result, each band signal becomes like an FM sinusoidal carrier, with a constant amplitude. The processed band signals are scaled in long-term amplitude so that they have the same root-mean-square amplitude as the original band signals and are then combined. Speech processed in this way will be referred to as “TFS-speech”. It should be emphasized that, although such processing preserves TFS information to some extent, the TFS information is nevertheless distorted and different from the information in the unprocessed speech.

A potential problem with TFS-speech was identified by Ghitza (2001). He showed that, although E cues are physically removed from TFS-speech, they are reconstructed at the outputs of the auditory filters (called cochlear filters by Ghitza) and may therefore contribute to intelligibility, especially when only a few broad analysis bands are used in the processing (Zeng et al. 2004). Thus, the intelligibility of TFS-speech may be influenced by reconstructed E cues. However, Gilbert and Lorenzi (2006) presented evidence indicating that the reconstructed E cues supported only minimal speech identification when the bandwidth of the analysis filters was less than 4 ERB_N, or equivalently, when the number of analysis bands was equal to or greater than eight over the frequency range 0.08 to 8.02 kHz. Using stimuli processed using a large number of bands, it has been shown that, after training, normal-hearing listeners achieve high levels of intelligibility for TFS-speech in quiet (Gilbert and Lorenzi 2006; Lorenzi et al. 2006). Thus, TFS cues do seem to convey information for intelligibility. However, it may be that reconstructed envelope cues make a contribution to the intelligibility of TFS-speech, even though the envelope cues alone are not sufficient to give good intelligibility. The fact that learning is required to achieve high intelligibility with TFS-speech may indicate that the auditory system normally uses TFS cues in conjunction with envelope cues; when envelope cues are minimal, TFS information may be difficult to interpret. Alternatively, the learning may reflect the fact that TFS cues are distorted in the TFS-speech (relative to unprocessed speech), and it may require some training to overcome the effects of the distortion.

Hopkins et al. (2008) adopted a different approach to assess the use of TFS in speech perception. The approach was intended to avoid some of the problems discussed above, where artifacts or side effects of the

signal processing may have an influence. They measured performance as a function of the number, J , of one-ERB_N-wide analysis bands that contained intact TFS and E information; the other bands were noise or tone vocoded, so that they conveyed only E information. Speech reception thresholds (SRTs; the speech-to-background ratio required for 50% correct key words in sentences) were measured for signals that were unprocessed for bands up to and including band number J and were vocoded for higher-frequency bands. The value of J was varied from 0 to 32 in steps of 4. A competing-talker background was used, because, as described earlier, E information does not seem to be sufficient to allow good intelligibility when a fluctuating background is used. The mean results for nine normal-hearing subjects when a noise vocoder was used for bands $J+1$ to 32 are shown by the open circles in Figure 3. The SRT declined considerably (by about 15 dB) as the value of J was increased from 0 to 32, i.e., as more TFS information was added. These results suggest that TFS information plays a considerable role in the ability to identify speech in a fluctuating background.

This section has emphasized the role of TFS information for speech perception in fluctuating background sounds. However, it should be noted that TFS information also plays a role in speaker identification and the understanding of tonal languages, in which differences in fundamental frequency (F0) or changes in F0 over time affect word meanings (Zeng et al. 2005).

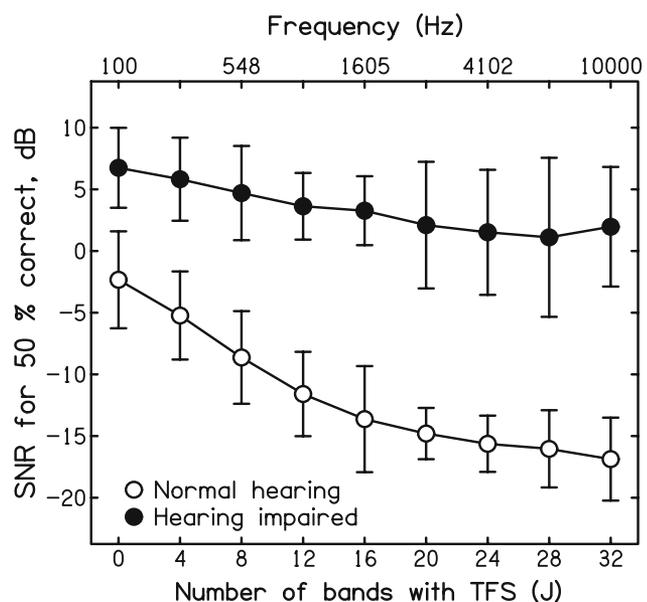


FIG. 3. Mean SRTs for normal-hearing (open circles) and hearing-impaired (filled circles) subjects, plotted as a function of the number of bands, J , containing TFS information. The frequency corresponding to highest band with TFS information is shown along the top axis. Error bars show \pm one standard deviation across subjects. Adapted from Hopkins et al. (2008).

THE EFFECT OF HEARING LOSS ON THE ABILITY TO USE TFS INFORMATION

Psychoacoustic studies

Experimental evidence suggests that hearing-impaired listeners have a reduced ability to use TFS information for: (1) detecting FM at low rates (Lacher-Fougère and Demany 1998; Moore and Skrodzka 2002), (2) lateralization based on interaural phase differences (Lacher-Fougère and Demany 2005), and (3) discrimination of the (missing) F0 of complex tones (Moore and Moore 2003a).

Hopkins and Moore (2007) directly assessed the ability of subjects with moderate cochlear hearing loss to use TFS information. They measured the ability to discriminate a harmonic complex tone, with $F_0=100, 200, \text{ or } 400$ Hz, from a similar tone in which all components had been shifted up by the same amount in Hertz, ΔF . To reduce cues relating to differences in the excitation patterns of the two tones, the tones contained many components, and they were passed through a fixed bandpass filter centered on the upper (unresolved) harmonics. To prevent components outside the passband of the filter from being audible, a background noise was added. In the presence of this noise, the differences in excitation patterns between the harmonic and frequency-shifted tones were very small when the bandpass filter was centered on the 11th harmonic. People with normal hearing perceive the shifted tone as having a higher pitch than the harmonic tone (de Boer 1956; Moore and Moore 2003b). The envelope repetition rate of the two sounds is the same (equal to F_0), so the difference in pitch is assumed to occur because of a difference in the TFS of the two sounds (Moore and Moore 2003b; Schouten et al. 1962).

Trained normally hearing subjects were able to perform this task well, presumably reflecting the ability to discriminate changes in the TFS of the harmonic and frequency-shifted tones. The smallest detectable frequency shift (corresponding to a detectability index $d'=1$) was typically about $0.05F_0$. Even untrained normally hearing subjects achieved thresholds of $0.2F_0$ or better (Moore and Sek 2008). However, subjects with moderate cochlear hearing loss generally performed very poorly. For most subjects and F_0 s, performance was not significantly above chance even for the maximum frequency shift of $0.5F_0$. Above-chance performance occurred only when there was little or no hearing loss at the center frequency of the filter passband. The results suggest that moderate cochlear hearing loss results in a reduced ability, or no ability, to discriminate harmonic from frequency-shifted tones based on TFS.

This conclusion applies only to relatively high center frequencies, since the lowest center frequency

tested by Hopkins and Moore was 1,100 Hz. Some hearing-impaired subjects can process TFS for lower center frequencies, since, for example, they can recognize melodies played as a pattern of binaural pitches (Santurette and Dau 2006). However, many hearing-impaired subjects do show a poorer than normal ability to process binaural TFS information for frequencies below 1,000 Hz (Lacher-Fougère and Demany 2005; Santurette and Dau 2006).

Speech perception studies

People with cochlear hearing loss usually have difficulty in understanding speech when background sounds are present, especially when the background is fluctuating (Duquesnoy 1983). Recent evidence supports the idea that the difficulty stems partly from a reduced ability to process TFS cues. Lorenzi et al. (2006) measured identification scores for unprocessed, E, and TFS-speech in quiet for three groups of listeners: young with normal hearing and young and elderly with moderate “flat” hearing loss. After training, normal-hearing listeners scored perfectly with unprocessed speech and about 90% correct with E- and TFS-speech. Both young and elderly listeners with hearing loss performed almost as well as normal with unprocessed and E-speech but performed very poorly with TFS-speech, indicating a greatly reduced ability to use TFS. For the younger hearing-impaired group, scores for TFS-speech were highly correlated ($r=0.83$) with the ability to take advantage of temporal dips in a background noise when identifying unprocessed speech. These results support the idea that the ability to use TFS is important for listening in the dips.

Lorenzi et al. (2008) measured the identification of E- and TFS-speech in quiet for normal-hearing listeners and listeners with high-frequency mild-to-severe hearing loss and normal (<20 dB HL) audiometric thresholds below 2 kHz. The stimuli were lowpass filtered at 1.5 kHz to restrict their spectrum to the region where audiometric thresholds were normal, and a condition with complementary highpass noise was included, to restrict “off-frequency listening”. Only a few of the hearing-impaired listeners were able to score above chance (6.25%) for the TFS-speech, whereas normal-hearing listeners achieved scores of 20–50% (relatively low scores would be expected given the limited audible frequency range of the stimuli). The results indicate that, for listeners with cochlear hearing loss, deficits in the ability to use TFS cues in speech can occur even when audiometric thresholds are within the normal range.

In the experiment of Hopkins et al. (2008), listeners with moderate cochlear hearing loss were also tested. The mean results are shown by the filled circles in Figure 3. The hearing-impaired listeners performed

more poorly than the normal-hearing listeners in all conditions, and the difference in performance was greater when J was large. The improvement in performance going from completely vocoded stimuli ($J=0$) to completely unprocessed stimuli ($J=32$) was only about 5 dB for the hearing-impaired listeners, compared to about 15 dB for the normal-hearing listeners. For the former are, the improvement varied across listeners, with some listeners showing near-normal benefit and others showing no benefit at all. The reasons for the individual differences are not at present clear. Hopkins et al. (2008) reported that the benefit gained from the addition of TFS information was not correlated with the average audiometric threshold over the range 250 to 4,000 Hz.

Current cochlear implant systems convey mainly envelope information in different frequency bands, and this may partly account for the relatively poor ability of cochlear implantees to understand speech when background sounds are present. Nie et al. (2005) evaluated the potential contribution of TFS information to speech recognition in noise via acoustic simulations of a cochlear implant. They transformed the rapidly-varying TFS into a slowly varying FM signal which was applied to the carrier in each band (which was already amplitude modulated by the speech envelope in that band). They found that, for sentence recognition in the presence of a competing voice, adding this FM signal improved performance by as much as 71 percentage points. This illustrates the potentially large benefit of providing TFS information in a cochlear implant.

POSSIBLE REASONS FOR THE EFFECT OF COCHLEAR HEARING LOSS ON SENSITIVITY TO TFS

There are several possible reasons why cochlear hearing loss might lead to a reduced ability to process TFS information. Such a loss may lead to:

1. Reduced precision of phase locking. Physiological studies disagree about whether or not this occurs for pure-tone stimuli (Harrison and Evans 1979; Miller et al. 1997; Woolf et al. 1981). For complex sounds, such as synthesized vowels, noise-induced hearing loss can lead to reduced synchrony capture (phase locking to the formant peaks), possibly as a result of diminished two-tone suppression (Miller et al. 1997)
2. A change in the relative phase of response at different points along the basilar membrane (Ruggero 1994). This would affect mechanisms for decoding TFS based on correlation of the outputs of adjacent places (Carney et al. 2002; Deng and Geisler 1987; Loeb et al. 1983; Shamma and Klein 2000; Shamma 1985)
3. More complex and more rapidly varying TFS resulting from broader auditory filters, which might make it more difficult for central mechanisms to decode the TFS information
4. A mismatch between TFS information and the place on the basilar membrane that would “normally” respond to that information. TFS information may be decoded on a place-specific basis (Huss and Moore 2005; Moore 1982; Oxenham et al. 2004; Srulovicz and Goldstein 1983). For example, the TFS at a 1-kHz rate may be decoded best by the central neurons that are tuned (in a normal ear) close to 1 kHz. Hearing loss may produce a shift in frequency-place mapping (Liberman and Dodds 1984; Sellick et al. 1982) and disrupt the decoding process. Note that this explanation depends on the assumption that the central decoding mechanism is, to some extent, “hard wired”
5. There may be central changes that occur following cochlear hearing loss, such as loss of inhibition, and such changes might disrupt the mechanisms for decoding the TFS

ACKNOWLEDGEMENTS

The work of the author was supported by the MRC, Deafness Research UK, and RNID. I thank Kathryn Hopkins, Christian Lorenzi, and Michael Stone for their collaboration in some of the work reported here. I thank Kathryn Hopkins for helpful comments on an earlier version of this paper. I also thank Associate Editor Fan-Gang Zeng and three anonymous reviewers for their helpful comments.

REFERENCES

- BLAUERT J. On the lag of lateralization caused by interaural time and intensity differences. *Audiology* 11:265–270, 1972.
- BRACEWELL RN. The Fourier transform and its applications. New York, McGraw-Hill, 1986.
- CARNEY LH, HEINZ MG, EVILSIZER ME, GILKEY RH, COLBURN HS. Auditory phase opponency: A temporal model for masked detection at low frequencies. *Acta Acust. Acust.* 88:334–346, 2002.
- DE BOER E. Pitch of inharmonic signals. *Nature* 178:535–536, 1956.
- DENG L, GEISLER CD. A composite auditory model for processing speech sounds. *J. Acoust. Soc. Am.* 82:2001–2012, 1987.
- DUDLEY H. Remaking speech. *J. Acoust. Soc. Am.* 11:169–177, 1939.
- DUQUESNOY AJ. Effect of a single interfering noise or speech source on the binaural sentence intelligibility of aged persons. *J. Acoust. Soc. Am.* 74:739–743, 1983.
- GHIZZA O. On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception. *J. Acoust. Soc. Am.* 110:1628–1640, 2001.
- GILBERT G, LORENZI C. The ability of listeners to use recovered envelope cues from speech fine structure. *J. Acoust. Soc. Am.* 119:2438–2444, 2006.
- GLASBERG BR, MOORE BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear Res.* 47:103–138, 1990.

- GRANTHAM DW, WIGHTMAN FL. Detectability of varying interaural temporal differences. *J. Acoust. Soc. Am.* 63:511–523, 1978.
- GRANTHAM DW, WIGHTMAN FL. Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. *J. Acoust. Soc. Am.* 65:1509–1517, 1979.
- HARRISON RV, EVANS EF. Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig. *Arch. Otolaryngol.* 224:71–78, 1979.
- HEINZ MG, COLBURN HS, CARNEY LH. Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. *Neur. Comput.* 13:2273–2316, 2001.
- HOPKINS K, MOORE BCJ. Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J. Acoust. Soc. Am.* 122:1055–1068, 2007.
- HOPKINS K, MOORE BCJ, STONE MA. Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J. Acoust. Soc. Am.* 123:1140–1153, 2008.
- HOUTSMA AJM, SMURZYNSKI J. Pitch identification and discrimination for complex tones with many harmonics. *J. Acoust. Soc. Am.* 87:304–310, 1990.
- HUSS M, MOORE BCJ. Dead regions and pitch perception. *J. Acoust. Soc. Am.* 117:3841–3852, 2005.
- LACHER-FOUGÈRE S, DEMANY L. Modulation detection by normal and hearing-impaired listeners. *Audiology* 37:109–121, 1998.
- LACHER-FOUGÈRE S, DEMANY L. Consequences of cochlear damage for the detection of interaural phase differences. *J. Acoust. Soc. Am.* 118:2519–2526, 2005.
- LIBERMAN MC, DODDS LW. Single neuron labeling and chronic cochlea pathology. III. Stereocilia damage and alterations in threshold tuning curves. *Hear Res.* 16:54–74, 1984.
- LOEB GE, WHITE MW, MERZENICH MM. Spatial cross correlation: A proposed mechanism for acoustic pitch perception. *Biol. Cybern.* 47:149–163, 1983.
- LORENZI C, DEBRUILLE L, GARNIER S, FLEURIOT P, MOORE BCJ. Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. *J. Acoust. Soc. Am.* (in press), 2008.
- LORENZI C, GILBERT G, CARN C, 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.* 103:18866–18869, 2006.
- LORENZI C, MOORE BCJ. Role of temporal envelope and fine structure cues in speech perception: A review. In: Dau T, Buchholz JM, Harte JM, Christiansen TU (eds) *Auditory Signal Processing in Hearing-Impaired Listeners*. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). Denmark, Centertryk A/S, pp. 263–272, 2008.
- MILLER RL, SCHILLING JR, FRANCK KR, YOUNG ED. Effects of acoustic trauma on the representation of the vowel /ε/ in cat auditory nerve fibers. *J. Acoust. Soc. Am.* 101:3602–3616, 1997.
- MOORE BCJ. Frequency difference limens for short-duration tones. *J. Acoust. Soc. Am.* 54:610–619, 1973.
- MOORE BCJ. *An introduction to the psychology of hearing*. 2nd edn. London, Academic, 1982.
- MOORE BCJ. Interference effects and phase sensitivity in hearing. *Phil. Trans. R. Soc. Lond. A* 360:833–858, 2002.
- MOORE BCJ. *An introduction to the psychology of hearing*. 5th edn. San Diego, Academic, 2003.
- MOORE BCJ, GLASBERG BR. Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes. *J. Acoust. Soc. Am.* 82:69–79, 1987.
- MOORE BCJ, GLASBERG BR, FLANAGAN HJ, ADAMS J. Frequency discrimination of complex tones; assessing the role of component resolvability and temporal fine structure. *J. Acoust. Soc. Am.* 119:480–490, 2006a.
- MOORE BCJ, GLASBERG BR, LOW K-E, COPE T, COPE W. Effects of level and frequency on the audibility of partials in inharmonic complex tones. *J. Acoust. Soc. Am.* 120:934–944, 2006b.
- 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.* 182:153–163, 2003a.
- MOORE BCJ, SEK A. Detection of combined frequency and amplitude modulation. *J. Acoust. Soc. Am.* 92:3119–3131, 1992.
- MOORE BCJ, SEK A. Effects of carrier frequency and background noise on the detection of mixed modulation. *J. Acoust. Soc. Am.* 96:741–751, 1994.
- MOORE BCJ, SEK A. Effects of carrier frequency, modulation rate and modulation waveform on the detection of modulation and the discrimination of modulation type (AM vs FM). *J. Acoust. Soc. Am.* 97:2468–2478, 1995.
- MOORE BCJ, SEK A. Development of a fast method for determining sensitivity to temporal fine structure. *Int J. Audiol.* (in press), 2008.
- 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.* 111:327–335, 2002.
- MOORE GA, MOORE BCJ. Perception of the low pitch of frequency-shifted complexes. *J. Acoust. Soc. Am.* 113:977–985, 2003b.
- NIE K, STICKNEY G, ZENG FG. Encoding frequency modulation to improve cochlear implant performance in noise. *IEEE Trans. Biomed. Eng.* 52:64–73, 2005.
- OXENHAM AJ, BERNSTEIN JG, PENAGOS H. Correct tonotopic representation is necessary for complex pitch perception. *Proc. Natl. Acad. Sci. U.S.A.* 101:1421–1425, 2004.
- PLACK CJ, OXENHAM AJ. The psychophysics of pitch. In: Plack CJ, Oxenham AJ, Fay RR, Popper AN (eds) *Pitch perception*. New York, Springer, pp. 7–55, 2005.
- QIN MK, OXENHAM AJ. Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *J. Acoust. Soc. Am.* 114:446–454, 2003.
- RUGGERO MA. Cochlear delays and traveling waves: Comments on ‘Experimental look at cochlear mechanics’. *Audiology* 33:131–142, 1994.
- SANTURETTE S, DAU T. Binaural pitch perception in normal-hearing and hearing-impaired listeners. *Hear Res.* 223:29–47, 2006.
- SCHOONEVELDT GP, MOORE BCJ. Comodulation masking release (CMR): effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band. *J. Acoust. Soc. Am.* 82:1944–1956, 1987.
- SCHOUTEN JF, RITSMA RJ, CARDOZO BL. Pitch of the residue. *J. Acoust. Soc. Am.* 34:1418–1424, 1962.
- SEK A, MOORE BCJ. Frequency discrimination as a function of frequency, measured in several ways. *J. Acoust. Soc. Am.* 97:2479–2486, 1995.
- SELICK PM, PATUZZI R, JOHNSTONE BM. Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *J. Acoust. Soc. Am.* 72:131–141, 1982.
- SHAMMA S, KLEIN D. The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. *J. Acoust. Soc. Am.* 107:2631–2644, 2000.
- 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.* 78:1622–1632, 1985.
- SHANNON RV, ZENG F-G, KAMATH V, WYGONSKI J, EKELID M. Speech recognition with primarily temporal cues. *Science* 270:303–304, 1995.
- SIVEKE I, EWERT SD, GROTHE B, WIEGREBE L. Psychophysical and physiological evidence for fast binaural processing. *J. Neurosci.* 28:2043–2052, 2008.

- SRULOVICZ P, GOLDSTEIN JL. A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. *J. Acoust. Soc. Am.* 73:1266–1276, 1983.
- 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.* 114:1023–1034, 2003.
- WOOLF NK, RYAN AF, BONE RC. Neural phase-locking properties in the absence of outer hair cells. *Hear Res.* 4:335–346, 1981.
- YOUNG ED, SACHS MB. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibres. *J. Acoust. Soc. Am.* 66:1381–1403, 1979.
- ZENG FG, NIE K, LIU S, STICKNEY G, DEL RIO E, KONG YY, CHEN H. On the dichotomy in auditory perception between temporal envelope and fine structure cues (L). *J. Acoust. Soc. Am.* 116:1351–1354, 2004.
- ZENG FG, NIE K, STICKNEY GS, KONG YY, VONGPHOE M, BHARGAVA A, WEI C, CAO K. Speech recognition with amplitude and frequency modulations. *Proc. Nat. Acad. Sci. U.S.A.* 102:2293–2298, 2005.