The roles of temporal envelope and fine structure information in auditory perception

Brian C. J. Moore*

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

Abstract: Within the cochlea, broadband sounds like speech and music are filtered into a series of narrowband signals, each with a relatively slowly varying envelope (ENV) imposed on a rapidly oscillating carrier (the temporal fine structure, TFS). Information about ENV and TFS is conveyed in the timing and short-term rate of action potentials in the auditory nerve. This paper describes the role of ENV and TFS information in pitch perception, binaural processing, and the perception of speech in the presence of background sounds. The paper also describes the effects of hearing loss and age on the processing of TFS and ENV information. The monaural and binaural processing of TFS information is adversely affected by both hearing loss and increasing age. The monaural processing of ENV information is little affected by hearing loss or by increasing age. The binaural processing of ENV information deteriorates somewhat with increasing age but is not markedly affected by hearing loss. The reduced TFS processing abilities found for older/hearing-impaired subjects may partially account for the difficulties that such subjects experience in complex listening situations.

Keywords: Temporal envelope, Temporal fine structure, Pitch, Sound localisation, Binaural processing, Age, Hearing loss

PACS number: 43.66.Hg, 43.66.Pn, 43.66.Sr, 43.71.Ky, 43.71.Lz, 43.71.Rt

[doi:10.1250/ast.40.61]

1. INTRODUCTION

Within a normal healthy cochlea, broadband sounds like speech and music are decomposed into narrowband signals, each of which can be considered as a relatively slowly varying envelope (ENV) imposed on a rapidly oscillating carrier (the temporal fine structure, TFS). This is illustrated in Fig. 1. The thin traces show the outputs of bandpass filters centred at 369, 1,499 and 4,803 Hz in response to the sound “en” in “sense.” The filters had bandwidths similar to those of the “auditory filters” in a normal ear, as determined in masking experiments [1,2]. The thick lines in Fig. 1 show the envelopes of the waveforms. The envelope has been regarded as the most important carrier of information, at least for speech signals [3], and cochlear implants allow good understanding of speech in quiet based solely on ENV cues [4]. However, for speech in the presence of background sounds, cochlear implant users perform poorly [5], and this may indicate an important role for TFS information.

Moore [6] proposed a distinction between the physical ENV and TFS of the input signal (ENV and TFS), the ENV and TFS at a given place on the basilar membrane (BM) within the cochlea (ENV and TFS), and the neural representation of ENV and TFS (ENV and TFS). ENV is conveyed in the short-term rate of action potentials in the auditory nerve [7], while TFS is conveyed by the detailed timing of the action potentials [7–9]. Specifically, the action potentials tend to occur at a specific phase of the TFS, a property called phase locking. Phase locking is robust at low frequencies, but becomes less precise for frequencies above about 1,000 Hz and it disappears altogether at very high frequencies [10]. The upper limit in humans is not known precisely, but is above 3 kHz [11] and may be much higher.

The physical ENV of a broadband signal is not well defined, and different methods of ENV extraction can lead to very different outcomes. For narrowband signals, different methods of ENV extraction lead to very similar results.
where possible on studies that separate the effects of hearing loss and age (Sects. 6–8).

2. THE PITCH PERCEPTION OF SINUSOIDS FOR NORMAL-HEARING SUBJECTS

The ability to judge the pitch of a sinusoid and to detect changes in its frequency may depend on two mechanisms [12]. The first, called the place mechanism, is based on the filtering that occurs on the BM. A sinusoid produces a pattern of excitation with a distinct peak. The pitch may be based on the position of the peak [13] or on the position of the low-frequency edge of the excitation pattern [14].

The second mechanism, called the temporal mechanism, is based on TFS$_n$. It is assumed that the pitch of a sinusoid is extracted from the patterns of TFS$_n$ information, either within neurons [12,15,16] or from comparison or cross-correlation of the patterns of TFS$_n$ in neurons contacting adjacent regions of the BM [17–19].

The role of TFS$_n$ information in pitch perception is often assessed from the patterns of performance in various tasks as a function of frequency; since phase locking weakens at high frequencies, while the sharpness of excitation patterns on the BM does not worsen, a worsening in performance at high frequencies is taken as evidence for a role of TFS$_n$ information at lower frequencies.

2.1. The Perception of Musical Intervals

Two successive sinusoids with a frequency ratio of 2:1 sound similar. The musical interval between them is called one octave. In Western musical notation, two tones that are separated by one octave are given the same name. For example, sinusoids with frequencies 220 Hz and 440 Hz are called A3 and A4, respectively. This has led to the idea that musical pitch has two dimensions [20]. One increases monotonically with increasing frequency and is called “tone height.” The other is related to pitch class (that is, name note) and is called “tone chroma” [21].

If two sinusoids with frequencies $f_1$ and $f_2$ are presented in alternation and the listener is asked to adjust $f_2$ so that it sounds an octave higher than $f_1$, $f_2$ is usually adjusted to be roughly twice $f_1$. However, if $f_1$ lies above 2,500 Hz, octave matches become very erratic [22], suggesting that the octave is not clearly perceived when $f_2$ lies above 5,000 Hz. Matches to other musical intervals (such as the fifth, corresponding to a frequency ratio of 3:2) also become erratic when one of the tones has a frequency above 5,000 Hz [23].

Consistent with these findings, a sequence of sinusoids with frequencies above 5,000 Hz does not produce a clear sense of melody for most listeners [24]. Also, people with absolute pitch (the ability to assign name notes without reference to other notes) are very poor at naming sinusoids with frequencies above 4,000–5,000 Hz [25]. These results are consistent with the idea that the sense of tone chroma depends on the use of TFS$_n$ information, and that the perception of tone chroma becomes weak or absent for frequencies where TFS$_n$ information is impoverished, above about 5,000 Hz.

2.2. The Frequency Discrimination of Sinusoids

The smallest detectable difference in frequency between successive sinusoids is called the difference limen for frequency, DLF. A change in frequency produces a shift in the excitation pattern along the BM, and this shift may provide the basis for detection of the change in frequency. This place mechanism is illustrated in Fig. 2.

According to place theory, the DLF may depend on the detection of a change in excitation level at a single point on the excitation pattern [14,26,27] or on the combination of information from changes in excitation level over the whole audible part of the excitation pattern [28]. In either case, the DLF should depend on the slope of the excitation
pattern and on the smallest detectable change in excitation level. For example, Zwicker [26,27] proposed that a change in frequency can be detected whenever the excitation level at some point on the excitation pattern changes by more than about 1 dB. The change in excitation level is greatest on the steeply sloping low-frequency side of the excitation pattern. The steepness of the low-frequency side is roughly independent of centre frequency when expressed in terms of the equivalent rectangular bandwidth of the auditory filter for people with normal hearing, ERB\textsubscript{N} [1,2]. The slope is about 27 dB/ERB\textsubscript{N}. Thus, Zwicker’s model predicts that the DLF in hertz at a given centre frequency should be approximately $\text{ERB}_N/27$.

Data from many studies show that DLFs do not conform to this prediction [29–33]. If the DLF is expressed as a proportion of $\text{ERB}_N$, the result is not invariant with centre frequency, but tends to increase with increasing frequency above 500 Hz [32,33]. An example is shown by the solid line in Fig. 3. This led to the suggestion that DLFs are determined by the use of TFS at low frequencies, with a transition to a place mechanism at high frequencies.

At high centre frequencies, auditory filters and excitation patterns remain roughly equally sharp with increasing centre frequency, or even increase slightly in sharpness with increasing centre frequency, when plotted on a log-frequency scale [2,34,35]. Also, the ability to detect changes in excitation level is roughly invariant with centre frequency, although it does worsen slightly at very high frequencies, especially for medium levels [36,37]. If DLFs at high frequencies are based purely on a place mechanism, then, over the range where the place mechanism is dominant, DLFs expressed as a proportion of centre frequency ($\text{DLF}/f$) should be approximately constant. If there is a transition from a temporal to a place mechanism, then $\text{DLF}/f$ should increase with increasing $f$ above about 1,000 Hz (as phase locking weakens), but eventually should reach a plateau when the place mechanism has taken over fully from the temporal mechanism.

To test this prediction, Moore and Ernst [38] measured DLFs over a wide frequency range, including frequencies well above 8 kHz. The subjects were selected to have audiometric thresholds better than a hearing level of 20 dB (described as 20 dB HL, hereafter) for frequencies up to 14 kHz. The task was designed to be easy to learn and not to require naming of the direction of a pitch change, which is difficult for some subjects [39]. In one interval of a trial (selected randomly), there were four successive 500-ms bursts of tone A, with a fixed frequency, $f$. In the other interval, tones A and B alternated, giving the pattern ABAB. Tone B had a frequency that was higher than that of tone A by $\Delta f$ Hz. The task of the subject was to choose the interval in which the sound changed across the four tone bursts within an interval. To make it difficult for subjects to use loudness cues to detect the frequency changes, the stimuli were presented via Etymotic Research ER2 earphones, which are designed to have a “flat” response at the eardrum, and the level of every tone was varied randomly over a range of $\pm 4$ dB (uniform distribution) around the mean level. For each frequency, the DLF was estimated for a mean level of 70 dB SPL and for a mean sensation level (SL) of 20 dB (20 dB SL, hereafter).

Figure 4 shows the geometric mean values of $\text{DLF}/f$ across subjects. Circles and squares show values obtained at 20 dB SL and 70 dB SPL, respectively. Diamonds show
geometric means for the two levels. DLF/f increased with increasing frequency up to 8 or 10 kHz and then flattened off. There was no significant difference between the values of DLF/f for centre frequencies from 8 to 14 kHz, but the values of DLF/f for all these centre frequencies were significantly greater than the value at 6 kHz.

These results are consistent with the idea that DLFs depend on TFS information at low frequencies and place information at high frequencies. Over the frequency range where the place mechanism is dominant, DLF/f is roughly independent of f. However, the transition between the two mechanisms appears to occur at 8–10 kHz, rather than at 4–5 kHz; the latter would be expected from the work on the perception of musical intervals described in Sect. 2.2. Possibly, the weak phase locking for frequencies between 4 and 8 kHz is sufficient to allow reasonably small DLFs (DLF/f less than 0.02 for the data of Moore and Ernst) but not sufficient to provide a strong sense of musical interval.

2.3. The Detection of Frequency Modulation

The depth of frequency modulation (FM) of a sinusoidal carrier required for the FM just to be detectable is called the FM detection limen (FMDL). FMDLs depend upon the carrier frequency, f_c, and the modulation frequency, f_m. Unlike DLFs, FMDLs are approximately a constant proportion of ERB_N [32,40] for f_m = 10 Hz and above, as illustrated by the dashed line in Fig. 3, consistent with the predictions of Zwicker’s place model [27]. However, for low f_m, for example 2 Hz, the ratio FMDL/ERB_N is not constant across frequency, but increases with increasing frequency [40]. Moore and Sek [40] proposed that, for very low f_m, and for f_c below about 5 kHz, FM is detected using the changes in phase locking to the carrier that occur over time, that is, using TFS_{n}. They suggested that the mechanism for decoding the phase-locking information was “sluggish” and could not follow rapid changes in TFS_{n}. Hence, it played little role for high f_m, even when f_c was in a range where phase locking is reasonably precise.

To assess the relative importance of place and TFS cues for FM detection, Moore and Sek [41] measured FMDLs for wide ranges of f_c and f_m, with and without sinusoidal amplitude modulation (AM) with a 6-dB peak-to-valley ratio applied to both stimuli in a forced-choice trial (the FM was present only in one interval). The AM had the same modulation frequency as the FM, and the starting phase of the AM was chosen randomly. The AM was intended to disrupt excitation-pattern cues for FM detection by introducing large fluctuations in excitation level that were uninformative about the FM. The AM adversely affected performance and, for f_c ≤ 4 kHz, the adverse effect increased with increasing f_m, consistent with the idea that place information plays a greater role for higher f_m. For f_c = 6 kHz, the adverse effect of the AM was similar for all f_m, consistent with the idea that, for very high f_c, place information dominates for all f_m. Similar effects of added AM have been obtained by other researchers [42,43].

3. THE PITCH PERCEPTION OF COMPLEX TONES FOR NORMAL-HEARING SUBJECTS

A periodic complex tone with fundamental frequency F0 usually has a similar pitch to a sinewave with frequency close to F0, even when the sinusoidal component at F0 is absent from the complex tone [44] or is masked by low-frequency noise [45]. This is called the “phenomenon of the missing fundamental” and the pitch is often called “residue” pitch. In one class of model to explain residue pitch, the pattern-recognition models [46–49], there is an initial analysis to determine the frequencies of some of the individual sinusoidal components of the complex tone. The lower components in a complex tone are partially resolved by the BM, each leading to a peak in response at the appropriate place. The waveform evoked on the BM by a resolved component reflects the frequency of that component, and the TFS_{n} evoked at the place of maximal response to the component is similar to the TFS_{n} that would be evoked by that component presented in isolation. Hence, component frequencies are represented in the auditory system both by excitation-pattern (place) cues and by TFS_{n} cues. A role of TFS_{n} cues is supported by the finding that, for a given spacing of the components in a complex sound, the ability to “hear out” an individual component worsens when the frequency is above about 3,500 Hz [50].

In the second stage of the pattern-recognition models, the residue pitch is computed from the frequencies of the resolved components, for example by computing the F0 whose harmonic frequencies would match the frequencies determined in stage 1. For this type of model, a residue pitch can only be derived if there is at least one resolved harmonic. It is usually assumed that only harmonics up to about the eighth are resolved [51,52], although the exact upper limit is still debated. Bernstein and Oxenham [53] proposed that harmonics up to about the 10th could be resolved, although the method used by them to determine the limits of resolution has been questioned [54,55].

Other models are based on ENV_{BM} and TFS_{BM} evoked by the higher harmonics [46,56]. These harmonics interfere on the BM and the temporal pattern of the resulting waveform reflects the F0. It is assumed that pitch is derived from the periodicity of ENV_{BM} or from the time interval between peaks in TFS_{BM} close to adjacent ENV_{BM} peaks [46,56,57]. Hence, interference of harmonics on the BM is required for a residue pitch to be heard.
A residue pitch can be perceived both when only low resolved harmonics are present and when only high unresolved harmonics are present, although the pitch is weaker in the latter case [58–63]. This means that neither type of model is sufficient and has led to models in which temporal information from both resolved and unresolved harmonics is used to determine residue pitch [64–68].

To determine whether the pitch of unresolved harmonics is derived from ENV\textsubscript{BM} or TFS\textsubscript{BM}, the effect on pitch of shifting all harmonics in a harmonic complex tone (H) upwards by a fixed amount in hertz has been determined. For example, starting with an H tone with components at 1,000, 1,100, 1,200, and 1,300 Hz (\(F_0 = 100\) Hz), a frequency-shifted (inharmonic, I) tone can be created by shifting all components upwards by 30 Hz, giving components at 1,030, 1,130, 1,230, and 1,330 Hz. Example waveforms of such sounds are shown in Fig. 5.

The two left panels show waveforms for H tones with \(F_0 = 100\) Hz; the waveforms differ because the starting phases of the harmonics were chosen randomly for each stimulus. The two right panels show examples of I tones with a shift of 30 Hz. The envelopes of the waveforms (shown as thick blue lines), have the same repetition rate of 100 Hz in all cases, so if pitch is derived from ENV\textsubscript{BM}, the pitch should be the same for the H and I tones. In fact, the I tones are usually heard as having a slightly higher pitch than the H tones [46,56,57].

To eliminate possible effects related to shifts in the excitation pattern when the components are frequency shifted, Moore and Moore [70] used complex tones with many components, and passed the tones through a fixed bandpass filter centred on the higher harmonics. A background noise was used to limit the audibility of components falling on the filter slopes, and to mask combination tones. The pitch of each I tone was assessed by asking subjects to adjust the F0 of an H tone, so as to achieve a pitch match; the I and H tones were presented in alternation.

The harmonic number of the lowest component in the passband is denoted \(N\). When \(N\) was above about 14, Moore and Moore [70] found no measurable shift in pitch, whereas when \(N\) was 9, significant pitch shifts occurred. Moore and Moore argued that, when a complex tone contains only harmonics above about the 14th, the pitch is determined from the periodicity of ENV\textsubscript{BM}. The pitch shift found when \(N\) was 9 probably reflected sensitivity to TFS\textsubscript{BM}. However, it is hard to rule out the possibility that the pitch in this case was based on a pattern-matching process to (partially) resolved harmonics [46,71].
The evidence reviewed above suggests that the discrimination of bandpass filtered H and I tones depends on TFS<sub>BM</sub> cues when the value of \( N \) is 9–14, but depends mainly on ENV<sub>BM</sub> cues for higher \( N \). Based on this, Moore and Sek [72, 73] developed a discrimination task that was intended for use as a clinical tool for assessing sensitivity to TFS<sub>BM</sub>. The stimuli were bandpass filtered H and I tones and the starting phases of the components were chosen randomly for each and every tone. The tones were presented in a background of threshold-equalising noise [TEN, 74], to limit the audibility of components falling on the filter slopes and to mask combination tones. A variation on a two-interval two-alternative forced-choice method was used. In one interval of a trial (selected randomly), there were four successive bursts of tone H. In the other interval, tones H and I alternated, giving the pattern HIHI. The task was to choose the interval in which the sound changed across the four tone bursts. The repeated pattern HIHI was used to help subjects to focus on the appropriate detection cue. This test is called the “TFS1” test.

Any audible difference between the H and I tones could be used to identify the correct interval, although subjects reported using pitch changes as a cue. Subjects did not need to indicate the direction of any pitch change that was heard. This was desirable, as some people find it difficult to judge the direction of pitch changes [39]. Also, as described above, the pitch of tones with only a few audible components is ambiguous [57, 69] and sometimes tone I may be heard as lower in pitch than tone H, even though the components in tone I are shifted upwards.

The frequency shift of each component in the I tones, \( \Delta f \), was varied adaptively to determine a “threshold” value of \( \Delta f \). The H and I tones are most different when \( \Delta f = 0.5F_0 \). Moore and Sek [72] showed that all 20 of the NH subjects tested (ten with some training and ten without training) achieved thresholds below 0.5F0 when \( F_0 = 200 \) Hz and \( N = 9 \). However, only three subjects achieved thresholds below 0.5F0, when \( F_0 = 200 \) Hz and \( N = 15 \). Moore and Sek showed that learning effects were small, and that thresholds did not depend on the level of the stimuli over the range 20 to 50 dB above the detection threshold.

Moore and Sek [75] used the TFS1 test to assess sensitivity to TFS<sub>BM</sub> at high frequencies. Phase locking becomes less precise at high frequencies, so the information in TFS<sub>BM</sub> might not be conveyed adequately by TFS<sub>n</sub>. The value of \( F_0 \) was 800 or 1,000 Hz, \( N \) was set to 12, and a background TEN was used. It was estimated that the lowest audible component in the stimuli fell at 8,000 Hz for \( F_0 = 800 \) Hz and 10,000 Hz for \( F_0 = 1,000 \) Hz. For \( F_0 = 800 \) Hz, seven of the nine subjects were able to complete all three adaptive runs. Even for the subject who scored most poorly, the score was significantly above chance. For \( F_0 = 1,000 \) Hz, four of the eight subjects were able to complete all three adaptive runs and two of the others performed significantly above chance.

These results suggest that TFS<sub>n</sub> information can be used by some subjects to perform the TFS1 task when the frequency of the lowest audible component is 8,000 Hz or even 10,000 Hz. This is consistent with the way that DLFs vary with centre frequency [38], as shown in Fig. 4. It is also consistent with the results of Oxenham et al. [76]. They showed that a pitch corresponding to the missing fundamental frequency can be perceived when all of the harmonics have frequencies above 5,000 Hz. The pitch is sufficiently strong to allow the identification of simple melodies and musical intervals. The stimuli used by Oxenham et al. contained relatively low-numbered harmonics, which were presumable partly resolved along the basilar membrane. One interpretation of these findings is that phase locking above 5,000 Hz, while weak, is still sufficient to allow extraction of a residue pitch when the phase-locking information can be combined across several harmonics.

Recently, Lau et al. [77] showed that a residue pitch can be perceived for complex tones with all harmonics above 8,400 Hz. They suggested that “place-based spectral coding is sufficient to elicit complex pitch.” However, it has not yet been demonstrated that such tones can evoke a sense of musical interval. Also, the F0 discrimination thresholds for these stimuli were about four times larger than found at lower frequencies for comparable stimuli [78], suggesting that place cues alone lead to a relatively weak pitch.

It must be acknowledged that many researchers are sceptical about the possibility that TFS<sub>n</sub> information can be used for frequencies up to about 8,000 Hz. It is possible that performance in the study of Moore and Sek [75] was based on the use of excitation-pattern cues, based on shifts in the pattern of ripples in the excitation pattern, even though the depth of the ripples would be small [71].

To assess the role of excitation-pattern cues, Moore and Sek [79] measured performance on the TFS1 task as a function of level. The auditory filters tend to broaden with increasing level, especially at high frequencies [35, 80–83]. If discrimination of the H and I tones is based on excitation-pattern cues, performance on the TFS1 test should worsen with increasing level. Moore and Sek set \( F_0 = 800 \) Hz and \( N = 9 \), and the level of the stimuli ranged from 20 to 50 dB above threshold. Ten of the 12 NH subjects performed the task reliably for all levels. There was no significant effect of level. Moore and Sek concluded that the most likely explanation of their results was that the TFS1 task was performed using weak TFS<sub>n</sub> cues.

Jackson and Moore [84] also assessed the role of excitation-pattern cues in performance of the TFS1 task. They examined the effect of randomly perturbing the level
of each of the components in each of the H and I tones over ranges of ±3 dB and ±5 dB. Models based on the use of excitation-pattern cues predicted that this level perturbation should markedly impair performance, because it disrupts the pattern of ripples in the excitation patterns. However, the performance of the NH subjects was only slightly (non-significantly) affected by the level perturbation. This suggests that performance of the TFS1 task was not based on the use of excitation-pattern cues. Simulations of the outputs of auditory filters showed that TFS cues (the time intervals between peaks in TFS outputs of auditory filters) showed that TFS cues play a role in the perception of the pitch of complex tones in two ways: (1) via coding of the frequencies of individual resolved harmonics; (2) via neural coding of prominent time interval between peaks in TFSBM close to adjacent envelope maxima) were only slightly affected by the level perturbation, consistent with the idea that TFS cues were used to perform the task.

Overall, the findings described in this section suggest that TFS cues play a role in the perception of the pitch of complex tones in two ways: (1) via coding of the frequencies of individual resolved harmonics; (2) via neural coding of prominent time interval between peaks in TFSBM close to adjacent envelope peaks. When the lowest audible harmonic in a complex tone is above about the 14th, ENVBM cues may be used to extract pitch, but the pitch of such tones is very weak. Similarly, the pitch evoked by tones with all harmonics above 8,000 Hz is very weak.

4. BINAURAL PROCESSING FOR NORMAL-HEARING SUBJECTS

The ability to localise sounds in azimuth depends mainly on differences in the sound arriving at the two ears, specifically, interaural time differences (ITDs) and interaural level differences (ILDs). For sound sources presented in a free field, ITDs range from 0 for a sound at 0° azimuth to about 650 μs for a sound at ±90° azimuth. For a sinusoidal tone, an ITD is equivalent to an interaural phase difference (IPD). For example, for a 400-Hz tone, with a period of 2,500 μs, an ITD of 500 μs is equivalent to an IPD of 72° (one fifth of a cycle). For low-frequency sinusoids, the IPD provides effective and unambiguous information about azimuth, and relatively small changes in IPD (about 5° for a reference azimuth of 0°) can be detected by trained listeners [85]. The fact that the auditory system can compare the phase of a sinusoid at the two ears means that TFSn information must be preserved at least up to the point in the auditory system where neural signals from the two ears are combined. Young NH listeners can detect changes in IPD for sinusoids with frequencies up to about 1,400 Hz [86]. However, the upper frequency limit varies considerably across individuals [87]. It is not known whether the individual differences reflect differences in the neural coding of binaural TFS information or differences in auditory processing efficiency (the ability to make use of a limited amount of sensory information).

Sounds can also be localised based on ITDs in their envelopes. Envelope ITDs can be discriminated even when the centre frequency of the sounds is well above 1,400 Hz [88]. The ability to discriminate ITDs in the envelopes of high-frequency sounds is better for sounds that have sharp peaks in their envelopes than for sounds that have rounded peaks [89,90].

Differences in the sounds reaching the two ears also contribute to the ability to detect signals in the presence of background sounds. For example, the detection threshold for a tone presented in diotic white noise is lower when the tone is presented to both ears but with the phase inverted at one ear (called N0S+) than when the signal is in phase at the two ears (called N0S0). The difference between the two thresholds is known as a masking level difference (MLD). Its value is up to 15 dB at low frequencies (near 500 Hz), decreasing to 2–3 dB for frequencies above 1,500 Hz [91]. The fact the MLD is much larger for low-frequency tones than for high-frequency tones supports the idea that sensitivity to IPDs contributes to the MLD. Hence, measures of the MLD can be used as an indirect way of assessing sensitivity to IPD, and hence of the ability to use TFSn.

5. THE PERCEPTION OF SPEECH IN BACKGROUND SOUNDS BY NORMAL-HEARING SUBJECTS

The importance of ENV and TFS information for speech perception has been explored using various forms of vocoder processing. In such processing, speech in quiet or in a background sound is filtered into several frequency channels, and ENVp and TFSp are estimated for each channel. The signal in each channel is manipulated to alter either ENVp or TFSp. Finally, each manipulated channel signal is filtered to restrict its spectrum to the original channel bandwidth, and the resulting signals are combined.

If the intention is to disrupt ENV cues, ENVp in each channel may be replaced by a constant equal to the mean of ENVp. This has sometimes been described as “removing” ENV cues and the resulting signal is called “TFS speech.” If the intention is to disrupt TFS cues, TFSp in each channel may be replaced by a sinewave or a band of noise with centre frequency equal to the channel centre frequency. This has sometimes been described as “removing” TFS cues and the resulting signal is called “ENV speech.”

There is a fundamental problem with this approach. When a processed signal is passed through an array of bandpass filters, as occurs in the auditory system, information about ENVp can be extracted from TFSBM and TFSn, and information about TFSp can be extracted from ENVBM and ENVn [92–94]. Effectively, ENVp or TFSp that have been physically removed from the channel signals in the
vocoder are recreated in the auditory system. Despite this, the analyses of Hopkins et al. [95] and of Swaminathan and Heinz [96] indicate that it is possible to process speech so as to alter ENV_p and TFS_p to different extents, and this can provide a basis for evaluating their relative importance.

Hopkins et al. [97] studied the effect of vocoding over a limited frequency range. They measured speech reception thresholds (SRTs: the speech-to-background ratio required for 50% correct key words in sentences) for target sentences presented in a background of speech from a single talker. The mixture was filtered into 30 channels, each with a width of one ERBN. In condition “TFS-low,” channels up to and including channel number J were unprocessed, and hence contained intact TFS and ENV information. For channels above the Jth, TFS_p was replaced by a sinusoid at the channel centre frequency, so that original TFS information was disrupted, while ENV_p was preserved. In condition “TFS-high,” TFS_p in channels up to channel number J was replaced by a sinusoid at the channel centre frequency, while higher channels were intact.

The SRT was measured as a function of J, which was varied from 0 to 30. The mean results for seven NH subjects are shown in Fig. 7. Performance improved as original TFS information was added, either starting from low frequencies (TFS-low, open circles) or starting from high frequencies (TFS-high, filled circles). These results suggest that original TFS information plays a role in the ability to identify speech in a background sound, and that TFS information is usable over a wide frequency range. However, the improvement produced by adding original TFS may have been partly a result of more accurate representation of the ENV information in the original signal.

Sheft et al. [98] created TFS speech (with a “flat” envelope in each channel) using 16 channels that were about 2 ERBN wide or 32 channels that were about 1 ERBN wide. They created two different modulators representing the fluctuations in TFS_p for a given channel, i.e., the patterns of FM. One modulator, denoted FM_u, contained the unmodified pattern of frequency modulation. For a second modulator, denoted FM_r, the amount of FM was reduced so that deviations in instantaneous frequency were restricted to the passband of the channel. The modulator for a given channel was applied to a sinusoidal carrier at the centre frequency of that channel. A third condition, FM_t, was obtained by using the FM_r modulator for each channel, but randomizing the starting phase of the carrier, separately for each channel carrier. The processed channel signals were then combined. Simulations of processing in the peripheral auditory system suggested that envelope reconstruction was lower for the FM_r and FM_u modulators than for the FM_t modulator. The FM_t modulators led to similar envelope reconstruction, but the intelligibility of processed VCV syllables was lower for the FM_u than for the FM_t modulator. Intelligibility scores were only poorly correlated with estimates of the amount of envelope reconstruction, but were more highly correlated with estimates of the fidelity of preservation of TFS cues (effectively, the extent to which TFSAM cues for the processed signals resembled TFSRM for the original signals). The authors concluded that TFS cues convey important phonetic information that is not solely a consequence of envelope reconstruction.

The intelligibility of TFS speech may be reduced by the large excursions of instantaneous frequency that occur in the TFS_p of a channel signal when its envelope magnitude is near zero. Such excursions are a consequence of the amplification of low-level noise in the signal. Hopkins et al. [95] reduced the excursions by adding low-level “low-noise” noise [LNN, 99] to each channel signal before processing to flatten ENV_p and leave TFS_p. LNN is created by manipulating the phases of the components in a Gaussian noise so that the envelope fluctuations are greatly reduced. In the study of Hopkins et al., the LNN for each channel was synthesized to have the same long-term average spectrum as the target speech within that channel. Thus, the LNN had a narrow bandwidth that was matched to the channel bandwidth. When such LNN is added to a channel signal, the excursions in instantaneous frequency of TFS_p for that channel signal are markedly reduced. Speech processed in this was is called TFS-LNN speech.

Using 12-channel processing, Hopkins et al. found that the intelligibility of sentences was higher for TFS-LNN speech than for “conventional” TFS speech (when LNN was not added). To assess whether the better intelligibility of TFS-LNN speech was related to greater envelope
reconstruction, Hopkins et al. used a similar method to Gilbert and Lorenzi [100]. Sentences were first processed to produce TFS-LNN speech and then the TFS-LNN speech was used as input to a tone vocoder with analysis filters designed to have similar frequency selectivity to human auditory filters. The resulting output is called “recovered-envelope speech.” The intelligibility of the recovered-envelope speech was very low when 12 channels were used for the original processing to create TFS-LNN speech, suggesting that envelope recovery was small. Hopkins et al. concluded that TFS information probably contributed to the relatively high intelligibility of the TFS-LNN speech.

In summary, the studies using TFS speech suggest that TFS information does contribute to intelligibility. However, the interpretation of these studies depends on simulations of the extent of envelope recovery in the auditory system, and the accuracy of these simulations is not clear. Some amount of envelope recovery always occurs, and this probably affects the measured speech intelligibility.

6. EFFECTS OF AGE AND HEARING LOSS ON PITCH PERCEPTION

In what follows, subject groups are designated as follows: young normal hearing (YNH), older normal hearing (ONH), young hearing impaired (YHI), and older hearing impaired (OHI). HI refers to hearing impaired regardless of age. In all studies reviewed, the hearing loss was assumed to be primarily cochlear in origin. NH usually means audiometric thresholds ≤20 dB HL over the frequency range 0.5 to 4 kHz (or sometimes 0.5 to 6 kHz), although in some studies the audiometric threshold was ≤20 dB HL only for frequencies close to the test frequency. Older usually implies ages of 60 years or more, although in a few studies the older group included 40 year old subjects.

In many of the studies described below, results were compared for YNH and OHI groups, so the effects of hearing loss and age are confounded.

6.1. The Perception of Pitch and Musical Intervals by Hearing-Impaired Subjects with Dead Regions

In some cases of hearing loss, there is a region of the cochlea with no or very few functioning inner hair cells, synapses or neurons. Such a region is called a dead region [DR, 101], since no information about BM vibration in that region is transmitted to the brain. DRs can occur in the base of the cochlea (a high-frequency DR), in the apex of the cochlea (a low-frequency DR), in an intermediate region, or in multiple places [102]. DRs are often diagnosed using the psychophysical tuning curve (PTC), which is the level of a narrowband masker required just to mask a fixed low-level signal, plotted as a function of the masker centre frequency; when the tip of the PTC is shifted away from the signal frequency, this is taken to indicate a DR at the signal frequency [74,103–105].

When there is a low-frequency DR, a low-frequency sinusoid cannot produce maximum neural excitation at the place corresponding to its frequency, since there are no neural responses at that place. The peak in the neural excitation pattern occurs more towards the base of the cochlea. If the place theory is correct, this should lead to a marked upward shift in the pitch of the tone. In fact, this usually does not happen. Florentine and Houtsma [103] obtained pitch matches between the two ears of a subject with a low-frequency DR in one ear only. They presented the stimuli at levels just above absolute threshold, to minimize the spread of excitation along the BM. Pitch shifts between the two ears were small, consistent with the pitch depending on TFS information rather than place information.

Turner et al. [106] studied six subjects with low-frequency hearing loss. Three subjects did not have any DR and the other three had low-frequency DRs. Pitch perception was studied either by pitch matching between the two ears (for subjects with unilateral losses) or by octave matching (for subjects with bilateral losses, but with some musical ability). The subjects with DRs gave results similar to those of the subjects without DRs; no distinct pitch anomalies were observed, inconsistent with what would be expected based on place theory.

Huss and Moore [107] studied pitch perception for subjects with a variety of configurations of DRs. Pitch matches within one ear were reasonably consistent when the tones did not fall within a DR, but were often erratic for tones falling more than half an octave into a low-frequency or high-frequency DR, indicating that such tones usually do not evoke a clear pitch sensation. This is consistent with the large DLFs that are found for tones whose frequencies fall in a DR [108–111]. Pitch matches across the ears of subjects with asymmetric hearing loss and octave matches within ears indicated that sinusoids falling within a DR were often perceived with a pitch that was different (usually higher) than “normal.” This was true for sinusoids falling in both low-frequency and high-frequency DRs. The results for cases with low-frequency DRs differ from those of Florentine and Houtsma [103] and Turner et al. [106]. For tones with frequencies below 5,000 Hz, the pitches were usually not consistent with what would be expected from the use of either place information or TFS information alone. However, for tones with frequencies above 5,000 Hz, the pitches were roughly consistent with what would be expected from the place where the tones were detected.

Taken together, the results of studies of pitch perception using people with DRs indicate the following:
(1) Tones with frequencies falling in a DR do not evoke a clear pitch sensation.
(2) Tones falling within a DR are sometimes perceived with a near-normal pitch and sometimes with a pitch distinctly different from normal.
(3) The shifted pitches found for some subjects indicate that the pitch of low-frequency tones is not represented solely by TFSn regardless of the CFs of the neurons within which the intervals occur.

Possibly, there needs to be a correspondence between place and TFS\textsubscript{n} information for a normal pitch to be perceived [15,17,112]. Alternatively, TFS\textsubscript{n} information may be “decoded” by a network of coincidence detectors whose operation depends on the relative phase of the responses at different points along the BM [17,113]. Alteration of the phase response of the BM by cochlear hearing loss [114,115] may prevent effective use of TFS\textsubscript{n} information.

6.2. Effect of Hearing Loss and Age on the Frequency Discrimination of Sinusoids

If frequency discrimination depends on a place mechanism, the DLF should depend on the slope of the excitation pattern and there should be a correlation between DLFs and measures of frequency selectivity at the same centre frequency. Tyler \textit{et al.} [116] tested this idea using YNH and OHI subjects and found a low correlation between DLFs and frequency selectivity measured using PTCs. They concluded that frequency discrimination was not closely related to frequency selectivity, indicating that place models were not adequate to explain DLFs.

Moore and Peters [117] measured DLFs for four groups: YNH, YHI, ONH, and OHI. The auditory filter shapes of the subjects were estimated in earlier experiments using the notched-noise method [118,119] for centre frequencies of 100, 200, 400 and 800 Hz. The DLFs for both HI groups were higher than for the YNH group at all frequencies tested (50–4,000 Hz). The DLFs for the ONH group were intermediate. The DLFs at a given centre frequency were only weakly correlated with the sharpness of the auditory filter at that centre frequency, and some subjects with broad filters at low frequencies had near-normal DLFs at low frequencies. These results again suggest that a place model is not adequate to account for DLFs.

Ernst and Moore [120] measured DLFs over a wide frequency range for OHI subjects, using the same method as Moore and Ernst [38], as described earlier. They reasoned that if hearing loss is associated with a reduced ability to use TFS\textsubscript{n}, then the breakpoint in the function relating DLF/f to f (as shown in Fig. 4) should occur at a lower frequency for OHI than for NH subjects; disruption of the ability to use TFS\textsubscript{n} should have a greater effect for frequencies where TFS\textsubscript{n} is already weak. A reduced ability to use TFS\textsubscript{n} should also result in larger DLFs at low frequencies, while poorer frequency selectivity should lead to larger DLFs at high frequencies.

The open squares in Fig. 8 show the geometric mean values of DLF/f for the OHI subjects. For comparison, the circles show values of DLF/f obtained at the same sensation level (20 dB SL) for the YNH subjects of Moore and Ernst [38]. The OHI subjects showed significantly higher DLFs than the YNH subjects. The values of DLF/f for the OHI subjects increased when the frequency was increased from 2 to 4 kHz and then flattened off. There was no significant difference between the values of DLF/f for centre frequencies from 4 to 12 kHz. The values of DLF/f for all these centre frequencies were significantly greater than the value of DLF/f at 2 kHz. The results are consistent with what would be expected if the OHI subjects had reduced sensitivity to TFS combined with a broadening of the auditory filters.

Auditory filters do not broaden with increasing age when audiometric thresholds remain normal [121,122]. Hence, if frequency discrimination worsens with increasing age even for subjects with normal audiograms, this cannot be attributed to reduced frequency selectivity. However, poorer frequency discrimination might be associated with a general reduction in “processing efficiency” for many types of sounds.

Abel \textit{et al.} [123] measured DLFs for centre frequencies of 500 and 4,000 Hz for YNH and ONH subjects. The stimuli were presented at 70 dB SPL. The ONH group had higher DLFs than the YNH group, by a factor of about 1.8 at 500 Hz and 2.3 at 4,000 Hz. However, Abel \textit{et al.} also found that increasing age adversely affected duration discrimination, consistent with a general deficit in auditory processing.

He \textit{et al.} [124] assessed the effect of age on frequency and intensity discrimination for subjects with near-normal
hearing using sinusoids with frequencies from 0.5 to 4 kHz and levels of 40 and 80 dB SPL. The mean audiograms of the YNH and ONH subjects were closely matched for frequencies up to 4 kHz. Values of DLF/\( f \), were consistently higher for the ONH than for the YNH subjects, by a factor of about 2.5 to 3.6 for lower frequencies (0.5 and 1 kHz) and 1.4 to 2.1 for higher frequencies (2 and 4 kHz). The intensity discrimination thresholds were also higher for the ONH than for the YNH subjects, by a factor of about 2.1 to 3.6 for lower frequencies (0.5 and 1 kHz) and 1.2 to 2.5 for higher frequencies (2 and 4 kHz). It is hard to know how to interpret the finding of a larger age effect at lower than at higher frequencies, since Abel et al. [123] did not find such an effect.

Cliner et al. [125] examined the effect of age on DLFs for frequencies of 0.5 and 1 kHz, using 32 subjects with ages from 22 to 77 years. All subjects had audiometric thresholds ≤25 dB HL for frequencies from 250 to 8,000 Hz. The sound level of the stimuli was not specified. They also measured the frequency following response (FFR) to 0.5- and 1-kHz sinusoids. The FFR is a scalp-recorded electrical signal that reflects the synchronous activity (phase locking) of sub-cortical neurons. DLFs increased with increasing age for both signal frequencies, and the correlations of DLFs with age were significant (\( r = 0.55, p = 0.002 \) for 500 Hz and \( r = 0.42, p = 0.015 \) for 1 kHz). Measures of FFR strength tended to decrease with increasing age, but the decrease was significant only for frequencies near 1 kHz. However, the measures of FFR strength were not significantly correlated with the DLFs at 0 kHz. Experts disagree on how to interpret the finding of a larger age effect at lower than at higher frequencies, a finding that was absent. The corresponding ratio for the YNH subjects was only 1.45. The relatively small disruptive effect of the AM tended to increase with increasing \( f_m \), for \( f_c \) below 6 kHz, consistent with the results of Moore and Sek [41]. For the OHI subjects, the adverse effect of the AM was generally larger than for the YNH subjects, and the magnitude of the effect did not consistently increase with increasing \( f_m \). For \( f_c = 2 \) Hz, the FMDLs for the OHI subjects, averaged across the four lowest values of \( f_c \), were a factor of 2.5 larger when AM was present than when it was absent. The corresponding ratio for the YNH subjects was only 1.45. The relatively small disruptive effect of the AM for the YNH subjects probably reflects the use of TFS\(_n\) information for \( f_m = 2 \) Hz. The larger effect for the OHI subjects suggests that they were not using TFS\(_n\) information effectively.

Ernst and Moore [42] measured FMDLs for OHI subjects. They used \( f_c = 1, 4 \) and 6 kHz, and \( f_m = 2 \) and 10 Hz, with levels of 20 dB SL and 90 dB SPL. The SL corresponding to 90 dB SPL varied across subjects and frequencies from about 30 to 60 dB. FMDLs were estimated both with and without AM as applied to the carrier in all three intervals of a forced-choice trial, to disrupt excitation pattern cues. In addition, thresholds for detecting sinusoidal AM (AMDLs) were measured for the same \( f_c \) and \( f_m \) values as used for the measurement of FMDLs. This was done to provide a quantitative measure of the ability to detect fluctuations in excitation level and of

---

**6.3. Effects of Hearing Loss and Age on Frequency Modulation Detection**

Moore and Skrodzka [127] measured FMDLs for YNH and OHI subjects, for a wide range of carrier frequencies, \( f_c \), and modulation frequencies, \( f_m \). The FMDLs were measured both without and with sinusoidal AM added in both intervals of a forced-choice trial to disrupt excitation-pattern cues. Generally, FMDLs were larger for the OHI than for the YNH subjects. The adverse effect of AM tended to increase with increasing \( f_m \), for \( f_c \) below 6 kHz, consistent with the results of Moore and Sek [41]. For the OHI subjects, the adverse effect of the AM was generally larger than for the YNH subjects, and the magnitude of the effect did not consistently increase with increasing \( f_m \). For \( f_c = 2 \) Hz, the FMDLs for the OHI subjects, averaged across the four lowest values of \( f_c \), were a factor of 2.5 larger when AM was present than when it was absent. The corresponding ratio for the YNH subjects was only 1.45. The relatively small disruptive effect of the AM for the YNH subjects probably reflects the use of TFS\(_n\) information for \( f_m = 2 \) Hz. The larger effect for the OHI subjects suggests that they were not using TFS\(_n\) information effectively.
the way that this ability varied with \( f_\text{m} \). For low values of \( f_\text{m} \) (below about 5\% of \( f_c \)), the excitation pattern of FM tones simply shifts backwards and forwards along the frequency axis, without changing shape (as illustrated in Fig. 2). Hence, if FMDLs for both \( f_\text{m} \) values used are determined solely by place mechanisms, the ratio of AMDLs for the two values of \( f_\text{m} \) should equal the ratio of FMDLs for the two rates.

The general pattern of results was similar for the two levels used. For \( f_c = 4 \) and 6kHz, the FMDLs measured without added AM were much smaller for \( f_\text{m} = 10 \) Hz than for \( f_\text{m} = 2 \) Hz. Correspondingly, the AMDLs were also much smaller for \( f_\text{m} = 10 \) Hz than for \( f_\text{m} = 2 \) Hz. This is consistent with the idea that FM detection for these two values of \( f_c \) was largely based on excitation-pattern cues. Small AMDLs imply a good ability to detect changes in excitation level associated with place cues for FM detection. For \( f_c = 1 \) kHz, AMDLs were also much smaller for \( f_\text{m} = 10 \) Hz than for \( f_\text{m} = 2 \) Hz. However, FMDLs were similar for the two values of \( f_\text{m} \). Since, at a given level, the excitation patterns would have been equally sharp for the two values of \( f_\text{m} \), this finding implies that FM detection for \( f_c = 1 \) kHz was not based solely on the use of excitation-pattern cues; if it were, FMDLs should have been smaller for \( f_\text{m} = 10 \) Hz than for \( f_\text{m} = 2 \) Hz. Also, FMDLs for \( f_\text{m} = 2 \) Hz, expressed as a proportion of \( f_c \), were smaller for \( f_c = 1 \) kHz than for \( f_c = 4 \) or 6kHz.

These results are consistent with the idea that the use of TFS\textsubscript{n} cues led to relatively small FMDLs for \( f_\text{m} = 2 \) Hz and \( f_c = 1 \) kHz. If so, this implies that the OHI subjects did have some ability to use TFS\textsubscript{n} information at 1kHz, perhaps because their hearing loss was mild (about 40 dB) at 1,000 Hz. The added AM led to increases in the FMDLs for the two values of \( f_\text{m} \), but the adverse effect was smallest for \( f_c = 1 \) kHz and \( f_\text{m} = 2 \) Hz, consistent with the idea that the FM at this low rate and for low \( f_c \) was detected partly by tracking changes in TFS\textsubscript{n} information over time.

I consider next studies that focused on the effect of age. Whiteford et al. [128] measured FMDLs and AMDLs for slow (1 Hz) and fast (20 Hz) rates for 85 subjects with a wide range of ages. All had audiometric thresholds ≤20 dB HL for low frequencies. In some conditions, the stimuli were presented diotically (the same to each ear). FMDLs and AMDLs were correlated with each other even for the 1-Hz rate. This might suggest place-based FM detection mechanisms for both rates, but the correlations might also have occurred because of individual variability in processing efficiency. FMDLs were correlated with age for both slow and fast rates, again consistent with individual variability in processing efficiency. When the effect of AM sensitivity was controlled for, only slow-rate FMDLs were significantly correlated with age. The authors concluded that “this outcome provides some evidence for a role for TFS coding in slow diotic FM detection and perhaps no role, or a smaller role, for TFS coding in fast diotic FM.”

To control for the effects of processing efficiency, Moore et al. [129] conducted a two-stage experiment. In stage 1, psychometric functions were measured for the detection of AM alone and FM alone imposed on a 1-kHz carrier, using 2- and 10-Hz rates. In stage 2, the task was to discriminate AM from FM at the same modulation rate when the detectability of the AM alone and FM alone was equated. It was argued that if FM is detected only via FM-to-AM conversion, then in stage 2 it should be relatively difficult to discriminate AM from FM. However, if FM is partly coded via TFS cues, FM will give rise to detectable fluctuations in TFS\textsubscript{BM} and TFS\textsubscript{n} while AM will not, and it should be easier to discriminate AM from FM. Based on previous evidence suggesting that TFS information is used to detect 2-Hz FM but not 10-Hz FM, this leads to the prediction that discrimination of FM from AM should be better for a 2-Hz rate than for a 10-Hz rate.

For YNH subjects, discrimination of AM from FM was consistently better for the 2- than for the 10-Hz rate, as shown previously [40,130]. Some older subjects with normal hearing at 1kHz showed a similar pattern but others showed no difference in AM-FM discrimination for the two rates. Moore et al. argued that greater age reduces the ability to use TFS cues for some but not all people.

In summary, the data on FM detection by HI listeners are consistent with the idea that FM detection for \( f_\text{m} \geq 10 \) Hz depends largely on the use of excitation-pattern (place) cues. The broader auditory filters that often occur with cochlear hearing loss lead to impaired FM discrimination for \( f_\text{m} \geq 10 \) Hz. The addition of AM has large disruptive effects on FM detection for \( f_\text{m} \geq 10 \) Hz. For very low \( f_\text{m} \), such as 2 Hz, FM detection may be partly based on the use of TFS\textsubscript{n} cues when \( f_c \leq 4,000 \) Hz, but the ability to use such cues probably depends on the degree of hearing loss. Subjects with mild loss appear to have some ability to use TFS\textsubscript{n} cues, while those with greater loss probably rely largely on excitation-pattern cues. The addition of AM has only small disruptive effects when TFS\textsubscript{n} cues can be used. Age appears to adversely affect the ability to use TFS cues for detection of FM at very low rates for \( f_c \leq 4 \) kHz.

6.4. The Effect of Age and Hearing Loss on the Pitch Perception of Complex Tones

Hopkins and Moore [131] assessed the ability to discriminate harmonic (H) and frequency-shifted inharmonic (I) bandpass filtered tones of the type described in Sect. 3 and illustrated in Figs. 5 and 6. They tested YNH subjects (tested at 65 dB SPL) and HI subjects with a wide range of ages (22–79 years, tested at 20 or 30 dB SL). The lowest harmonic in the passband was the ninth. While the
YNH subjects were able to perform the task well for F0s of 100, 200 and 400 Hz, four of the six HI subjects were unable to perform above chance in any condition. The other two HI subjects could perform the task when the filter passband fell in a frequency region where their audiometric thresholds were near-normal, but otherwise performed very poorly. Hopkins and Moore concluded that subjects with moderate hearing loss have very little or no ability to use TFSBM cues to discriminate H and I tones.

It is possible that the poor performance of the HI subjects reflected their reduced frequency selectivity, which would decrease the salience of excitation-pattern cues. If this were the case, one would expect poor discrimination of the H and I tones to be associated with reduced frequency selectivity. To assess this, Hopkins and Moore [132] measured performance on the TFS1 task and measured frequency selectivity using the notched-noise method [118] for the same subjects for centre frequencies of 750, 1,000 and 2,000 Hz. The TFS1 scores were expressed in terms of the discriminability index, $d'$, for which larger numbers indicate better performance. The measure of frequency selectivity was the equivalent rectangular bandwidth of the auditory filter (ERB), for which larger numbers indicate poorer frequency selectivity. Three groups of subjects were tested: YNH, ONH, and HI with a wide range of ages.

As found previously [97], the HI subjects generally performed poorly on the TFS1 test, but with considerable individual variability. For the data collapsed across subject groups, the TFS1 scores were modestly but significantly correlated with the measures of frequency selectivity at the same centre frequency (correlations ranged from −0.29 to −0.48). However, both TFS1 scores and the ERB values were correlated with audiometric thresholds, and this could account for the significant correlations between the two measures. When the effect of audiometric threshold was partialled out, the correlations between TFS1 scores and ERB values were no longer significant. The results for the centre frequency of 1,000 Hz are shown in Fig. 9.

It is noteworthy that there were some subjects with normal ERB values (ERB/centre frequency ≈ 0.15) who performed at chance on the TFS1 test (indicated by symbols within the ellipse in Fig. 9), while there was one HI subject with relatively poor frequency selectivity (ERB/centre frequency ≈ 0.28) who performed well above chance on the TFS1 test (indicated by the symbol in the circle). These results show that poor performance on the TFS1 test can be caused by factors other than reduced frequency selectivity, and that reduced frequency selectivity does not always lead to poor performance on the TFS1 test, supporting the idea that scores on the test do reflect sensitivity to TFSBM.

Füllgrabe et al. [133] measured performance on the TFS1 test for ONH and YNH subjects who were closely matched in mean audiograms, years of education, and performance IQ. The ONH subjects performed significantly more poorly than the YNH subjects for all centre frequencies used. This confirms that TFS processing worsens with increasing age even when audiometric thresholds remain normal.

7. EFFECTS OF AGE AND HEARING LOSS ON BINAURAL PROCESSING

7.1. Effects of Age and Hearing Loss on ITD and IPD Discrimination

Lacher-Fougère and Demany [134] measured thresholds for detecting changes in ITD (relative to a reference ITD of zero) using AM sinusoids. The ITDs were imposed either on the carrier alone (i.e., TFS, as illustrated in panel a of Fig. 10) or on the modulator alone (i.e., ENV, as illustrated in panel b of Fig. 10). The carrier frequency,
\( f_c \), was 250, 500, or 1,000 Hz, the modulation frequency, \( f_m \), was 20 or 50 Hz, and the level was 75 dB SPL. Seven NH subjects (age range 24 to 45 years) and nine HI subjects (age range 42–68 years) were tested. For both types of ITD, thresholds were on average higher for the HI than for the NH subjects. However, the deficit for the former was markedly larger for ITDs in TFS\(_P\) than for ITDs in ENV\(_P\). The authors concluded that one consequence of hearing loss is reduced sensitivity to TFS. However, the observed effects might be partly a consequence of the fact that the HI subjects were, on average, older than the NH subjects.

There have been several studies of sensitivity to changes in ITD as a function of age using subjects whose audiometric thresholds at the test frequency were normal or near-normal. Ross \textit{et al.} [135] used AM tones in which the modulation was in phase at the two ears. They measured the highest carrier frequency for which subjects could detect a difference between a stimulus with 0° IPD in the carrier and a stimulus with 180° IPD in the carrier; this frequency is denoted \( f_{\text{limit}} \). They tested 12 young subjects (mean age 27 years), 11 middle-aged subjects (mean age 51 years) and 10 older subjects (mean age 71 years). They found a significant negative correlation between \( f_{\text{limit}} \) and age \((r = -0.53)\). All of the young subjects performed the task reliably, but four subjects from the middle-aged group and five subjects from the older group were not able to perform the task reliably.

Grose and Mamo [136] used similar stimuli to those of Ross \textit{et al.} [135] to test 15 younger (mean age 22 years), 12 middle-aged (mean age 48 years), and 12 older (mean age 68 years) subjects, all with audiometric thresholds \( \leq \)20 dB HL for frequencies of 2,000 Hz and below. The value of \( f_{\text{limit}} \) was significantly lower for the middle-aged than for the young group, and was lower still for the older group. In a second experiment, IPD discrimination thresholds were measured for several fixed carrier frequencies. The middle-aged subjects had higher IPD thresholds than the young subjects for all but the lowest carrier frequencies tested. The older subjects, as a group, had the highest IPD thresholds. The authors concluded that the ability to process binaural TFS declines with increasing age, and that the deficits emerge relatively early in the aging process.

Several studies have measured sensitivity to binaural TFS using the TFS-LF test developed by Hopkins and Moore [137]. This test makes use of an adaptive two-alternative forced-choice task. Each interval contains four successive sinusoidal tone bursts, all of the same frequency. In one interval all tones are diotic (IPD = 0), and in the other interval tones one and three are diotic while tones two and four have an IPD in the carrier only. The envelopes of the tone bursts are always synchronous across ears. The task is to identify the interval with the phase-shifted tones. For subjects who are sensitive to IPD, the tones in this interval are heard to move within the head. Hopkins and Moore showed that, for NH subjects, the effects of SL and training on performance were small (provided that the level was above 30 dB SL), and the test could be performed reliably for frequencies of 250, 500 and 750 Hz.

Füllgrabe and Moore [138] conducted a meta-analysis of 19 studies that used the TFS-LF test to assess the effects of age, hearing loss or both. The number of subjects ranged from 147 to 648, depending on the test frequency (250, 500, 750, and 850 Hz). The IPD thresholds were converted to values of the detectability index, \( d' \), that would be obtained for an IPD of 180°, assuming that the value of \( d' \) is proportional to the change in IPD. For all test frequencies, both audiometric threshold and age were significantly negatively correlated with the \( d' \) scores (with \( r \) varying from \(-0.19\) to \(-0.64\)) but the correlation was always significantly higher for age than for audiometric threshold.

Figure 11 shows a three-dimensional representation of mean \( d' \) scores as a function of audiometric threshold and age. Binaural TFS sensitivity worsened with increasing age above about 40, while it was additionally affected by audiometric threshold for older subjects.

Some older subjects are completely unable to perform the TFS-LF test, and hence no graded measure of sensitivity to TFS is obtained. Füllgrabe \textit{et al.} [87] modified the TFS-LF test to overcome this limitation. In their test, referred to as the TFS-AF test (where AF stands for adaptive frequency), the IPD is fixed, usually at 180°, and the frequency of the tone is adaptively varied, as was done in some of the studies described earlier [137].

---

**Fig. 11** Binaural TFS sensitivity (\( d' \)) as a function of hearing loss (10-dB-wide threshold groups) and age (10-year-wide age groups). Data are from Füllgrabe and Moore [138].
the great majority of subjects are able to perform the TFS-AF test.

Füllgrabe et al. [139] used the TFS-AF test to assess 118 subjects aged 60 to 96 years with normal or near-normal low-frequency hearing, but a variety of patterns of hearing loss at higher frequencies. TFS-AF scores were significantly lower (i.e. poorer) than those for young adults. On average, scores decreased by about 162 Hz for each ten-year increase in age over the range 60 to 85 years.

King et al. [140] measured the ability to discriminate ITD using AM sinusoids. The ITDs were imposed either on the carrier alone, TFS_p, or the modulator alone, ENV_p, as illustrated in Fig. 10. The carrier frequency, f_c, was 250 or 500 Hz, the modulation frequency, f_m, was 20 Hz, and the level was approximately 30 dB SL. They tested 46 subjects with a wide range of ages (18 to 83 years) and degrees of hearing loss. The subjects were selected so that absolute thresholds at the carrier frequencies of 250 and 500 Hz were not significantly correlated with age.

Thresholds for discriminating ITDs in TFS_p were positively correlated with absolute thresholds for both carrier frequencies (r = 0.45, p = 0.002 for f_c = 250 Hz; r = 0.40, p = 0.006, for f_c = 500 Hz). The correlations remained positive and significant when the effect of age was partialled out. However, thresholds for discriminating ITDs in ENV_p were not significantly correlated with absolute thresholds. These results suggest that hearing loss adversely affects sensitivity to ITDs in TFS_p, independently of age, but does not adversely affect sensitivity to ITDs in ENV_p.

Thresholds for detecting ITDs in TFS_p were positively correlated with age only for the carrier frequency of 500 Hz (r = 0.44, p = 0.002). The correlation remained positive and significant when the effect of absolute threshold was partialled out. Thresholds for detecting ITDs in ENV_p were positively correlated with age for both carrier frequencies (r = 0.62, p < 0.001 for f_c = 250 Hz; r = 0.58, p < 0.001, for f_c = 500 Hz). The correlations remained positive and significant when the effect of absolute threshold was partialled out. These results suggest that increasing age adversely affects the ability to discriminate IPDs in both TFS_p and ENV_p, independently of the effects of hearing loss.

One concern about studies of the effect of age is that any worsening in performance with age may reflect changes in processing efficiency rather than a specific deficit in binaural TFS processing. To address this concern, Moore et al. [141] used similar stimuli to those of Lacher-Fougère and Demany [134] and King et al. [140] to assess 12 young and 11 older subjects, all with audiometric thresholds ≤20 dB HL at the carrier frequencies of 250 and 500 Hz. The mean threshold for discriminating ITDs in TFS_p was 70 µs for the young subjects and 121 µs for the older subjects, a worsening with greater age by a factor of 1.7. The mean threshold for discriminating ITDs in ENV_p differed only slightly for the young group and the older group (1,312 and 1,517 µs, respectively), by a factor of 1.16. The small effect of age on TFS_p thresholds suggests that changes in processing efficiency with age are small.

### 7.2. Effects of Hearing Loss on MLDs

An advantage of MLDs is that they are based on the difference between two thresholds (e.g., N_0S_p and N_0S_0). This reduces the confounding effect of processing efficiency, since this effect should be similar for the two thresholds and should roughly cancel when the difference in thresholds is taken.

A survey of studies of the MLD using HI subjects was presented by Durlach et al. [142]. Generally, cochlear hearing loss was associated with reduced MLDs. For example, Quaranta and Cervellera [143] reported abnormally small MLDs for 86% of people with hearing loss.

Several more recent studies have confirmed that MLDs are smaller than normal for HI subjects [144–147]. MLDs tend to reduce with increasing absolute threshold at the test frequency, although the association is not strong, and subjects with similar absolute thresholds can have very different MLDs. MLDs also tend to be smaller the more asymmetrical is the loss [145]. Overall, the results support the idea that the binaural processing of TFS information is impaired by cochlear hearing loss.

### 7.3. Effects of Age on MLDs

Pichora-Fuller and Schneider [148] measured MLDs using a 500-Hz signal for 12 young subjects and 12 older subjects. All subjects had audiometric thresholds better than 25 dB HL for frequencies from 250 to 2,000 Hz but the older group had some hearing loss at higher frequencies. Pichora-Fuller and Schneider included a variety of signal and masker configurations, including N_0S_0 and N_0S_p. They also used both a continuous masker and a masker presented in bursts. MLDs were significantly smaller for the older than for the younger subjects, but the magnitude of the difference varied markedly across configurations.

Grose et al. [149] measured MLDs for nine older subjects with audiometric thresholds ≤20 dB HL for frequencies up to 2,000 Hz, and ten YNH subjects. In one experiment, they measured MLDs for a 500-Hz tone presented in narrowband noise. MLDs were about 4 dB smaller for the older than for the young subjects, mainly due to higher thresholds for the older subjects in the N_0S_p condition. In a second experiment, they measured thresholds for recognizing spondee words presented in broadband noise with the same long-term average spectrum as the speech, under N_0S_0 and N_0S_p conditions. MLDs were
about 3 dB smaller for the older than for the young group, again due to poorer performance in the N0S0 condition. For both experiments, N0S0 threshold was correlated with the average audiometric threshold at 0.5, 1, 2, and 4 kHz, suggesting a contribution of peripheral factors to binaural TFS processing.

Strouse et al. [150] also measured thresholds for recognizing spondee words in broadband noise in N0S0 and N0Sx conditions. They tested 12 YNH and 12 ONH subjects. The older group had higher thresholds than the young group for both N0S0 and N0Sx conditions, but the deficit was greater for the N0Sx condition. Mean MLDs were about 2 dB smaller for the ONH than for the YNH subjects. The authors concluded that age-related factors other than peripheral hearing loss contribute to temporal processing deficits in older subjects.

In summary, there is good evidence that age has an influence on binaural TFS processing as assessed using MLDs. Even when audiometric thresholds are within the “normal” range, binaural TFS processing deteriorates with increasing age.

8. EFFECTS OF AGE AND HEARING LOSS ON THE USE OF TFS INFORMATION FOR SPEECH PERCEPTION

One approach to studying the role of TFS in speech perception is to examine the correlation between psychoacoustic measures of sensitivity to TFS and measures of the intelligibility of speech presented in quiet or in background sounds. In one study using this approach, Hopkins and Moore [132] measured sensitivity to TFS, frequency selectivity, and speech reception in noise for YNH and ONH subjects, and HI subjects with a range of ages.

Two measures of TFS sensitivity were used. One was very similar to the TFS1 test described in Sect. 3. The other test was the TFS-LF test described in Sect. 7.1. Frequency selectivity was measured using a shortened version of the notched-noise method [151] for centre frequencies of 0.5, 1, and 2 kHz. The ONH group performed significantly more poorly than the YNH group for both measures of TFS sensitivity, but not for frequency selectivity, suggesting that TFS sensitivity declines with increasing age in the absence of elevated audiometric thresholds or broadened auditory filters, consistent with the studies reviewed earlier.

SRTs were measured for speech in a steady background noise, and noise with spectral and temporal dips, as used by Peters et al. [152]. When the effect of the audiometric threshold was partialled out, SRTs for speech in the modulated noise were correlated with scores on the TFS1 test, but not with scores on the TFS-LF test or the measures of frequency selectivity. The results suggest that reduced sensitivity to TFS can partly explain the speech perception difficulties experienced by HI and by older subjects.

In the study of Füllgrabe et al. [133] described in Sect. 6.4, sensitivity to TFS was measured using the TFS1 test and the TFS-LF test for YNH and ONH subjects with matched audiograms and verbal IQ. Speech intelligibility was also measured, using vowel-consonant-vowel stimuli and sentences presented in quiet and in background sounds.

Performance on the two TFS tasks was significantly poorer for the ONH than for the YNH group. Since auditory filters do not broaden with increasing age when the audiogram is normal [121,122,132], the worse performance of the older group on the TFS tasks cannot be attributed to reduced frequency selectivity.

The intelligibility of consonants in the presence of steady noise and noise that had sinusoidal AM at 5 or 80 Hz was measured using several signal-to-background ratios (SBRs). The result are shown in Fig. 12. The ONH subjects performed more poorly than the YNH subjects with both the steady and AM noise backgrounds. To obtain a global measure of performance, scores were averaged across all noise types and all SBRs. These global speech scores were correlated with scores on the TFS tasks, averaged across all centre frequencies and tasks (expressed as discriminability index, d’, values); for the two groups together, r = 0.76, p < 0.01; for the ONH group only, r = 0.53, p < 0.05; for the YNH group only, r = 0.51, ns.

Figure 13 shows mean scores for the identification of sentences in a single-talker background that was either co-located with the target speech or spatially separated from it. There was a significant correlation between TFS scores and speech scores averaged across conditions: for the two groups together, r = 0.83, p < 0.01; for the ONH subjects only, r = 0.59, p < 0.05; for the YNH subjects only, r = 0.87, p < 0.01.

The results of Füllgrabe et al. [133] are consistent with other studies indicating that sensitivity to TFS declines with increasing age regardless of the presence or absence of hearing loss as measured by the audiogram [153,154]. Furthermore, this reduced sensitivity to TFS is associated with a reduced ability to understand speech. Note that the
ONH subjects in the study of Füllgrabe et al. showed deficits in speech perception that were similar for steady and for fluctuating backgrounds; see Fig. 12. This suggests that the ability to use TFS is not specifically associated with the ability to “listen in the dips” of a fluctuating background sound.

Strelcyk and Dau [155] assessed TFS processing, frequency selectivity and speech reception for YNH subjects and older subjects with high-frequency hearing loss but with near-normal hearing for frequencies up to 1,000 Hz. Frequency selectivity was assessed using the notched-noise method [119] for a centre frequency of 750 Hz. TFS processing was assessed using two binaural tasks, tone lateralisation and binaural masked detection. TFS processing was also assessed using an indirect measure based on FMDLs for low carrier frequencies. Two modulation rates were used, a low rate (2 Hz) for which TFS cues are thought to play a strong role [32,41], and a higher rate (16 Hz) for which TFS cues are thought to play a minimal role. SRTs were measured for sentences presented diotically in several types of background sounds, including both steady and modulated sounds. Given the presentation levels used and the patterns of hearing loss of the HI subjects, it is likely that frequency components in the speech above about 2,000 Hz were not audible.

There were significant correlations between the bin- austral and monaural measures of TFS processing. However, there was no significant correlation between the measures of TFS processing and frequency selectivity. There was no significant correlation between the measures of TFS processing and the SRT for speech in amplitude-modulated noise but there were significant correlations for speech in two other background sounds. One of these was a noise with the same long-term average spectrum as speech that was made to appear to come from one side of the head by use of an ITD. The other was composed of a male and a female talker, both time reversed to make the interfering speech unintelligible. The correlations between the measures of TFS sensitivity and the SRTs are consistent with the idea that speech reception was influenced by sensitivity to TFS cues, although it is not clear why correlations between TFS processing and SRTs were found for only two of the background sounds. Strelcyk and Dau [155] noted that some of the differences between the NH and HI subjects might be related to age.

Overall, the results of the studies reviewed above suggest that both hearing loss and increasing age adversely affect the ability to use TFS information and that this adversely affects the ability to understand speech in the presence of background sounds.

9. CONCLUSIONS

9.1. Frequency Discrimination and Pitch Perception

For sinusoidal signals, several aspects of perception for NH subjects worsen at high frequencies, consistent with a transition from a temporal mechanism based on TFS cues to a place mechanism based on the filtering that takes place on the BM. The sense of tone chroma (musical pitch) is largely lost for frequencies above 5,000 Hz, although the upper limit varies across subjects. Values of DLF/f increase with increasing frequency above 2 kHz but reach a plateau above 8–10 kHz, suggesting a transition from a temporal to a place mechanism at rather high frequencies. FMDLs for \( f_m \geq 10 \) Hz are roughly a constant fraction of \( \text{ERB}_N \) at the same centre frequency, consistent with the predictions of a place model. However, FMDLs for very low \( f_m \) increase relative to \( \text{ERB}_N \) at high frequencies, suggesting that FMDLs are determined primarily by a temporal mechanism for low \( f_m \) and for \( f_c \) below about 4 kHz.

Cochlear hearing loss has adverse effects on pitch perception and the frequency discrimination of sinusoids, which may be caused partly by reduced frequency selectivity and partly by a reduced ability to use TFS information. The sense of tone chroma appears to be reduced by hearing loss, although data on this are sparse. DLF/f values for HI subjects reach a plateau for frequencies above about 4 kHz, which is lower than the value of 8–10 kHz found for NH subjects and is consistent with a reduced ability of HI subjects to use TFS information.

For subjects with DRs, tones producing maximum BM vibration within the DR do not produce a clear pitch. Such tones sometimes have a near-normal pitch value, perhaps reflecting the use of TFS information, but also sometimes have a pitch value that is very different from normal. There may need to be a correspondence between place and temporal information for a clear normal pitch to be perceived.
The detection of FM by HI subjects is strongly adversely affected by the addition of AM to disrupt excitation-pattern cues, consistent with FMDLs being determined largely by a place mechanism. However, when the hearing loss is mild, the disruptive effect of the added AM can be small for very low \( f_m \) and \( f_c = 1 \) kHz, consistent with there being some remaining ability to use TFS

The pitch of complex tones may be derived from place and TFS\(_{BM}\) information about the frequencies of individual low resolved harmonics and from TFS\(_{BM}\) and ENV\(_{BM}\) information derived from places on the BM where the higher unresolved harmonics interfere. The relative role of TFS\(_{BM}\) and ENV\(_{BM}\) information for extracting pitch from unresolved harmonics can be assessed using frequency-shifted inharmonic (I) tones, created by shifting all of the components in a harmonic (H) complex tone upwards by the same amount in Hz, while keeping the spectral “centre-of-gravity” constant by passing the I and H tones through a fixed bandpass filter. The results suggest that, for NH subjects, TFS\(_{BM}\) information can be used when the number, \( N \), of the lowest harmonic within the filter passband lies between 9 and 14. For higher \( N \), only ENV\(_{BM}\) information can be used.

The sensitivity to TFS\(_{BM}\) for \( N \) in the range 9 to 14 can be assessed using the TFS1 task, which measures the ability to discriminate I and H tones. HI subjects generally perform very poorly on the TFS1 task, suggesting a reduced ability to use TFS\(_{BM}\) information. It is possible that this reduced ability stems partly from reduced frequency selectivity, but some HI subjects with mild hearing loss and normal frequency selectivity show very poor performance on the TFS1 task, suggesting that reduced frequency selectivity is not the only factor involved.

Increasing age also adversely affects frequency discrimination and performance on the TFS1 test, even when audiometric thresholds remain within the normal range, suggesting that increasing age adversely affects the ability to use TFS information. However, the deleterious effects of age might partly be a consequence of reduced processing efficiency.

### 9.2. Binaural Processing

Binaural processing of TFS is important for the localization of sounds with frequencies below about 1,400 Hz and for the detection of signals in background sounds whose ITD differs from that of the signal.

Cochlear hearing loss has adverse effects on the ability to use ITDs in both TFS and envelope, but the adverse effects are larger for ITDs in TFS than for ITDs in ENV. The effects are generally larger when the hearing loss is asymmetric than when it is symmetric across ears.

Increasing age is also associated with decreased accuracy of ITD processing, and the decrease starts relatively early in middle age. The effect of age appears to be greater for ITDs in TFS than for ITDs in ENV, but increasing age may also lead to a general reduction in processing efficiency.

MLDs are smaller for HI than for NH subjects and tend to decrease with increasing age, even when audiometric thresholds at the test frequency are normal, suggesting a reduced ability to process binaural TFS.

The reduced binaural processing abilities found for HI subjects and older subjects with or without hearing loss partially account for the difficulties that such subjects experience, relative to YNH subjects, when listening in situations when target speech and interfering sounds come from different directions in space, as is common in everyday life.

### 9.3. Perception of Speech in Background Sounds

The role of ENV and TFS cues in the perception of speech has often been studied using vocoder processing. The interpretation of such studies is complicated, since signal manipulations intended to disrupt TFS\(_{BM}\) and TFS\(_{n}\) cues while preserving ENV\(_{BM}\) and ENV\(_{n}\) cues have adverse effects on ENV\(_{BM}\) and ENV\(_{n}\) cues. Similarly, signal manipulations intended to disrupt ENV\(_{BM}\) and ENV\(_{n}\) cues while preserving TFS\(_{BM}\) and TFS\(_{n}\) cues have adverse effects on TFS\(_{BM}\) and TFS\(_{n}\) cues. Also, when TFS\(_{n}\) or ENV\(_{p}\) cues are removed from the channel signals, TFS\(_{n}\) and ENV\(_{n}\) cues are nevertheless present in the auditory system. However, it does seem possible to some extent to process sounds so as to differentially affect the representation of TFS\(_{BM}\) and TFS\(_{n}\) cues on the one hand and ENV\(_{BM}\) and ENV\(_{n}\) cues on the other hand. The extent to which the different types of cues are disrupted can be assessed using auditory models, although some doubts must remain as to the validity of these models for human listeners.

Overall, the results of studies using vocoder processing to produce ENV speech or TFS speech are consistent with the idea that TFS cues are used for speech perception, especially for the perceptual segregation of target speech from background sounds, and that cochlear hearing loss and increasing age reduce the ability to use TFS cues. Studies of the correlation between speech reception and sensitivity to TFS as measured using non-speech stimuli support the idea that the ability to use TFS cues for speech perception is adversely affected by cochlear hearing loss and age.

### 9.4. Some Practical Implications

The loss of sensitivity to TFS associated with hearing loss and greater age is not compensated by the use of hearing aids. This may partly account for the limited effectiveness of hearing aids, especially when background
sounds are present [156]. Cochlear implants mainly convey ENV information, except perhaps for the most apical channels, and the lack of TFS information may contribute to the very poor ability of cochlear implant users to understand speech in background sounds [5].

One possible cause of reduced sensitivity to TFS is loss of synapses between inner hair cells and auditory neurons, which may occur following noise exposure [157] or with greater age [158,159]. Although such synaptopathy cannot at present be cured, pharmaceutical treatments might be developed to alleviate it or prevent it. Early diagnosis using tests like the TFS1 test and TFS-AF test might be useful in this regard, since treatment is likely to be more effective if it commences before the damage is too severe. Also, since there are large individual variations in susceptibility to the effects of noise exposure, early diagnosis of hearing damage may be useful for detecting individuals who are at greatest risk, so that their hearing can be protected.

REFERENCES


[148] M. K. Pichora-Fuller and B. A. Schneider, “Masking-level...


