Published in final edited form as:

J Acoust Soc Am. 2015 May ; 137(5): 2687–2697. doi:10.1121/1.4919315.

The role of excitation-pattern cues in the detection of frequency shifts in bandpass-filtered complex tones

Frederic Marmel^{a)}, Christopher J. Plack, and Kathryn Hopkins

School of Psychological Sciences, Manchester Academic Health Science Centre, The University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

Robert P. Carlyon and Hedwig E. Gockel

Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, United Kingdom

Brian C. J. Moore

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

Abstract

One task intended to measure sensitivity to temporal fine structure (TFS) involves the discrimination of a harmonic complex tone from a tone in which all harmonics are shifted upwards by the same amount in hertz. Both tones are passed through a fixed bandpass filter centered on the high harmonics to reduce the availability of excitation-pattern cues and a background noise is used to mask combination tones. The role of frequency selectivity in this “TFS1” task was investigated by varying level. Experiment 1 showed that listeners performed more poorly at a high level than at a low level. Experiment 2 included intermediate levels and showed that performance deteriorated for levels above about 57 dB sound pressure level. Experiment 3 estimated the magnitude of excitation-pattern cues from the variation in forward masking of a pure tone as a function of frequency shift in the complex tones. There was negligible variation, except for the lowest level used. The results indicate that the changes in excitation level at threshold for the TFS1 task would be too small to be usable. The results are consistent with the TFS1 task being performed using TFS cues, and with frequency selectivity having an indirect effect on performance via its influence on TFS cues.

I Introduction

The peripheral auditory system acts like an array of bandpass filters that decompose incoming complex sounds into a series of narrowband signals. For harmonic complex tones, this filtering process results in approximately the first five to eight harmonics being “resolved” (Plomp, 1964; Moore and Gockel, 2011), i.e., each harmonic gives rise to a distinct peak in the profile of excitation along the basilar membrane (BM). Higher harmonics are said to be “unresolved,” i.e., several harmonics interact at a given place on the BM. The resulting waveform at each place on the BM can be considered as a slowly

^{a)} frederic.marmel@gmail.com.

fluctuating envelope superimposed on a more rapidly varying temporal fine structure (TFS). The temporal envelope and TFS are represented in the phase-locked activity of auditory nerve fibers (Joris *et al.*, 2006). For normal-hearing (NH) listeners, the processing of TFS has been argued to be important for pitch perception, masking, and speech perception (Moore, 2008, 2014), especially for segregating target sounds from background sounds (Hopkins and Moore, 2009; Ardoint and Lorenzi, 2010; Strelcyk and Dau, 2009; Jackson and Moore, 2013); see, however, Swaminathan and Heinz (2012).

Cochlear hearing loss (CHL) and increasing age have been shown to impair performance on a number of tasks that have been interpreted as reflecting TFS processing (Moore, 2008, 2014; Füllgrabe *et al.*, 2015), and poor TFS processing might contribute to the difficulties experienced by people with hearing loss and older people in understanding speech in the presence of background sounds (Lorenzi *et al.*, 2006; Lorenzi *et al.*, 2009; Hopkins *et al.*, 2008; Ardoint *et al.*, 2010; Hopkins and Moore, 2010, 2011; Füllgrabe *et al.*, 2015). This has led to interest in tests for assessing the ability to process TFS, including the TFS1 test (Moore and S k, 2009a; S k and Moore, 2012) described below. However, it has been suggested that performance of the TFS1 test might depend on the use of excitation-pattern cues rather than TFS cues. The main purpose of the present study was to assess the possible role of excitation-pattern cues in performance of the TFS1 test.

The TFS1 test involves the discrimination of a harmonic complex tone (H), with a fundamental frequency F_0 , from a tone in which all harmonics are shifted upwards by the same amount in hertz, resulting in an inharmonic tone (I) (Hopkins and Moore, 2007, 2011; Moore and S k, 2009a; S k and Moore, 2011; Jackson and Moore, 2014). Both tones have an envelope repetition rate equal to F_0 , but the tones differ in their TFS. To reduce the availability of excitation-pattern cues, all tones are passed through a fixed bandpass filter, usually centered at $11F_0$. The components falling within the passband of the filter are separated by less than one auditory-filter bandwidth, and are assumed to be unresolved. A background noise is used to mask combination tones and to limit the audibility of components falling on the sloping skirts of the filter. The amount of frequency shift is adaptively varied to determine a threshold value.

The TFS1 task is based on the theory that, for complex tones with unresolved components, the pitch can be derived from the time interval between peaks in the TFS on the BM close to adjacent envelope maxima (de Boer, 1956; Schouten *et al.*, 1962; Moore *et al.*, 2009; Jackson and Moore, 2014). If the components are shifted upwards in frequency, then that time interval gets smaller, and, consistent with the hypothesis, NH listeners usually report hearing an upward shift in pitch. Hence, thresholds measured using the TFS1 task have often been taken as indicating sensitivity to changes in TFS. However, it has been suggested that the task might be performed using residual excitation-pattern cues, namely shifts in the shallow pattern of ripples in the excitation patterns evoked by the H and I tones (Micheyl *et al.*, 2010). This could account for the poor performance of listeners with CHL on the TFS1 task (Hopkins and Moore, 2007, 2011), since CHL is usually associated with poorer-than-normal frequency selectivity (Glasberg and Moore, 1986) and this would reduce the magnitude of the ripples in the excitation pattern.

One way of assessing the role of frequency selectivity in the TFS1 task is to manipulate the level of the stimuli. The bandwidth of the auditory filter increases with increasing level (Moore and Glasberg, 1987; Glasberg and Moore, 2000; Baker and Rosen, 2006; Unoki *et al.*, 2006). As a result, the ability to “hear out” individual partials from complex tones worsens at high levels (Moore *et al.*, 2006), and patterns of forward masking produced by complex tones show reduced peak-to-valley ratios at high levels (Houtgast, 1974). If performance of the TFS1 task depends on frequency selectivity, performance should worsen with increasing sound level. This prediction was tested by Moore and S k (2009a, 2011). They measured performance on the TFS1 task for NH listeners, using complex tones with levels ranging from 10 to 50 dB sensation level (SL; they specified the SL as the overall level of the complex tone relative to each listener’s absolute threshold for detecting a pure tone at the center frequency of the bandpass filter used in the TFS1 task). They found no significant effect of level, and concluded that performance of the TFS1 task probably did not depend on the use of excitation-pattern cues. However, the study of Moore and S k (2009a) used headphones (Sennheiser HD580, Sennheiser electronic GmbH & Co. KG, Wedemark, Germany) with a relatively “open” design, so there is a possibility that, for high levels, the task was performed via “cross-hearing” of the stimuli in the non-test ear; the lower level in the non-test ear might have allowed performance to be maintained for high levels in the test ear. The study of Moore and S k (2011) used insert earphones (Etymotic ER2, Etymotic Research, Inc., Elk Grove Village, IL) for which cross-hearing is unlikely to have played a role. However, that study used bandpass filters centered at very high frequencies (above 6 kHz), and many researchers are skeptical about the possibility of TFS cues being used at such high frequencies. Also, a mechanism based on TFS may be used to detect frequency modulation at low rates, but this mechanism appears to be ineffective for a carrier frequency of 6 kHz, irrespective of the modulation frequency (Moore and S k, 1995, 1996; S k and Moore, 1995; Ernst and Moore, 2010, 2012). Therefore, the present study reassessed the effect of level on performance of the TFS1 task, using stimuli with a wide range of center frequencies and levels. Noise was presented to the non-test ear during performance of the task, to eliminate the possibility of cross-hearing.

Experiment 1 evaluated performance of the TFS1 task at a low and high level over a wide range of f_0 s. The results showed higher thresholds at the high level than at the low level, in contrast to the findings of Moore and S k (2009a, 2011). To further explore the discrepancy with the results of Moore and S k (2009a, 2011), experiment 2 measured performance on the TFS1 task using a single f_0 of 250 Hz at several levels, covering a wide range. A significant worsening in performance was found at the two highest levels used. As a further means of evaluating the role of excitation-pattern cues in the TFS1 task, experiment 3 estimated the changes in excitation level produced by the frequency shift using a forward-masking task. One of the complex tones used in the TFS1 task was used as a forward masker, and the threshold for detecting a pure tone with a frequency falling on the lower skirt of the bandpass filter was measured as a function of the frequency shift. The estimated changes in excitation level were very small, probably too small to provide usable cues in the TFS1 task. Therefore we argue that the TFS1 task is probably performed using TFS cues rather than excitation-pattern cues, but that frequency selectivity may have an indirect effect on performance of the task via its influence on TFS cues.

II Experiment 1: Effect of Level on the Discrimination of Harmonic and Frequency-Shifted Tones

A Listeners

Ten NH listeners completed experiment 1. They were between 19 and 35 yrs old (average: 25 yrs). Absolute thresholds were measured using a three-alternative forced-choice (3AFC) task and were converted from dB sound pressure level (SPL) to dB hearing level (HL). All listeners had absolute thresholds less than 15 dB HL in their test ear at octave frequencies between 250 and 8000 Hz. All the procedures of the study were approved locally by the School of Psychological Sciences Ethics Committee, University of Manchester, and nationally by the National Health Service North West 3 Research Ethics Committee.

B Stimuli

1 Harmonic and frequency-shifted complex tones—The stimuli were similar to, but not exactly the same as, those used in previous studies employing the TFS1 test. Differences are noted below. Stimuli consisted of H and I tones presented in threshold-equalizing noise (TEN, Moore *et al.*, 2000). The F_0 s were 60, 100, 250, and 625 Hz. To generate the H tones, 30 harmonics were added, each starting in sine phase. Moore and S k (2009a) and S k and Moore (2012) recommended the use of random phase to reduce the salience of possible envelope cues, but Jackson and Moore (2014) showed that there was no significant effect of component starting phase on performance of the TFS1 task. I tones were formed in the same way as H tones except that each component was shifted upwards in frequency by the same amount in hertz. All complex tones were passed through a fixed bandpass filter, designed using the FIR2 function in MATLAB (MathWorks, Natick, MA), with 256 taps at a 48-kHz sampling rate. The filter had a central flat region with a width of $5F_0$ and was centered on the 11th harmonic. Thus the start and end frequencies of the passband corresponded to $8.5F_0$ and $13.5F_0$, respectively. In the region of the skirt just below the passband, the slope was about 30 dB/oct, similar to that used by Moore and S k (2009a) and S k and Moore (2012). For lower frequencies the slope increased, but for frequencies below about $7F_0$ the components would have been masked by the background TEN (see below for details of the signal-to-TEN ratio). Each complex tone was 500 ms long, including 30-ms raised-cosine onset and offset ramps.

2 TEN—The complexes were gated synchronously with a TEN (extending from 50 to 16 000 Hz) in order to mask combination tones and to limit the audibility of frequency components falling on the skirts of the bandpass filter. Moore and S k (2009a) and S k and Moore (2012) used a TEN that was turned on 300 ms before the first tone burst in a trial and was turned off 300 ms after the end of the last tone burst. Here, the TEN level at 1 kHz was set to either 20 dB SPL/ERB_N (low-level condition) or 60 dB SPL/ERB_N (high-level condition), where ERB_N stands for the average value of the equivalent rectangular bandwidth of the auditory filter for young NH listeners tested at low sound levels (Glasberg and Moore, 1990). To prevent cross-hearing, an independent TEN was presented to the contralateral ear, at the same level as for the test ear. The TEN was uncorrelated at the two

ears to avoid the release from masking that could otherwise have occurred at some frequencies.

To set the signal-to-TEN ratio for the complex tones, the detection threshold was first measured for a pure tone presented monaurally in the TEN with frequency equal to the center frequency of the passband. This was done using a three-interval, 3AFC task with a two-down one-up adaptive procedure. In the same way as for the main experiment, an uncorrelated TEN at the same level was presented in the contralateral ear. Three estimates were obtained for each level and F_0 and the three were averaged to obtain the final estimate. The masked thresholds were typically 3–4 dB (mean 3.9 dB) below the TEN level/ ERB_N and were almost invariant across frequency, as intended.

For the complex tones, each component in the passband was set to a level 12.5 dB above the measured threshold for the pure tone. The overall level of the complex tone was about 8 dB above the level of the central component in the passband, so the overall signal-to-TEN ratio in the main experiment, expressing the TEN level in dB/ ERB_N , was about $-3.9 + 8 + 12.5$ dB, i.e., 16.6 dB. This is slightly higher than the ratio of 15 dB recommended by Moore and S k (2009a) and S k and Moore (2012). However, Jackson and Moore (2014) showed that a 3-dB change in signal-to-TEN ratio had almost no effect on performance of the TFS1 task. Examples of the spectra of H and I complex tones (for $F_0 = 250$ Hz) presented in TEN are shown in Fig. 1.

3 Equipment—All stimuli were generated using a personal computer with an E-MU 0202 USB (Creative Technology Ltd., Singapore) sound card, using a sampling rate of 48 kHz and a bit depth of 24. Stimuli were presented through Sennheiser HD 650 Headphones.

C Procedure

The experimental procedure was controlled via custom MATLAB software. A three-interval, 3AFC task was used. Two intervals contained the H tone and one interval (chosen at random) contained the I tone. The inter-stimulus interval was set to 500 ms. Listeners were asked to identify the interval that was different from the other two by pressing a key on a computer keyboard. Visual feedback indicated whether the response was right or wrong. The frequency shift, expressed as a percentage of the F_0 , was varied adaptively using a geometric track with a two-down, one-up rule (tracking the 70.7% point on the psychometric function). The frequency shift at the start of the procedure was set to 50% F_0 , which leads to the largest possible difference between the H and I tones. This shift was the maximum allowed in the adaptive procedure. Any response calling for a larger shift led to the presentation of another trial with a shift of 50% F_0 . A block of trials consisted of 16 reversals (changes in track direction). The step size was a factor of 2 for the first 4 reversals and a factor of 1.414 for the remaining 12 reversals. For each block, the threshold was taken as the geometric mean of the frequency shift at the last 12 reversals. A testing session consisted of eight blocks, corresponding to the four F_0 s and two levels. Listeners completed four testing sessions. The first session was treated as practice and the geometric averages of the thresholds for the last three sessions were taken as the final estimates. The order of conditions was randomized.

D Statistics

Statistical tests were performed on the log-transformed thresholds from the adaptive procedure. A within-subjects analysis of variance (ANOVA) was performed with factors F_0 (60, 100, 250, and 625 Hz) and level (20 and 60 dB/ERB_N). In this and subsequent ANOVAs, the degrees of freedom used to calculate p values for the factor F_0 were corrected using the Greenhouse-Geisser correction; however the original (uncorrected) degrees of freedom are reported. All statistical analyses were carried out with SPSS (SPSS, Chicago, IL).

E Results

Individual and averaged shift thresholds are plotted in Fig. 2. The pattern of the results varied across listeners, but most listeners had higher thresholds at the high level than at the low level for some or all F_0 s. There were only a few cases where thresholds were lower for the high level than for the low level (S4 and S5 for $F_0 = 60$ Hz). Listener S3 showed performance close to chance (thresholds $\approx 50\%$ F_0) for all conditions, and the results for this listener were excluded from the mean and from further analyses.¹ An ANOVA revealed a main effect of level, with worse performance at the high level than at the low sound level [$F(1, 8) = 35.8$; $p < 0.001$]. A main effect of F_0 was also observed [$F(3, 24) = 5.33$; $p < 0.05$]. The interaction between level and F_0 was not significant. *Post hoc t*-tests using the Bonferroni correction showed poorer performance for $F_0 = 625$ Hz than for $F_0 = 250$ Hz ($p < 0.05$). The thresholds for other pairs of F_0 s did not differ significantly. The pattern of results across F_0 s is similar to that found previously using the TFS1 test (Moore *et al.*, 2009; Jackson and Moore, 2014).

F Discussion

The observed influence of sound level on thresholds is consistent with the idea that frequency selectivity affects performance on the TFS1 task. It is noteworthy that the listeners could perform the task for the highest F_0 of 625 Hz, since the passband in this case was centered at 6875 Hz and the lowest audible component would have had a frequency of about 4375 Hz. Indeed, at the lower level, performance for the F_0 of 625 Hz was similar to that for $F_0 = 100$ Hz, where the lowest audible component would have had a frequency of about 700 Hz. Phase locking is generally assumed to be very weak or absent for frequencies above 4000 Hz (Johnson, 1980; Palmer and Russell, 1986; Verschooten, 2013; Verschooten and Joris, 2014), and much weaker than at 700 Hz. However, previous research has also shown that the TFS1 task can be performed when all audible components have frequencies above 5000 Hz (Moore and S k, 2009b, 2011; Jackson and Moore, 2014). These results imply either that some TFS cues remain at high frequencies (Heinz *et al.*, 2001; Recio-Spinoso *et al.*, 2005) and, at the lowest level tested here, are sufficient to perform the TFS1 task as well as at much lower frequencies, or that, at least for high center frequencies, the task is performed using other cues, such as excitation-pattern cues.

¹To explore the origin of this poor performance, the ability of S3 to discriminate the frequency of pure tones with frequencies of 660 and 2750 Hz was measured. The tones were presented at 80 dB SPL in the presence of dichotic TEN at 40 dB/ERB_N. The frequency discrimination thresholds (16.6% and 13.1% for 660 and 2750 Hz) were much higher than usually found for NH listeners, suggesting that this listener had a general problem with frequency discrimination.

The effect of level for the NH listeners found here appears to differ from the findings of two studies that showed no significant effect of level on performance of the TFS1 task (Moore and S k, 2009a, 2011). However, the range of levels used differed across studies. In the present study, the highest overall level of the complex tone (not including the level of the TEN) was about 76.6 dB SPL. In Moore and S k (2009a), the highest level of the complex tone was always below 65 dB SPL. The highest level of the complex tone used by Moore and S k (2011) was about 80 dB SPL, which is slightly higher than used here. However, the highest SL of the complex tones used by Moore and S k (2011) was 50 dB, which is lower than the highest SL used here, which, averaged across listeners, varied from about 66 to 75 dB, depending on the $F0$. Hence, the discrepancy across studies could be related to the difference in the highest SL used.

To assess whether performance of the TFS1 task worsens only at very high levels, experiment 2 used the same stimuli and methods as experiment 1, but included intermediate levels.

III Experiment 2: Discrimination of Harmonic and Frequency-Shifted Tones at Intermediate Sound Levels

A Method

Only an $F0$ of 250 Hz was used. The low- and high-level conditions of experiment 1 at that $F0$ were tested again, and three intermediate conditions with the TEN level set to 30, 40, and 50 dB/ERB_N were added. As in experiment 1, the complex tones were presented with each component in the passband 12.5 dB above the masked threshold for a pure tone in the TEN with frequency corresponding to the central harmonic in the stimulus passband ($11F0$). All other aspects of the stimuli and method were the same as for experiment 1. Eight of the 10 listeners from experiment 1 participated. Six completed three testing sessions and the averaged thresholds were taken as the final estimates. The two others (S9 and S10) completed only one testing session, so only one threshold estimate was obtained for each condition. The log-thresholds from the adaptive procedure were analyzed with a repeated-measures ANOVA with level as the single factor.

B Results

Individual and averaged shift thresholds are plotted in Fig. 3. Consistent with the results of experiment 1 for $F0 = 250$ Hz (re-plotted as triangles in Fig. 3), most listeners had higher thresholds at 60 than at 20 dB/ERB_N (the only listener from experiment 1 who did not have higher thresholds at 60 than at 20 dB/ERB_N for $F0 = 250$ Hz, S6, was not available for this experiment). Although thresholds increased monotonically with level only for S10, all listeners showed some sign of a level effect not restricted to the highest level, since thresholds were always higher at 50 dB/ERB_N than at either 40 or 30 dB/ERB_N, but not always both. The ANOVA revealed a significant effect of level [$F(4, 28) = 11.1$; $p < 0.01$]. *Post hoc t*-tests with Bonferroni correction showed that the mean threshold at 60 dB/ERB_N was higher than the mean thresholds at 30 dB/ERB_N ($p = 0.026$), 40 dB/ERB_N ($p = 0.007$), and 50 dB/ERB_N ($p = 0.047$), and that the mean threshold at 50 dB/ERB_N was higher than at 30 dB/ERB_N ($p = 0.037$), but not higher than at 20 or 40 dB/ERB_N ($p > 0.05$).

C Discussion

The thresholds showed a significant effect of level, consistent with experiment 1. Performance worsened for the two highest levels used, 50 and 60 dB/ERB_N. While Moore and S k (2009a, 2011) found no significant effect of level, they did not use as wide a range of levels as used here. To compare our results with those of Moore and S k (2009a), which were obtained with similar stimuli to those used here, the levels for both studies were expressed as the overall level of the complex tone in dB SPL. This level is shown at the top of the bottom-right panel in Fig. 3. For our stimuli, the overall level of the complex tones was, on average, 16.6 dB above the level of the TEN in dB/ERB_N. Moore and S k (2009a) expressed the levels of their stimuli in terms of the “SL” of the complex tones, which they defined as the overall level of the tones relative to the absolute threshold for a sinusoid at the center frequency of the bandpass filter. For the data of Moore and S k (2009a), this center frequency was 2200 Hz, and the mean absolute threshold was 12.5 dB. Hence the mean level of the complex tone in dB SPL was the SL + 12.5 dB. The geometric mean data of Moore and S k (2009a) are shown as open squares in the bottom-right panel of Fig. 3. Their thresholds overall were higher than those found here, which might reflect the higher tone-to-TEN ratio, the longer signal duration, or the larger amount of training used here. However, the variation of thresholds with level is similar for the present study and that of Moore and S k (2009a) over the range tested in both experiments. Three of the ten listeners tested by Moore and S k (2009a) and two of the ten listeners tested by Moore and S k (2011) showed higher thresholds at the highest level than at the lower levels. In the present study, most but not all of the listeners showed worse performance at the two higher levels relative to the levels of 30 or 40 dB/ERB_N, but not relative to the level of 20 dB/ERB_N.

IV Experiment 3: Effects of Sound Level on Excitation-Pattern Cues to Frequency Shift

Experiment 3 was intended to provide a quantitative estimate of the change in excitation level produced by the frequency shift of the I tones relative to the H tones as a function of stimulus level. The magnitude of excitation-level differences between the H and I tones was estimated by using the complex tones plus the TEN as maskers. Forward masking was used in order to avoid interactions between the masker and the signal (e.g., beats or suppression) that can affect the signal threshold in simultaneous masking. A similar method has been used in the past to estimate the excitation patterns of harmonic complex tones (Plomp, 1964; Moore and Glasberg, 1983a) and of vowel sounds (Moore and Glasberg, 1983b). The masker-signal interval was chosen separately for each masker level and listener such that the signal level at threshold was relatively low, but still above absolute threshold. This was done so that the signal level would fall within the linear portion of the basilar-membrane input-output function (Robles and Ruggero, 2001). This means that the signal threshold should be linearly related to the excitation produced by the masker at the place on the BM corresponding to the signal frequency (Plack and Oxenham, 1998).

The threshold of a pure tone forward masked by one of the complex tones in the TEN was measured when the frequency of the pure tone coincided with that of a component in the masker or fell between that of two components, corresponding to a peak or a trough of the

ripples in the excitation pattern of the masker. The ripples were expected to be largest for a component falling on the lower slope of the fixed bandpass filter, provided that the component was above the masked threshold imposed by the TEN (Jackson and Moore, 2014). Calculations based on the excitation-pattern model of Glasberg and Moore (1990) suggested that the ripples would be largest around the eighth harmonic. Hence, testing focused on the masking of a pure tone corresponding to the eighth harmonic. To keep testing time reasonable, a single F_0 of 250 Hz was used, together with 4 of the 5 levels used in experiment 2 (TEN at 20, 40, 50, and 60 dB/ERB_N). The forward-masked threshold was measured using the H tone (0% frequency shift) and the I tone (with frequency shifts of 25% F_0 , 50% F_0 , and 75% F_0) as maskers. It was predicted that the threshold would be highest for the 0% shift (when the signal frequency fell close to a peak in the excitation pattern of the masker) and lowest for the shift of 50% F_0 (when the signal frequency fell close to a trough in the excitation pattern of the masker); the magnitude of the difference between these two thresholds gives an estimate of the change in excitation level produced by the frequency shift, which in turn is related to the ripple depth in the excitation patterns of the H and I tones.

A Method

Eight listeners were tested. Two were tested only with the lowest and highest levels, while the other six were tested using TEN levels of 20, 40, 50, and 60 dB/ERB_N. All listeners had participated in the previous experiments, and the complex tone levels relative to the TEN level were the same as in those experiments. The complex tones were generated as in the previous experiments except that they were 100-ms long, including 2-ms raised-cosine onset and offset ramps. The maskers (complex tone + TEN) were presented to both ears, with the same TEN at each ear, while the signal was presented to one ear, the one in which the complex tone had been presented in the previous experiments. The diotic presentation of the masker (complex tone + TEN) was intended to eliminate any confusion of the signal with the masker. The signal was a pure tone with a frequency of 2000 Hz, corresponding to the eighth harmonic. The signal had 5-ms cosine-squared onset and offset ramps and no steady portion. The masker-signal delay (the silent interval between masker offset and signal onset) was 0 ms for the level of 20 dB/ERB_N. For each other level, the masker-signal delay was adjusted for each listener to find a delay that would give an amount of masking similar to that obtained at 20 dB/ERB_N. The masker-signal delays for levels higher than 20 dB/ERB_N were chosen by running the 50% F_0 -shift condition repeatedly with the masker-signal delay increased by 10 ms on each repetition, and choosing the delay that gave the closest threshold to the one measured at 20 dB/ERB_N. For a given level and listener, the same delay was used for all frequency shifts. On average, the masker-signal delays were equal to 22, 37, and 49 ms for the TEN levels of 40, 50, and 60 dB/ERB_N, respectively.

The masked threshold of the signal was measured with a 3AFC procedure. The level of the signal was varied adaptively (two-down, one-up rule), using a step size of 4 dB for the first 4 reversals, and 1 dB for the remaining 12 reversals. For each level, condition, and listener, ten masked thresholds were collected. For each listener, the mean and standard deviation of the thresholds were calculated across all levels and frequency shifts, and masked thresholds

more than two standard deviations away from the mean were discarded (3.25% of the total number of collected thresholds).

Since the masker-signal delay varied across masker level, effects of masker level on the masked thresholds were not meaningful. To compensate for the effect of masker level, the collected thresholds were first normalized across level (Masson, 2003). For each level L , normalized thresholds were derived from the obtained thresholds by: (1) Adding the grand mean threshold averaged across all levels, listeners and frequency shifts and (2) subtracting the mean threshold averaged across all listeners and frequency shifts for level L .

The masker-signal delay for a given level also varied across listeners (except for the lowest level). To reduce the influence of differences across listeners in the signal level at masked threshold for a given masker level, the thresholds were then normalized across listeners. For each listener P and each level L , final normalized thresholds were derived from the thresholds that were already normalized for level by: (1) Adding the mean threshold averaged across all listeners and frequency shifts and (2) subtracting the mean threshold averaged across all frequency shifts for listener P .

B Results

Individual and averaged normalized thresholds are plotted in Fig. 4. Note that the range of numbers on the ordinate is only 4 dB. The differences in masked thresholds across frequency shifts were small overall. The maximum difference between thresholds for the 0 and 50% $F0$ shifts for a given listener and level ranged between -0.7 dB (for S1 at 50 dB/ERB_N) and 2.7 dB (for S5 at 20 dB/ERB_N). For the three higher levels, no consistent change in masked threshold with frequency shift was apparent, while for the lowest level the threshold tended to be lower for the frequency shift of 50% $F0$ than for the frequency shift of zero, as expected. The mean differences between thresholds for the 0 and 50% $F0$ shifts were 1.1, 0.4, 0.0, and 0.4 dB for the levels of 20, 40, 50, and 60 dB/ERB_N, respectively. One-sample t -tests were used to assess whether the difference between mean thresholds for the 0 and 50% $F0$ shifts was significant for each level. The outcome was significant only for the lowest level [$t(7) = 2.74$, $p < 0.05$, two-tailed]. The results of the t -tests were used to estimate 95% confidence intervals for the mean difference in threshold between the 0 and 50% $F0$ frequency shifts. These confidence intervals were: [0.15 2.06], $[-0.38 \ 1.12]$, $[-0.69 \ 0.66]$, and $[-0.14 \ 0.83]$ dB, for the levels of 20, 40, 50, and 60 dB/ERB_N, respectively.

As there were missing data for the two listeners who were tested only at the two extreme levels, masked thresholds were analyzed with a linear mixed-model procedure. A model with correlated terms was built in SPSS. The normalized thresholds (averaged over the ten repetitions per listener and condition) were the dependent variable. The level and the frequency shift of the masker were the factors and were analyzed as fixed effects. The restricted maximum likelihood solution was used to estimate the model parameters. The analysis showed a main effect of frequency shift [$F(3, 34.8) = 6.5$; $p = 0.001$] but no significant interaction between level and frequency shift [$F(9, 22.8) = 1.8$; $p = 0.12$]. The estimated marginal means of the fitted model were compared using the Bonferroni confidence interval adjustment. There was a significant difference between the 0 and 50% $F0$ frequency shifts averaged across levels ($p = 0.001$; 95% confidence interval for the

difference = [0.16 0.81]), and a significant difference between the 75% F_0 and 50% F_0 frequency shifts averaged across levels ($p = 0.05$; 95% confidence interval for the difference = [0.00 0.74]).

C Discussion

In experiment 3 we attempted to measure directly the excitation level differences produced by frequency shifting the complex tones. It was assumed that the excitation level differences between the H and I tones at 2000 Hz would be approximately equal to the differences in masked threshold. The mean change in masked threshold between the conditions with frequency shifts of 0 and 50% F_0 was small, being 0.4 dB or less for the three higher levels, and 1.1 dB for the lowest level. The mixed-model analysis showed that the effect of frequency shift was significant, but the interaction of frequency shift and level just failed to reach significance.

One difficulty in interpreting the masked threshold differences comes from the fact that, although the masker-signal delays were adjusted with the goal that the signal would fall within the linear portion of the BM response function, some signals might still have been compressed around the masked threshold levels. The collected masked thresholds (before normalization) were all between 19 and 44 dB SPL, with the first, second (median), and third quartiles being at 27, 30, and 33 SPL, respectively. Physiological measurements of BM vibration have shown a linear response for levels below 20 dB SPL, a transition region between 20 and 40 dB SPL, and a highly compressive region for levels higher than 40 dB SPL (Ruggero *et al.*, 1997). It is thus possible that some signals were compressed to a small extent. In that case, the differences between the masked thresholds for the 0 and 50% F_0 frequency shifts could be larger than the associated excitation level differences, and hence the results of experiment 3 might overestimate the excitation level changes associated with the frequency shift of the complex tones. In other words, the changes in excitation level might be even smaller than estimated from the results of experiment 3.

When interpreting experiment 3, it is necessary to consider how large the change in excitation level with frequency shift would need to be to account for performance in the TFS1 task. The smallest detectable change in excitation level over a small region of the excitation pattern is probably 2 dB or more (Moore *et al.*, 1989; Buus and Florentine, 1995; Moore and S k, 2009b). However, there is evidence that excitation-pattern cues can be combined across several center frequencies (Florentine and Buus, 1981; Bernstein and Green, 1987; Moore and S k, 1994; Green, 1992), in which case the largest change in level at threshold for any single component may be as small as 0.4 dB. Micheyl *et al.* (2010) showed that H and I tones may be discriminable on the basis of excitation-pattern differences even when the maximum difference in excitation level at any single point is only 1 dB or perhaps even less. However, it should be remembered that our stimuli were presented in TEN, and that the random amplitude fluctuations in the TEN lead to random changes in excitation level and excitation-pattern shape that would limit the smallest change in excitation level or in ripple pattern that could be detected (Jackson and Moore, 2014).

The upper limit of the change in excitation level produced by a frequency shift corresponding to the measured threshold (as determined in experiment 2) can be estimated

from the 95% confidence intervals derived from the results of experiment 3. For example, for the level of 60 dB/ERB_N, the change in excitation level for a frequency shift of 50% f_0 was less than 0.83 dB. Since the mean frequency shift at threshold for this level was 15% f_0 , the mean change in excitation level at threshold would have been less than about 0.25 dB ($0.83 \text{ dB} \times 15\% f_0 / 50\% f_0$). For the levels of 50, 40, and 20 dB/ERB_N, the corresponding upper limits of the changes in excitation level at threshold were 0.12, 0.15, and 0.27 dB. Hence, even according to the model of Micheyl *et al.* (2010), the changes in excitation level would have been too small to support performance of the TFS1 task. We conclude that performance of the TFS1 task probably depends on the use of TFS cues rather than excitation-pattern cues.

V General Discussion

Jackson and Moore (2014) have also presented evidence indicating that excitation-pattern cues are not used to perform the TFS1 task. They investigated the effect of randomly varying the level of each component in the H and I tones for each and every tone. The range of level variation was ± 3 or ± 5 dB. An excitation-pattern model based on the detection of shifts in the pattern of ripples along the tonotopic axis predicted that the level randomization would substantially impair performance, but measured performance was hardly affected by the level randomization.

Assuming that performance of the TFS1 task depends on the use of TFS cues rather than excitation-pattern cues, an explanation is needed for the worsening in performance at high levels found in experiments 1 and 2. It is possible that performance for the TEN level of 60 dB/ERB_N used here was influenced by the loudness of the stimuli. In the present study, independent TEN was presented at the same level to each ear. For a given TEN level, this would have led to the TEN having a loudness that was about 1.5 times that for monaural presentation (as used in the studies of Moore and S k), corresponding to a difference in loudness level of about 6 phons (Moore and Glasberg, 2007). According to the loudness model of Moore and Glasberg (2007), a TEN with a level of 60 dB/ERB_N at each ear would have a loudness level of about 97 phons. Also, the complex tones used here had components added in cosine phase; in contrast, Moore and S k (2009a, 2011) used tones with components added in random phase. Cosine phase tends to lead to a greater loudness than random phase for complex tones at medium and high levels (Gockel *et al.*, 2002). Hence, some listeners in the present study might have found both the TEN and the complex tones to be unpleasantly loud and/or annoying at the highest level used.

It is not clear whether greater loudness/annoyance would have adversely affected performance. While a distinct worsening in discrimination performance at high levels has sometimes been found for psychoacoustic tasks (intensity discrimination of pure tones, Viemeister and Bacon, 1988), other studies have revealed no change or a slight improvement (intensity discrimination of white noise, Miller, 1947; frequency discrimination of pure tones, Wier *et al.*, 1977). In a study using stimuli and a task resembling ours (f_0 discrimination of complex tones filtered so as to contain unresolved harmonics, presented in TEN with levels of 10, 40 and 65 dB SPL/ERB_N), Bernstein and Oxenham (2006) found that thresholds were, on average, about 20% higher for the 65- than for the 40-dB/ERB_N

level. The change with level was smaller than found here, but was consistent across listeners. It is noteworthy that one of the listeners of Bernstein and Oxenham complained that the stimuli at the level of 65-dB/ERB_N were uncomfortably loud, and that listener was tested using a lower level.

Another possible explanation for the worsening in performance at high levels found in experiments 1 and 2 is that it reflects an indirect effect of frequency selectivity on the use of TFS cues. There are at least three possible indirect effects.

First, when frequency selectivity is reduced, the TFS at the outputs of the auditory filters becomes more complex and more rapidly time varying, and this might make the TFS harder to “decode” (Moore and S k, 1996; Moore, 2014). Consistent with this, Jackson and Moore (2014) showed that performance of the TFS1 task was better when the passband of the bandpass filter used in the task had a width of $F0$ than when it had a width of $5F0$; when the stimuli themselves are passed through a narrow bandpass filter, the TFS remains relatively simple and slowly varying even when the listener has relatively poor frequency selectivity.

A second indirect effect of reduced frequency selectivity at high levels is that it might disrupt mechanisms for “decoding” TFS information based on comparison of the phase of BM responses at different places (or comparison of the timing of nerve spikes in neurons with different characteristic frequencies) (Carney *et al.*, 2002; Carlyon *et al.*, 2011; Heinz *et al.*, 2010).

A third indirect effect is that reduced frequency selectivity at higher levels might have resulted in a slightly poorer signal-to-noise ratio at the output of the auditory filters centered just below the stimulus passband. To investigate this possibility, we measured masked thresholds in TEN with levels of 20 and 60 dB/ERB_N for pure-tone signals with frequencies corresponding to harmonics 7 and 8, for 7 NH listeners and $F0 = 100, 250, \text{ and } 625 \text{ Hz}$. The thresholds, expressed as the signal level relative to the TEN level, were not higher at the high level than at the low level, suggesting that a worsening in effective signal-to-noise ratio was not responsible for the level effect found in experiments 1 and 2.

Overall, a reasonable interpretation of the present results is as follows. The results of experiment 3 suggest that the changes in excitation level produced by a frequency shift corresponding to the measured threshold in the TFS1 task were too small to be used to perform the TFS1 task. This supports the conclusion of Jackson and Moore (2014) that the TFS1 task is not performed using excitation-pattern cues. The effect of level probably reflects an indirect influence of frequency selectivity on performance of the TFS1 task, although we cannot rule out the possibility that part of the worsening at the highest level used was caused by the unpleasant loudness of the stimuli.

As noted earlier, the TFS1 task could be performed for the highest $F0$ of 625 Hz, even though the passband in this case was centered at 6875 Hz and the lowest audible component would have had a frequency of about 4375 Hz. This is consistent with earlier work (Moore and S k, 2009b, 2011; Jackson and Moore, 2014). There are two possible explanations of this finding. First, excitation-pattern cues may have been usable at very high frequencies, because the relative bandwidth of the auditory filter is smaller at these frequencies than at

medium frequencies (Oxenham and Shera, 2003). However, very sharp tuning of the auditory filter at high frequencies has been observed only for low-level stimuli (Oxenham and Simonson, 2006), while our listeners were able to perform the TFS1 task even for the level of 60 dB/ERB_N. The other possibility is that some TFS cues remain at high frequencies (Heinz *et al.*, 2001; Recio-Spinoso *et al.*, 2005) and that these cues are sufficient to perform the TFS1 task. Consistent with this, Moore and Ernst (2012) presented evidence that the frequency discrimination of pure tones depends on a temporal mechanism at low and medium frequencies and a place mechanism at very high frequencies, but the transition between the two occurs at about 8000 Hz rather than at 4000–5000 Hz.

VI Summary and Conclusions

The results of experiments 1 and 2 showed that performance of the TFS1 task worsened at high levels, consistent with an effect of frequency selectivity. However, the worsening was marked only for the highest TEN level used (60 dB/ERB_N). For TEN levels up to 50 dB/ERB_N, the effect of level was small and somewhat variable across listeners, consistent with the results of Moore and S k (2009a, 2011).

The results of experiment 3 suggested that the changes in excitation level produced by a frequency shift corresponding to the measured threshold in the TFS1 task were very small. The estimated upper limits of the changes in excitation level at threshold at the center frequency corresponding to the eighth harmonic were 0.25, 0.12, 0.15, and 0.27 dB for the levels of 60, 50, 40, and 20 dB/ERB_N. These changes seem too small to provide usable cues even based on the excitation-pattern model proposed by Micheyl *et al.* (2010). Therefore, if performance of the TFS1 task is influenced by frequency selectivity, the effect may be an indirect one; poorer frequency selectivity may make the TFS at the output of the auditory filters more complex and more rapidly time varying, making it harder to decode.

Acknowledgments

This work was supported by MRC Grant No. G0900591. We thank Dr. Peter Watson for assistance with statistical analysis.

References

- Ardoint M, Lorenzi C. Effects of low-pass and high-pass filtering on the intelligibility of speech based on temporal fine structure or envelope cues. *Hear Res.* 2010; 260:89–95. [PubMed: 19963053]
- Ardoint M, Sheft S, Fleuriot P, Garnier S, Lorenzi C. Perception of temporal fine-structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss. *Int J Audiol.* 2010; 49:823–831. [PubMed: 20666687]
- Baker RJ, Rosen S. Auditory filter nonlinearity across frequency using simultaneous notched-noise masking. *J Acoust Soc Am.* 2006; 119:454–462. [PubMed: 16454300]
- Bernstein JG, Oxenham AJ. The relationship between frequency selectivity and pitch discrimination: Effects of stimulus level. *J Acoust Soc Am.* 2006; 120:3916–3928. [PubMed: 17225419]
- Bernstein LR, Green DM. Detection of simple and complex changes of spectral shape. *J Acoust Soc Am.* 1987; 82:1587–1592. [PubMed: 3693697]
- Buus, S.; Florentine, M. Sensitivity to excitation-level differences within a fixed number of channels as a function of level and frequency. *Advances in Hearing Research.* Manley, GA.; Klump, GM.; Köppl, C.; Fastl, H.; Oekinghaus, H., editors. World Scientific; Singapore: 1995. p. 401-412.

- Carlyon RP, Long CJ, Micheyl C. Across-channel timing differences as a potential code for the frequency of pure tones. *J Assoc Res Otolaryngol*. 2011; 13:159–171. [PubMed: 22160791]
- 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*. 2002; 88:334–346.
- de Boer, E. On the ‘residue’ in hearing. Ph.D. thesis; University of Amsterdam: 1956.
- Ernst SM, Moore BCJ. Mechanisms underlying the detection of frequency modulation. *J Acoust Soc Am*. 2010; 128:3642–3648. [PubMed: 21218896]
- Ernst SM, Moore BCJ. The role of time and place cues in the detection of frequency modulation by hearing-impaired listeners. *J Acoust Soc Am*. 2012; 131:4722–4731. [PubMed: 22712945]
- Florentine M, Buus S. An excitation-pattern model for intensity discrimination. *J Acoust Soc Am*. 1981; 70:1646–1654.
- Füllgrabe C, Moore BCJ, Stone MA. Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Front Aging Neurosci*. 2015; 6:1–25.
- Glasberg BR, Moore BCJ. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J Acoust Soc Am*. 1986; 79:1020–1033. [PubMed: 3700857]
- Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear Res*. 1990; 47:103–138. [PubMed: 2228789]
- Glasberg BR, Moore BCJ. Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise. *J Acoust Soc Am*. 2000; 108:2318–2328. [PubMed: 11108372]
- Gockel H, Moore BCJ, Patterson RD. Influence of component phase on the loudness of complex tones. *Acta Acust Acust*. 2002; 88:369–377.
- Green DM. The number of components in profile analysis tasks. *J Acoust Soc Am*. 1992; 91:1616–1623. [PubMed: 1564198]
- Heinz MG, Colburn HS, Carney LH. Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. *Neur Comput*. 2001; 13:2273–2316.
- Heinz, MG.; Swaminathan, J.; Boley, JD.; Kale, S. Acrossfiber coding of temporal fine-structure: Effects of noise-induced hearing loss on auditory-nerve responses. *The Neurophysiological Bases of Auditory Perception*. Lopez-Poveda, EA.; Palmer, AR.; Meddis, R., editors. Springer; New York: 2010. p. 621-630.
- Hopkins K, Moore BCJ. Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J Acoust Soc Am*. 2007; 122:1055–1068. [PubMed: 17672653]
- Hopkins K, Moore BCJ. The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *J Acoust Soc Am*. 2009; 125:442–446. [PubMed: 19173429]
- Hopkins K, Moore BCJ. The importance of temporal fine structure information in speech at different spectral regions for normalhearing and hearing-impaired subjects. *J Acoust Soc Am*. 2010; 127:1595–1608. [PubMed: 20329859]
- Hopkins K, Moore BCJ. The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J Acoust Soc Am*. 2011; 130:334–349. [PubMed: 21786903]
- Hopkins K, Moore BCJ, and Stone MA. Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J Acoust Soc Am*. 2008; 123:1140–1153. [PubMed: 18247914]
- Houtgast, T. Lateral suppression in hearing. Ph.D. thesis; Free University of Amsterdam: 1974.
- Jackson HM, Moore BCJ. Contribution of temporal fine structure information and fundamental frequency separation to intelligibility in a competing-speaker paradigm. *J Acoust Soc Am*. 2013; 133:2421–2430. [PubMed: 23556607]
- Jackson HM, Moore BCJ. The role of excitation-pattern and temporal-fine-structure cues in the discrimination of harmonic and frequency-shifted complex tones. *J Acoust Soc Am*. 2014; 135:1356–1570. [PubMed: 24606274]
- Johnson DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am*. 1980; 68:1115–1122. [PubMed: 7419827]

- Joris PX, Louage DH, Cardoen L, van der Heijden M. Correlation index: A new metric to quantify temporal coding. *Hear Res.* 2006; 216–217:19–30.
- 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.* 2009; 125:27–30. [PubMed: 19173391]
- 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 USA.* 2006; 103:18866–18869. [PubMed: 17116863]
- Masson ME. Using confidence intervals for graphically based data interpretation. *Can J Exp Psychol.* 2003; 57:203–220. [PubMed: 14596478]
- Micheyl C, Dai H, Oxenham AJ. On the possible influence of spectral- and temporal-envelope cues in tests of sensitivity to temporal fine structure. *J Acoust Soc Am.* 2010; 127:1809–1810.
- Miller GA. Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *J Acoust Soc Am.* 1947; 19:609–619.
- Moore BCJ. The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol.* 2008; 9:399–406. [PubMed: 18855069]
- Moore, BCJ. Auditory Processing of Temporal Fine Structure: Effects of Age and Hearing Loss. World Scientific; Singapore: 2014. p. 1-182.
- Moore BCJ, Ernst SM. Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code. *J Acoust Soc Am.* 2012; 132:1542–1547. [PubMed: 22978883]
- Moore BCJ, Glasberg BR. Forward masking patterns for harmonic complex tones. *J Acoust Soc Am.* 1983a; 73:1682–1685. [PubMed: 6863746]
- Moore BCJ, Glasberg BR. Masking patterns of synthetic vowels in simultaneous and forward masking. *J Acoust Soc Am.* 1983b; 73:906–917. [PubMed: 6221041]
- Moore BCJ, Glasberg BR. Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns. *Hear Res.* 1987; 28:209–225. [PubMed: 3654390]
- Moore BCJ, Glasberg BR. Modeling binaural loudness. *J Acoust Soc Am.* 2007; 121:1604–1612. [PubMed: 17407897]
- 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.* 2006; 120:934–944. [PubMed: 16938981]
- Moore BCJ, Gockel H. Resolvability of components in complex tones and implications for theories of pitch perception. *Hear Res.* 2011; 276:88–97. [PubMed: 21236327]
- Moore BCJ, Hopkins K, Cuthbertson SJ. Discrimination of complex tones with unresolved components using temporal fine structure information. *J Acoust Soc Am.* 2009; 125:3214–3222. [PubMed: 19425664]
- Moore BCJ, Huss M, Vickers DA, Glasberg BR, Alcántara J. A test for the diagnosis of dead regions in the cochlea. *Br J Audiol.* 2000; 34:205–224. [PubMed: 10997450]
- Moore BCJ, Oldfield SR, Dooley G. Detection and discrimination of spectral peaks and notches at 1 and 8 kHz. *J Acoust Soc Am.* 1989; 85:820–836. [PubMed: 2925997]
- Moore BCJ, S k A. Effects of carrier frequency and background noise on the detection of mixed modulation. *J Acoust Soc Am.* 1994; 96:741–751. [PubMed: 7930075]
- Moore BCJ, S k 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.* 1995; 97:2468–2478. [PubMed: 7714263]
- Moore BCJ, S k A. Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *J Acoust Soc Am.* 1996; 100:2320–2331. [PubMed: 8865639]
- Moore BCJ, S k A. Development of a fast method for determining sensitivity to temporal fine structure. *Int J Audiol.* 2009a; 48:161–171. [PubMed: 19085395]
- Moore BCJ, S k A. Sensitivity of the human auditory system to temporal fine structure at high frequencies. *J Acoust Soc Am.* 2009b; 125:3186–3193. [PubMed: 19425661]

- Moore BCJ, S k A. Effect of level on the discrimination of harmonic and frequency-shifted complex tones at high frequencies. *J Acoust Soc Am*. 2011; 129:3206–3212. [PubMed: 21568422]
- Oxenham AJ, Shera CA. Estimates of human cochlear tuning at low levels using forward and simultaneous masking. *J Assoc Res Otolaryngol*. 2003; 4:541–554. [PubMed: 14716510]
- Oxenham AJ, Simonson AM. Level dependence of auditory filters in nonsimultaneous masking as a function of frequency. *J Acoust Soc Am*. 2006; 119:444–453. [PubMed: 16454299]
- Palmer AR, Russell IJ. Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear Res*. 1986; 24:1–15. [PubMed: 3759671]
- Plack CJ, Oxenham AJ. Basilar-membrane nonlinearity and the growth of forward masking. *J Acoust Soc Am*. 1998; 103:1598–1608. [PubMed: 9514024]
- Plomp R. The ear as a frequency analyzer. *J Acoust Soc Am*. 1964; 36:1628–1636.
- Recio-Spinoso A, Temchin AN, van Dijk P, Fan YH, Ruggero MA. Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers. *J Neurophysiol*. 2005; 93:3615–3634. [PubMed: 15659532]
- Robles L, Ruggero MA. Mechanics of the mammalian cochlea. *Physiol Rev*. 2001; 81:1305–1352. [PubMed: 11427697]
- Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L. Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J Acoust Soc Am*. 1997; 101:2151–2163. [PubMed: 9104018]
- Schouten JF, Ritsma RJ, Cardozo BL. Pitch of the residue. *J Acoust Soc Am*. 1962; 34:1418–1424.
- S k A, Moore BCJ. Frequency discrimination as a function of frequency, measured in several ways. *J Acoust Soc Am*. 1995; 97:2479–2486. [PubMed: 7714264]
- S k A, Moore BCJ. Implementation of a fast method for measuring psychophysical tuning curves. *Int J Audiol*. 2011; 50:237–242. [PubMed: 21299376]
- S k A, Moore BCJ. Implementation of two tests for measuring sensitivity to temporal fine structure. *Int J Audiol*. 2012; 51:58–63. [PubMed: 22050366]
- Strelcyk O, Dau T. Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J Acoust Soc Am*. 2009; 125:3328–3345. [PubMed: 19425674]
- Swaminathan J, Heinz MG. Psychophysiological analyses demonstrate the importance of neural envelope coding for speech perception in noise. *J Neurosci*. 2012; 32:1747–1756. [PubMed: 22302814]
- Unoki M, Irino T, Glasberg BR, Moore BCJ, Patterson RD. Comparison of the roex and gamma chirp filters as representations of the auditory filter. *J Acoust Soc Am*. 2006; 120:1474–1492. [PubMed: 17004470]
- Verschooten, E. Assessment of fundamental cochlear limits of frequency resolution and phase-locking in humans and animal models. Ph.D. thesis, KU Leuven; 2013.
- Verschooten E, Joris PX. Estimation of neural phase locking from stimulus-evoked potentials. *J Assoc Res Otolaryngol*. 2014; 15:767–787. [PubMed: 24890715]
- Viemeister NF, Bacon SP. Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tones. *J Acoust Soc Am*. 1988; 84:172–178. [PubMed: 3411045]
- Wier CC, Jesteadt W, Green DM. Frequency discrimination as a function of frequency and sensation level. *J Acoust Soc Am*. 1977; 61:178–184. [PubMed: 833369]

Fig. 1.

Power spectra of an H tone and an I tone (frequency shift = 50% F_0) presented in TEN with a level of 20 dB/ERB_N. The analysis bandwidth was 1 Hz. Complex tones had an F_0 of 250 Hz. The frequency range in the figure is 5000 Hz, but the TEN extended to 16 000 Hz.

Fig. 2.

Results of experiment 1. Thresholds were measured for complex tones with F_0 s of 60, 100, 250, and 625 Hz, presented in a TEN with a level of either 20 dB/ERBN or 60 dB/ERBN. The data for S3 were not included in the average. Error bars show ± 1 standard error of the mean (sem) calculated from the log-transformed thresholds for the last three testing sessions (for individual thresholds) or from the log-transformed individual thresholds (for the group-averaged thresholds).

Fig. 3.

Results of experiment 2. Individual and group-averaged shift thresholds are plotted as a function of TEN level. The overall level of the complex tone is shown on the top of the bottom-right panel. Data from experiment 1 are plotted as triangles for comparison (slightly shifted to the left for better readability). Error bars show ± 1 sem, calculated as for Fig. 2. The bottom-right panel also shows the mean data of Moore and Sek (2009a) with level expressed as the overall level of the complex tone.

Fig. 4.

Results of experiment 3. Individual and group-averaged normalized signal levels at masked threshold are plotted as a function of frequency shift (0, 25, 50, and 75% of F_0), with TEN level as parameter (20, 40, 50, and 60 dB/ERBN). Thresholds for different TEN levels are shifted along the x axis for better readability. Error bars show ± 1 sem, calculated as for Fig. 3.