



OPEN

## Exploiting individual differences to assess the role of place and phase locking cues in auditory frequency discrimination at 2 kHz

Vinay<sup>1✉</sup> & Brian C. J. Moore<sup>2</sup>

The relative role of place and temporal mechanisms in auditory frequency discrimination was assessed for a centre frequency of 2 kHz. Four measures of frequency discrimination were obtained for 63 normal-hearing participants: detection of frequency modulation using modulation rates of 2 Hz (FM2) and 20 Hz (FM20); detection of a change in frequency across successive pure tones (difference limen for frequency, DLF); and detection of changes in the temporal fine structure of bandpass filtered complex tones centred at 2 kHz (TFS). Previous work has suggested that: FM2 depends on the use of both temporal and place cues; FM20 depends primarily on the use of place cues because the temporal mechanism cannot track rapid changes in frequency; DLF depends primarily on temporal cues; TFS depends exclusively on temporal cues. This led to the following predicted patterns of the correlations of scores across participants: DLF and TFS should be highly correlated; FM2 should be correlated with DLF and TFS; FM20 should not be correlated with DLF or TFS. The results were broadly consistent with these predictions and with the idea that frequency discrimination at 2 kHz depends partly or primarily on temporal cues except for frequency modulation detection at a high rate.

The auditory detection of changes in frequency may depend on two mechanisms. The first, called the “place” mechanism, is based on the frequency analysis that occurs in the cochlea<sup>1,2</sup>. The place of maximum response along the cochlear partition varies with input frequency and changes in frequency may be coded by changes in the place of maximum response<sup>1</sup> or by changes on the low-frequency (apical) side of the response pattern, also called the excitation pattern<sup>3,4</sup>. In principle, the place mechanism can operate over the whole audible frequency range. The second mechanism, called the “temporal” mechanism, depends on the fact that, at least for low-frequency inputs, the action potentials (spikes) evoked in auditory neurons are synchronised to a specific phase of the input waveform, a property called phase locking<sup>5</sup>. As a result, the time intervals between successive spikes are approximately integer multiples of the period of the input. Changes in frequency may be detected via changes in the pattern of inter-spike intervals<sup>6–8</sup>. It is known that phase locking in the auditory nerve of animals becomes less precise at high frequencies<sup>5,9,10</sup>. Such a loss of precision probably also occurs for humans, but the upper limit of phase locking in humans is still a subject of considerable debate. Some researchers have argued that the temporal mechanism plays a role in frequency discrimination for frequencies up to 8–10 kHz<sup>7,8,11,12</sup>, while others have argued for a much lower limit, perhaps 1.5 kHz<sup>13,14</sup>.

Frequency discrimination can be measured in several ways. One method is to measure the modulation depth required to distinguish a frequency-modulated carrier from an unmodulated carrier<sup>3,15–17</sup>. The threshold determined in this way is called the frequency modulation detection limen (FMDL). It has been proposed that the detection of frequency modulation (FM) depends partly on the use of temporal (phase-locking) information when the carrier frequency is below about 5 kHz and the FM rate is low (below about 5 Hz), while for FM rates above about 10 Hz only a place mechanism is used<sup>18–23</sup>. This proposal is based on the idea that the mechanism that “decodes” the temporal information is “sluggish” and cannot track rapid changes in frequency<sup>16,19</sup>.

The relative importance of place and temporal information for low FM rates may depend on the overall level of the stimuli. Ernst and Moore<sup>24</sup> measured FMDLs for carrier frequencies ( $f_c$ ) of 1000, 4000, and 6000 Hz, using modulation frequencies ( $f_m$ ) of 2 and 10 Hz and levels of 20 and 60 dB SL, both with and without random

<sup>1</sup>Audiology Group, Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Tungasletta 2, 7491 Trondheim, Norway. <sup>2</sup>Cambridge Hearing Group, Department of Psychology, University of Cambridge, Cambridge, UK. ✉email: vinay.nagaraj@ntnu.no

amplitude modulation (AM), applied in all intervals of a forced-choice trial. The AM was intended to disrupt place cues for detection of the FM. At 60 dB SL, the deleterious effect of the AM was smaller for  $f_m = 2$  than for  $f_m = 10$  Hz for  $f_c = 1000$  and 4000 Hz, but not for  $f_c = 6000$  Hz. This is consistent with the idea that, for  $f_c$  below about 5000 Hz and  $f_m = 2$  Hz, frequency modulation can be detected via changes in phase locking over time and this information is less affected than place information by the addition of random AM. However, at 20 dB SL, the deleterious effect of the added AM for  $f_c = 1000$  and 4000 Hz was similar for the two values of  $f_m$ , while for  $f_c = 6000$  Hz, the deleterious effect of the AM was greater for  $f_m = 10$  than for  $f_m = 2$  Hz. Ernst and Moore<sup>24</sup> suggested that, at low SLs, the auditory filters become relatively sharp, so that excitation-pattern cues influence FMDLs to a greater extent, even for low  $f_c$  and low  $f_m$ . However, the hypothesis that FM detection depends partly on temporal information for low FM rates and mainly on place information for high rates is not universally accepted<sup>25–27</sup>.

Another measure of frequency discrimination is the smallest detectable difference in frequency between successive steady tones<sup>28</sup>. This measure is called the difference limen for frequency (DLF). In one version of this measure, two successive tones are presented with slightly different frequencies and the participant is asked to indicate whether the first or the second is higher in pitch. However, untrained participants often find it hard to name the direction of a pitch change, and extensive practice may be required to achieve stable performance<sup>29–31</sup>. A task that is easier and requires less practice involves the use of four successive tones in each of two observation intervals. In one randomly selected interval all four tones have the same frequency. In the other interval, the tones alternate in frequency between two values, A and B, giving the pattern ABAB. The task is to identify the interval in which the tones changed in pitch<sup>12,32</sup>. Moore and Ernst<sup>12</sup> used this task over a wide range of centre frequencies. They found that DLFs, expressed as a proportion of centre frequency, worsened with increasing frequency from 2 to 8 kHz and then became roughly constant with further increases in frequency. They suggested that the worsening in DLFs from 2 to 8 kHz reflected a progressive reduction in the precision of temporal information (phase locking) and that the “break point” around 8 kHz reflected a transition from a temporal code to a place code. However, this interpretation is controversial<sup>14</sup>.

Finally, frequency discrimination can be measured using a task that is specifically designed to limit the use of place cues, promoting performance based on phase locking. The task, called the TFS1 task, involves the detection of changes in the temporal fine structure (TFS) of bandpass filtered complex tones<sup>33–35</sup>. The participant is required to discriminate a harmonic complex tone (H), with fundamental frequency  $F_0$ , from a similar tone in which all components are shifted up in frequency by  $\Delta F$  (where  $\Delta F \leq 0.5F_0$ ), to create an inharmonic tone (I). People with normal hearing perceive the H and I tones as having different pitches if  $\Delta F$  is sufficiently large<sup>36–38</sup>. The H and I tones have the same envelope repetition rate (equal to the  $F_0$  of the H tones), but they differ in their TFS. The phases of the components are chosen randomly for each H and I tone, with the result that the envelope shape fluctuates randomly from one tone to the next, so that the envelope shape does not provide a cue for discriminating the H and I tones. The H and I tones are made up of many components and are passed through a fixed bandpass filter centred on the higher components, to make place cues minimal. The filter has a central flat region and skirts that decrease in level at a rate of 30 dB/octave. The use of a relatively shallow slope ensures minimal place cues as components move in and out of the passband<sup>38</sup>. Hence, it is thought that performance of the TFS1 task depends primarily on the use of temporal cues rather than place cues<sup>39,40</sup>. The task is similar to that used by Ernst and Moore<sup>12</sup> to measure DLFs. One randomly selected interval contains the sequence HHHH and the other contains the sequence HIHI, and the participant is asked to identify the interval in which the tones changed in pitch.

The present study was designed to test the following hypotheses, all for a centre frequency of 2 kHz, for which there is debate about the relative importance of place and temporal cues<sup>14</sup>:

- (1) FMDLs depend partly on a temporal mechanism for an FM rate of 2 Hz, while FMDLs depend primarily on a place mechanism for an FM rate of 20 Hz.
- (2) DLFs depend primarily on a temporal mechanism.
- (3) Thresholds measured using the TFS1 task, denoted TFS, depend primarily on a temporal mechanism.

To test these hypotheses, four measures of frequency discrimination were obtained for 63 normal-hearing participants: detection of FM using a 2-Hz modulation rate (FM2); detection of FM using a 20-Hz rate (FM20); DLF; and TFS. There are natural variations across normal-hearing participants in sharpness of tuning<sup>41,42</sup>, which partly determines the efficacy of place cues, and according to the hypotheses these variations should affect FM2 and FM20 values but have little effect on DLF and TFS values. There are also individual variations in the precision of phase locking<sup>43,44</sup>, and according to the hypotheses these should affect the values of DLF and TFS, and to a lesser extent FM2, but should have little effect on FM20 values. The three hypotheses lead to the following predictions based on the variations in performance across participants:

P1. FM2 and FM20 values should be moderately correlated, because place information may be partly used for both FM rates and because the task is very similar for the two measures.

P2. FM2 values should be moderately correlated with DLF and TFS values, because FM2 values depend partly on the use of temporal cues and DLF and TFS values depend strongly on the use of temporal cues.

P3. FM20 values should not be correlated with DLF or TFS values, because FM20 values depend mainly on place cues while DLF and TFS values depend strongly on temporal cues.

P4. DLF and TFS values should be highly correlated because both depend strongly on the use of temporal cues and because the task is very similar for the two measures.

Predictions P1 to P4 imply that some correlations will be bigger than others. The predictions about the relative sizes of the correlations are presented in the Results section, together with the outcomes.

## Methods

**Participants.** All participants were students or researchers at the Norwegian University of Science and Technology in Trondheim, Norway. The participants were recruited via flyers and via the university media channel. All participants had audiometric thresholds better than 20 dB HL (measured using an Otometrics Aurical Otosuite and Telephonics TDH-39P headphones) for all octave-spaced frequencies from 250 to 8000 Hz. Their ages ranged from 20 to 33 years. None of the participants had non-auditory neural conditions and none had any history of ear discharge or pain in their ears. None of the participants reported having neurological problems. Middle-ear function was checked using a Grason-Stadler Tympanometer Pro tympanometer; all participants had normal middle-ear function. Participants were paid for participating by being given gift vouchers. Data were obtained for 63 participants. This number is sufficient such that any correlation between scores for two tests greater than 0.21 would be significant at  $p < 0.05$ , using a directional (one-tailed) test.

The study and methods followed the tenets of the Declaration of Helsinki, and informed consent was obtained from participants after the nature and possible consequences of participation were explained. Approval for the experiments was given by the Regional Committee for Medical and Health Research Ethics (REK) in Norway.

**Apparatus and test procedures.** Stimuli were generated using the “PSYCHOACOUSTICS” software<sup>45</sup> using the built-in 24 bit soundcard of a Microsoft Surface Pro laptop computer. Stimuli were delivered via Sennheiser HDA200 headphones, which have a smooth frequency response in the frequency region around 2 kHz. Testing was carried out separately for each ear of each participant in a sound-attenuating room.

All thresholds were measured with a two-alternative forced-choice procedure using a two-down one-up adaptive procedure to estimate the 71% correct point on the psychometric function. The two observation intervals were marked on the laptop screen by successively lighting up two spatially separated boxes. After the participant had responded, feedback was provided by flashing the correct box, with green for a correct response and red for an incorrect response.

The following measures were obtained.

- (1) Absolute thresholds for 2-kHz pure-tone signals. The signal duration was 1000 ms including 10-ms raised-cosine ramps. The two observation intervals were separated by 500 ms. The step size was 4 dB until four turnpoints had occurred and 2 dB thereafter. A run continued until eight turnpoints (changes from decreasing to increasing level and vice versa) had been obtained using the 2-dB step size. The threshold was taken as the arithmetic mean of the levels at the last six turnpoints. The threshold was measured twice for each ear, and the two estimates were averaged. For the measures of frequency discrimination, the level of the test stimuli (excluding any background noise) was set to be 30 dB SL, i.e., 30 dB above the absolute threshold.
- (2) FMDLs using modulation frequencies of 2 and 20 Hz and a carrier frequency of 2 kHz. The signal duration was 1000 ms including 20-ms raised-cosine ramps and the two intervals in a trial were separated by 300 ms. The carrier was unmodulated in one randomly chosen interval and sinusoidally frequency modulated in the other. The starting frequency deviation was 80 Hz, which was chosen to be well above the likely threshold value. The frequency deviation was changed by a factor of 1.95 until two turnpoints had occurred, then by a factor of 1.56 until two more turnpoints had occurred and then by a factor of 1.25. A run continued until eight turnpoints had been obtained using the smallest step size. The threshold was taken as the geometric mean of the frequency deviations at the last eight turnpoints. Two estimates of threshold were obtained for each ear, and the final threshold was taken as the geometric mean of the two. Note that thresholds are expressed as the deviation from the carrier frequency; the overall frequency excursion of the modulated carrier was twice the specified deviation. For example, if the threshold was 10 Hz, the instantaneous frequency of the modulated carrier varied from 1990 to 2010 Hz.
- (3) DLFs were measured using the task outlined in the introduction. In one interval of a trial (selected randomly), there were four successive 200-ms bursts (including 20-ms raised-cosine ramps) of tone A, with a fixed frequency,  $f$ . The bursts were separated by 100 ms. In the other interval, tones A and B alternated, with the same 100-ms inter-burst interval, giving the pattern ABAB. The two intervals were separated by 300 ms. 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. The test tones were presented in a background of threshold equalizing noise (TEN)<sup>46</sup>, which was intended to reduce variations in loudness with changes in frequency; the loudness of the test tones was determined mainly by the signal-to-TEN ratio<sup>47</sup>. The TEN started 300 ms before the first tone burst and ended 300 ms after the last tone burst. The TEN level is specified as the level in a 1-ERB<sub>N</sub> wide band centred at 1000 Hz, where ERB<sub>N</sub> stands for the average value of equivalent rectangular bandwidth of the auditory filter at moderate sound levels for listeners with normal hearing<sup>48</sup>. The level of the TEN was set 15 dB below the level of the test tones. The starting value of  $\Delta f$  was 80 Hz, which was chosen to be well above the likely threshold value. The value of  $\Delta f$  was changed by a factor of 1.95 until one turnpoint had occurred, then by a factor of 1.56 until one more turnpoint had occurred and then by a factor of 1.25. A run continued until eight turnpoints had been obtained. The threshold was taken as the geometric mean of the  $\Delta f$  values at the last six turnpoints. Two estimates of threshold were obtained for each ear, and the final threshold was taken as the geometric mean of the two.
- (4) The stimuli for the TFS1 task were bandpass filtered complex tones. The bandpass filter had a flat top and slopes of 30 dB/oct. The passband width is specified as the width of the flat portion (not the 3-dB down points) and was equal to 222.22 Hz. The tones were either harmonic (H) with a fundamental frequency  $F_0$  of 222.22 Hz, or all components were shifted up in frequency by  $\Delta F$ , giving an inharmonic tone I. The lower edge of the passband was set to 2000 Hz, so the 9th harmonic was the lowest component of the H tone that

fell within the passband ( $9 \times 222.22 = 2000$ ). To prevent the detection of combination tones, and to limit the audibility of components falling on the skirts of the passband, the tones were presented in a background of TEN with the same characteristics as for the measurement of DLFs. The level of the TEN was set 15 dB below the overall level of the complex tone, which corresponds to about 10.5 dB below the level of each component within the passband. It was hypothesized that the audible components in the stimulus would not be resolved in the auditory system<sup>49</sup> and that the task would be performed primarily using temporal cues<sup>39,40</sup>. In one interval of a trial (selected randomly), there were four successive 200-ms bursts (including 20-ms raised-cosine ramps) of tone H. The bursts were separated by 100 ms. In the other interval, tones H and I alternated, with the same 100-ms inter-burst interval, giving the pattern HIHI. The two intervals were separated by 300 ms. The task of the subject was to choose the interval in which the sound changed across the four tone bursts within an interval. The phases of the components were chosen randomly for every tone burst, so the envelope of each tone burst was different and performance of the task could not be based on envelope cues. The starting value of  $\Delta F$  was 111.11 Hz. This value of  $\Delta F$  leads to the greatest possible difference between the H and I tones. The value of  $\Delta F$  was changed by a factor of 1.95 until one turnpoint had occurred, then by a factor of 1.56 until one more turnpoint had occurred and then by a factor of 1.25. If the adaptive procedure called for a value of  $\Delta F$  greater than 111.11 Hz, the value was kept at 111.11 Hz until two successive correct responses occurred. In practice the limit was reached very rarely. A run continued until eight turnpoints had been obtained. The threshold was taken as the geometric mean of the values of  $\Delta F$  at the last six turnpoints. Two estimates of threshold were obtained for each ear, and the final threshold was taken as the geometric mean of the two.

For each of the frequency discrimination tasks, the nature of the task was carefully described to each participant, including what to “listen for”. For the FM detection tasks they were asked to pick the interval in which the sounds appeared to “wobble”. For the DLF and TFS tasks, they were told to pick the interval in which the tones changed in pitch. When the initial performance of a participant was erratic or when a participant reported that they were not sure what to listen for, the participant was given practice and further instruction until their performance became stable. The FMDLs were measured first, with FM2 and FM20 tested in random order across participants. Then the TFS and DLF tasks were performed. The ear that was tested first was varied randomly across participants.

## Results

The thresholds for each measure were generally similar for the two ears of a given participant. To reduce the effects of random errors of measurement, the thresholds were averaged across ears (arithmetic means for absolute thresholds, geometric means for the measures of frequency discrimination). Two of the measures of frequency discrimination, FM2 and DLF, were weakly but significantly correlated with the absolute thresholds (FM2,  $r = 0.237$ ,  $p = 0.031$ ; DLF,  $r = 0.211$ ,  $p = 0.048$ , based on a 1-tailed test, since it was hypothesized that performance would worsen with increasing absolute threshold). However, these correlations are not significant if allowance is made for the calculation of four correlations (based on Bonferroni correction).

In what follows, the patterns of correlation between the different measures of frequency discrimination were examined to assess the extent to which they were consistent with predictions P1–P4. Because there were specific predictions about the direction and magnitude of each correlation, no correction was made for the number of correlations being calculated and the significance of the correlations was assessed using one-tailed tests.

Prediction P1 was that FM2 and FM20 values would be moderately correlated. Figure 1 is a scatter plot of FM20 values versus FM2 values. The correlation between the two was  $r = 0.619$  ( $p < 0.0001$ , 90% confidence interval, CI 0.471, 0.773). Hence P1 was confirmed. The results are consistent with the idea that FM2 and FM20 were partly based on a common mechanism, using place cues.

Prediction P2 was that FM2 values would be moderately correlated with DLF and TFS values. Figure 2 is a scatter plot of DLF values versus FM2 values. The correlation between the two was  $r = 0.248$  ( $p = 0.025$ , CI 0.041, 0.434).

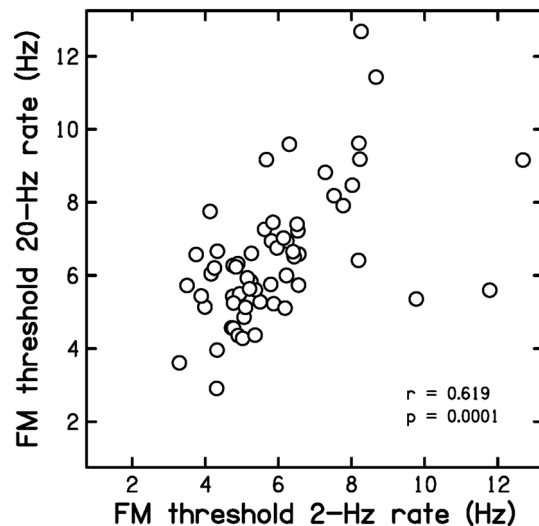
Figure 3 is a scatter plot of TFS values versus FM2 values. The correlation between the two was  $r = 0.234$  ( $p = 0.032$ , CI 0.026, 0.423). While the correlations with FM2 values were significant for both the DLF and TFS values, the correlations were small. This weakly supports P2.

Prediction P3 was that FM20 values would not be correlated with DLF or TFS values. Figure 4 is a scatter plot of DLF values versus FM20 values. The correlation between the two was  $r = 0.053$  ( $p = 0.340$ , ns, CI  $-0.158$ , 0.259). Figure 5 is a scatter plot of TFS values versus FM20 values. The correlation between the two was  $r = 0.057$  ( $p = 0.329$ , ns, CI  $-0.154$ , 0.263). These results are consistent with P3.

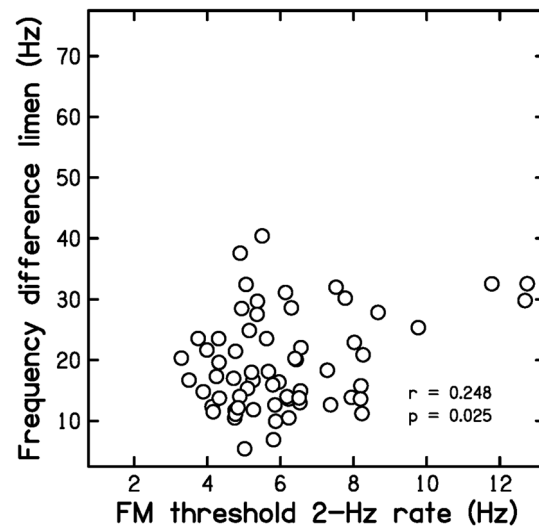
Prediction P4 was that DLF and TFS values would be highly correlated. Figure 6 is a scatter plot of TFS values versus DLF values. The correlation between the two was  $r = 0.796$  ( $p < 0.001$ , CI 0.704, 0.862). This high correlation is consistent with P4.

As noted in the “Introduction”, predictions P1 to P4 imply that some correlations will be bigger than others. Specifically, the following predictions can be made:

- (1)  $r(\text{TFS vs DLF})$  should be greater than  $r(\text{FM2 vs FM20})$ . This is the case because TFS and DLF values are hypothesized to depend primarily on the use of temporal cues, FM2 values to depend partly on temporal cues and partly on place cues, and FM20 values to depend primarily on place cues. The two correlations were 0.796 and 0.619, and based on Fisher’s  $r$  to  $z$  transform, these differ significantly ( $p = 0.023$ ).
- (2)  $r(\text{FM2 vs TFS})$  should be greater than  $r(\text{FM20 vs TFS})$ . This is the case because TFS values are hypothesized to depend primarily on the use of temporal cues, FM2 values to depend partly on temporal cues and partly



**Figure 1.** Scatter plot of FM20 values versus FM2 values.



**Figure 2.** Scatter plot of DLF values versus FM2 values.

on place cues, and FM20 values to depend primarily on place cues. The two correlations were 0.234 and 0.057, which was in the predicted direction, but the difference just failed to reach significance ( $p = 0.055$ ) (significance in this case and the case below was calculated taking into account the fact that the two correlations have a variable in common, in this case TFS<sup>50</sup>).

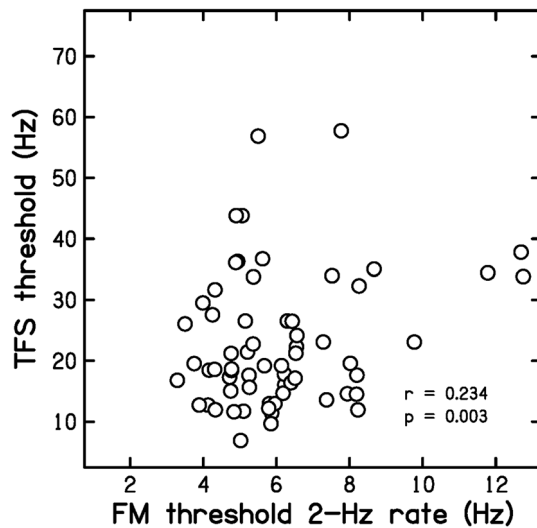
- (3)  $r(\text{FM2 vs DLF})$  should be greater than  $r(\text{FM20 vs DLF})$ . This is the case because DLF values are hypothesized to depend primarily on the use of temporal cues, FM2 values to depend partly on temporal cues and partly on place cues, and FM20 values to depend primarily on place cues. The two correlations were 0.248 and 0.053, and the difference was significant ( $p = 0.039$ ).

Overall, these outcomes are broadly consistent with the predictions.

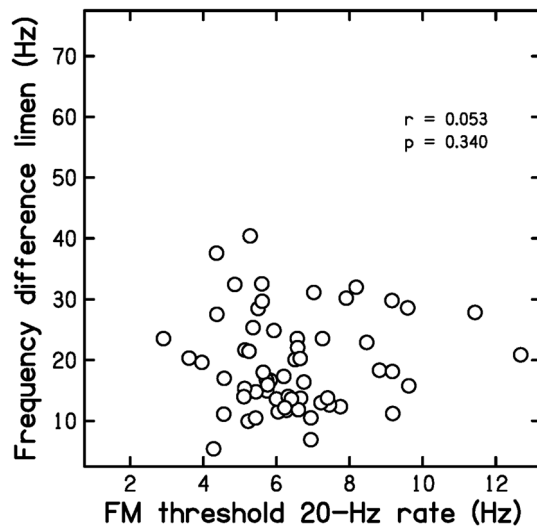
To gain more insight into the relationship between the scores for the four measures, a principal components analysis was conducted using the “ClustVis software<sup>51</sup>. The first two principal components (PC1 and PC2) accounted for 51.4 and 34.5% of the variance in the data, respectively, the remaining two components accounting for only 14% of the variance. Figure 7 shows where each participant fell in the PC1–PC2 space. There is no obvious clustering.

Table 1 shows the loading of each measure on PC1 and PC2. FM2 loaded moderately on PC1 and PC2. FM20 loaded strongly on PC2 and weakly on PC1. TFS and DLF loaded moderately strongly on PC1 and more weakly negatively on PC2. A reasonable interpretation is that PC1 primarily reflects sensitivity to temporal cues while





**Figure 3.** Scatter plot of TFS values versus FM2 values.

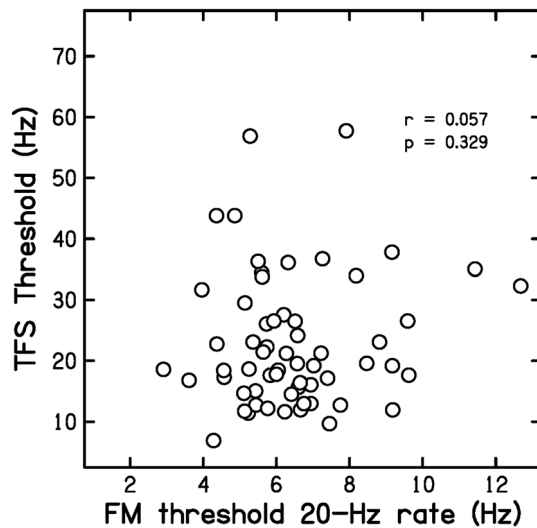


**Figure 4.** Scatter plot of DLF values versus FM20 values.

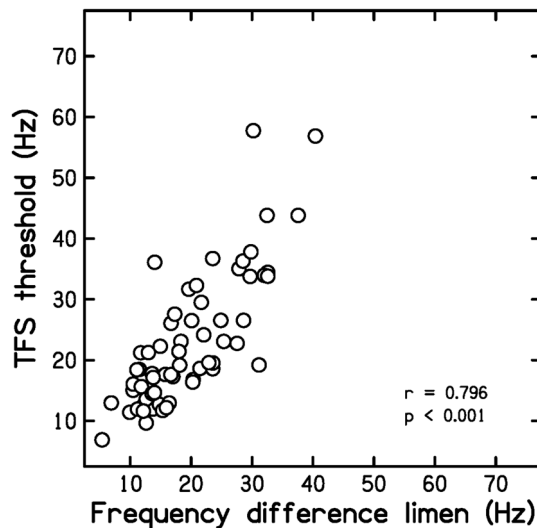
PC2 primarily reflects sensitivity to place cues. The analysis is consistent with the idea that TFS and DLF depend mainly on the use of temporal cues, FM20 depends mainly on the use of place cues, and FM2 depends on both temporal and place cues.

## Discussion

As noted in the introduction, the hypothesis that FM detection depends partly on the use of temporal information for low FM rates but exclusively on place information for high FM rates is the subject of some controversy<sup>25–27,52</sup>. For example, in a study also exploiting individual differences Whiteford and Oxenham<sup>25</sup> measured FMDLs for 1- and 20-Hz rates and a 500-Hz carrier presented at 70 dB SPL using 100 young normal-hearing participants. They also measured the detection of AM for the same rates, and assessed sensitivity to binaural phase-locking cues using a task involving the discrimination of FM that was in-phase at the two ears from FM that was out of phase at the two ears. This measure is denoted here dichotic FM. The correlation between the 1-Hz FMDLs and the dichotic FM thresholds, was not higher than the correlation between the 20-Hz FMDLs and dichotic FM thresholds, contrary to what would be expected if both the 1-Hz FMDLs and the dichotic FM thresholds depended on phase locking. However, the discrimination of interaural phase may be limited by specific binaural processes and might not provide a good measure of the ability to use monaural phase locking cues<sup>53</sup>. Whiteford and Oxenham<sup>25</sup> also estimated the sharpness of the upper and lower slopes of the excitation patterns evoked by the stimuli, using forward masking. For most participants, the lower slope was steeper, as expected<sup>54</sup>. The



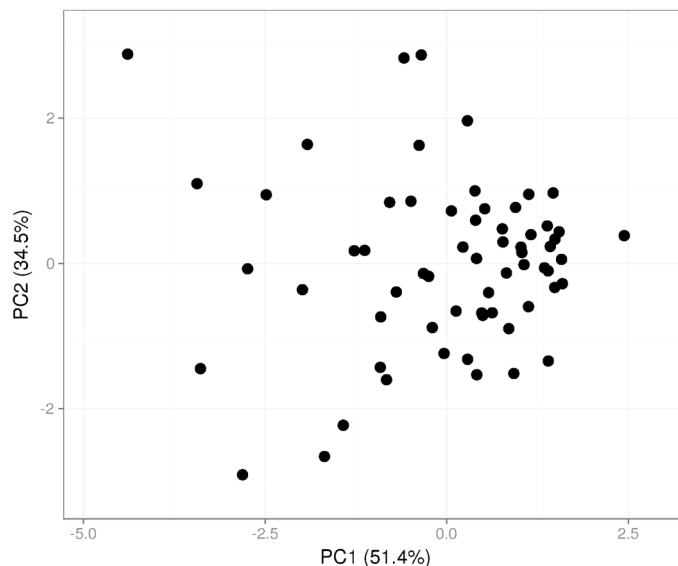
**Figure 5.** Scatter plot of TFS values versus FM20 values.



**Figure 6.** Scatter plot of TFS values versus DLF values.

correlation between the steeper of the lower and upper slopes and the FMDLs was moderate and similar for the two FM rates, contrary to what would be expected if FMDLs for the 20-Hz rate depended entirely on the use of place cues, while FMDLs for the 1-Hz rate depended partly on phase-locking cues. However, the slope measure showed very high individual variability, suggesting that it was imprecise. Also, as pointed out by the authors, the slope measure combined with the measure of AM sensitivity did not predict the 1-Hz or 20-Hz FMDLs better than just the measure of AM sensitivity, suggesting no clear relationship between frequency selectivity and either slow- or fast-rate FM detection thresholds. Thus the results of Whiteford and Oxenham do not allow strong conclusions to be drawn about the relative roles of place and temporal cues in FM detection.

The general pattern of the results presented in the present paper was broadly consistent with predictions P1 to P4. Prediction P1, that FM2 and FM20 values would be moderately correlated because both are partly determined by the use of place cues, was supported. Prediction P2 was that FM2 values would be moderately correlated with DLF and TFS values because FM2 values depend partly on the use of temporal cues and DLF and TFS values depend primarily on the use of temporal cues. This prediction was only weakly supported; the correlations were significant but small. Prediction P3, that FM20 values would not be correlated with DLF or TFS values because FM20 values depend mainly on the use of place cues while DLF and TFS values depend mainly on the use of temporal cues, was supported. Both correlations were below 0.06 and both were non-significant. Finally, prediction P4, that DLF and TFS values would be highly correlated because both depend primarily on the use of temporal cues, was strongly supported.



**Figure 7.** Results of a principal components analysis showing where each participant fell in the PC1–PC2 space.

	PC1	PC2
FM2	0.49	0.48
FM20	0.35	0.66
TFS	0.55	−0.43
DLF	0.57	−0.40

**Table 1.** Loading of each measure on PC1 and PC2.

P2, the weakly supported prediction, is based on the assumption that FM2 values depend partly on the use of temporal cues. The weak support for this prediction may be linked to the low SL of the stimuli in the present experiment, which was 30 dB. At low SLs, the auditory filters, which are thought to reflect the sharpness of the filtering in the cochlea<sup>55</sup>, are relatively sharp<sup>54,56</sup> and this greater sharpness promotes the use of place cues to detect FM, even for low FM rates<sup>24</sup>. If a higher level had been used, the auditory filters would have been less sharp, and there would probably have been more reliance on temporal cues for FM2<sup>24</sup>. It would be informative to repeat the current experiment using a higher SL. This would lead to a clearer test of the hypotheses and predictions presented in this paper.

The interpretation of the results in terms of the use of temporal and place cues is consistent with the results of the principal components analysis. This analysis showed that two components accounted for 86% of the variance in the data. PC1 can be interpreted as primarily related to the use of temporal cues, with moderately high loadings for TFS and DLF, a moderate loading for FM2 and a low loading for FM20. PC2 can be interpreted as primarily related to the use of place cues, with a high loading for FM20, a moderate loading for FM2, and negative loadings for TFS and DLF.

There is, however, an alternative explanation for the pattern of the results. It is possible that the correlations between scores for the different tasks were determined partly or mainly by the similarity of the tasks and/or stimuli. One could interpret PC1 as being related to the rate of frequency changes, higher loadings indicating slower changes. The number of high-low cycles per second was 1.67 Hz for DLF and TFS, 2 Hz for FM2 and 20 Hz for FM20, so the rate of change was monotonically related to the loading on PC1. One could interpret PC2 as being related to the use of gated tone bursts as opposed to continuous frequency changes; the loadings were negative for the two measures that used gated tone bursts (TFS and DLF) and positive for the two measures that used continuous frequency changes (FM2 and FM20). However, some aspects of the results are not entirely consistent with such an interpretation. Firstly, regarding PC1, the number of high-low cycles per second was only slightly lower for DLF and TFS (1.67 Hz) than for FM2 (2 Hz), but the loading on PC1 was higher for TFS (0.55) and DLF (0.57) than for FM2 (0.49). Secondly, regarding PC2, the loading was markedly higher for FM20 (0.66) than for FM2 (0.48), despite the fact that both involved continuous changes. Overall, while we cannot rule out an explanation for the pattern of results in terms of task/stimulus similarity, an interpretation in terms of temporal and place cues seems more reasonable.

The ability to detect FM based on place cues depends on two main factors: the sharpness of tuning in the cochlea, and the ability to detect fluctuations in excitation level at a given place in the cochlea. The sharpness of



the excitation pattern evoked by a frequency-modulated carrier does not depend on FM rate provided that the spectral components of the stimulus are not resolved, but the ability to detect AM for a stimulus of fixed duration is markedly better for a 20-Hz rate than for a 2-Hz rate<sup>24,57,58</sup>, probably because there are more modulation cycles in the stimulus at the higher rate. Hence if place cues were exclusively used to detect FM, the FM20 values should be markedly smaller than the FM2 values. In fact, the geometric mean values were 6.5 Hz for FM20 and 5.8 Hz for FM2. This supports the idea that FM2 values were partly determined by the use of temporal information.

The frequency discrimination thresholds showed considerable individual variability, as has been found for a variety of auditory abilities among participants with audiometric thresholds within the “normal” range<sup>59</sup>. The arithmetic means (with standard deviations, SDs, and ranges) of the frequency discrimination measures in the present study were: FM2 6.1 Hz (1.8, 3.3–12.8 Hz), FM20 6.8 Hz (2.3, 2.9–15.8 Hz), DLF 19.5 Hz (8.0, 5.4–37.6 Hz) and TFS 23.3 Hz (11.2, 6.9–57.8). The individual variability probably stems from both peripheral and central factors. However, it seems likely that the measures FM2 and FM20 were influenced at least partly by individual variability in the sharpness of tuning in the cochlea, which can be assessed by measurement of the sharpness of the auditory filters. For young participants with normal audiometric thresholds, the equivalent rectangular bandwidth of the auditory filters for a center frequency of 2 kHz has a mean of about 308 Hz and an SD of 32 Hz<sup>41</sup>. The SD is only about 10% of the mean. It is usually assumed that the steepness of the slope of the excitation pattern is highly correlated with the equivalent rectangular bandwidth of the auditory filter at the corresponding center frequency<sup>3</sup>, so one would expect the SD of the slope also to be about 10% of the mean value. However, the SDs for the FM2 and FM20 measures were 30% and 34% of the mean, respectively, suggesting sources of individual variability other than the sharpness of auditory filtering. One such source of variability is individual differences in the ability to detect fluctuations in excitation level at a given place within the cochlea<sup>58</sup>.

The individual variability for the measures DLF and TFS probably depends at least partly on individual differences in the precision of phase locking or in the ability of central mechanisms to make use of phase-locking information. It is noteworthy that the SDs of the DLF and TFS values were 41% and 48% of the mean, respectively. These percentages are larger than for the FM2 and FM20 values, suggesting that individual variability in the use of temporal cues is greater than individual variability in the use of place cues. The precision of phase locking can be indirectly assessed by recording the frequency-following response (FFR), a scalp-recorded electrophysiological response that reflects synchronous activity in subcortical neurons. Marmel et al.<sup>43</sup> measured the FFR synchronisation strength in response to a 606-Hz pure tone for participants with a wide range of ages and hearing losses. They also measured DLFs for a 606-Hz tone, using a three-alternative forced-choice task (pick the odd one out). They found a small but significant correlation between FFR synchronisation strength and DLFs even when the effects of both absolute thresholds and age were partialled out. These results support the idea that DLFs at low frequencies depend on the use of temporal cues and that individual differences in DLFs partly reflect differences in the precision of phase locking.

## Conclusions

The relative role of place and temporal mechanisms in auditory frequency discrimination was assessed for a centre frequency of 2 kHz. It was hypothesized that: (1) FMDLs depend partly on a temporal mechanism and partly on a place mechanism for an FM rate of 2 Hz, while FMDLs depend primarily on a place mechanism for an FM rate of 20 Hz; (2) DLFs depend primarily on a temporal mechanism; (3) performance of the TFS1 task depends primarily on a temporal mechanism. Based on these hypotheses, four predictions were made: (P1) FM2 and FM20 values should be moderately correlated; (P2) FM2 values should be moderately correlated with DLF and TFS values; (P3) FM20 values should not be correlated with DLF or TFS values; (P4) DLF and TFS values should be highly correlated.

Based on individual differences across 63 young normal-hearing participants, P1, P3 and P4 were supported while P2 was weakly supported. It is suggested that the weak support for P2 was related to the low SL of the stimuli (30 dB), which promoted the use of place cues for FM detection even for the 2-Hz FM rate. A principal components analysis showed that two components accounted for 86% of the variance in the data. PC1 can be interpreted as related to the use of temporal cues, while PC2 can be interpreted as reflecting the use of place cues. Overall, the results support the idea that frequency discrimination at 2 kHz depends partly on the use of temporal cues for FM detection at a 2-Hz rate, mainly on the use of place cues for FM detection at a 20-Hz rate, and mainly on the use of temporal cues for DLFs and TFS. The relative importance of temporal and place cues for higher centre frequencies needs further exploration.

## Data availability

The data for the current study are available from the corresponding author on reasonable request.

Received: 14 June 2023; Accepted: 13 August 2023

Published online: 23 August 2023

## References

1. von Békésy, G. *Experiments in Hearing* 1–745 (McGraw-Hill, 1960).
2. Robles, L. & Ruggero, M. A. Mechanics of the mammalian cochlea. *Physiol. Rev.* **81**, 1305–1352 (2001).
3. Zwicker, E. Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs (The foundations for determining the information capacity of the auditory system). *Acustica* **6**, 356–381 (1956).
4. Zwislocki, J. J. & Nguyen, N. Place code for pitch: A necessary revision. *Acta Otolaryngol.* **119**, 140–145 (1999).
5. Rose, J. E., Brugge, J. F., Anderson, D. J. & Hind, J. E. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.* **30**, 769–793 (1967).
6. Siebert, W. M. Frequency discrimination in the auditory system: Place or periodicity mechanisms. *Proc. IEEE* **58**, 723–730 (1970).

7. Heinz, M. G., Colburn, H. S. & Carney, L. H. Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. *Neural Comput.* **13**, 2273–2316 (2001).
8. Heinz, M. G., Colburn, H. S. & Carney, L. H. Evaluating auditory performance limits: II. One-parameter discrimination with random-level variation. *Neural Comput.* **13**, 2317–2338 (2001).
9. Kiang, N.Y.-S., Watanabe, T., Thomas, E. C. & Clark, L. F. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* 1–154 (MIT Press, 1965).
10. Palmer, A. R. & Russell, I. J. Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear. Res.* **24**, 1–15 (1986).
11. Goldstein, J.L. & Sruulovicz, P. Auditory-nerve spike intervals as an adequate basis for aural frequency measurement. In *Psychophysics and Physiology of Hearing* (eds. Evans, E. F. & Wilson, J. P.) 337–346 (Academic Press, 1977).
12. Moore, B. C. J. & Ernst, S. M. A. Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code. *J. Acoust. Soc. Am.* **132**, 1542–1547 (2012).
13. Verschooten, E., Robles, L. & Joris, P. X. Assessment of the limits of neural phase-locking using mass potentials. *J. Neurosci.* **35**, 2255–2268 (2015).
14. Verschooten, E. *et al.* The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints. *Hear. Res.* **377**, 109–121 (2019).
15. Shower, E. G. & Biddulph, R. Differential pitch sensitivity of the ear. *J. Acoust. Soc. Am.* **2**, 275–287 (1931).
16. Moore, B. C. J. & Şek, A. Effects of carrier frequency, modulation rate and modulation waveform on the detection of modulation and the discrimination of modulation type (AM vs FM). *J. Acoust. Soc. Am.* **97**, 2468–2478 (1995).
17. Lacher-Fougère, S. & Demany, L. Modulation detection by normal and hearing-impaired listeners. *Audiology* **37**, 109–121 (1998).
18. Moore, B. C. J. & Şek, A. Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *J. Acoust. Soc. Am.* **100**, 2320–2331 (1996).
19. Şek, A. & Moore, B. C. J. Frequency discrimination as a function of frequency, measured in several ways. *J. Acoust. Soc. Am.* **97**, 2479–2486 (1995).
20. Moore, B. C. J., Mariathasan, S. & Şek, A. P. Effects of age on the discrimination of amplitude and frequency modulation for 2- and 10-Hz rates. *Acta Acust. United Ac.* **104**, 778–782 (2018).
21. Moore, B. C. J. & Şek, A. Discrimination of the phase of amplitude modulation applied to different carriers: Effects of modulation rate and modulation depth for young and older subjects. *J. Acoust. Soc. Am.* **146**, 1696–1704 (2019).
22. Moore, B. C. J., Mariathasan, S. & Şek, A. P. Effects of age and hearing loss on the discrimination of amplitude and frequency modulation for 2- and 10-Hz rates. *Trends Hear.* **23**, 1–12 (2019).
23. Moore, B.C.J. & Vinay. Assessing mechanisms of frequency discrimination by comparison of different measures over a wide frequency range. *Sci. Rep.* **13**, 11379 (2023).
24. Ernst, S. M. & Moore, B. C. J. Mechanisms underlying the detection of frequency modulation. *J. Acoust. Soc. Am.* **128**, 3642–3648 (2010).
25. Whiteford, K. L. & Oxenham, A. J. Using individual differences to test the role of temporal and place cues in coding frequency modulation. *J. Acoust. Soc. Am.* **138**, 3093–3104 (2015).
26. Whiteford, K. L., Kreft, H. A. & Oxenham, A. J. Assessing the role of place and timing cues in coding frequency and amplitude modulation as a function of age. *J. Assoc. Res. Otolaryngol.* **18**, 619–633 (2017).
27. Whiteford, K. L., Kreft, H. A. & Oxenham, A. J. The role of cochlear place coding in the perception of frequency modulation. *Elife* **9**, e58468 (2020).
28. Moore, B. C. J. Frequency difference limens for short-duration tones. *J. Acoust. Soc. Am.* **54**, 610–619 (1973).
29. Moore, B. C. J. Comparison of frequency DLs for pulsed tones and modulated tones. *Brit. J. Audiol.* **10**, 17–20 (1976).
30. Micheyl, C., Delhommeau, K., Perrot, X. & Oxenham, A. J. Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.* **219**, 36–47 (2006).
31. Semal, C. & Demany, L. Individual differences in the sensitivity to pitch direction. *J. Acoust. Soc. Am.* **120**, 3907–3915 (2006).
32. Vinay & Moore, B.C.J. Effects of the use of personal music players on amplitude modulation detection and frequency discrimination. *J. Acoust. Soc. Am.* **128**, 3634–3641 (2010).
33. Hopkins, K. & Moore, B. C. J. Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J. Acoust. Soc. Am.* **122**, 1055–1068 (2007).
34. Moore, B. C. J., Hopkins, K. & Cuthbertson, S. J. Discrimination of complex tones with unresolved components using temporal fine structure information. *J. Acoust. Soc. Am.* **125**, 3214–3222 (2009).
35. Şek, A. & Moore, B. C. J. Implementation of two tests for measuring sensitivity to temporal fine structure. *Int. J. Audiol.* **51**, 58–63 (2012).
36. de Boer, E. Pitch of inharmonic signals. *Nature* **178**, 535–536 (1956).
37. Schouten, J. F., Ritsma, R. J. & Cardozo, B. L. Pitch of the residue. *J. Acoust. Soc. Am.* **34**, 1418–1424 (1962).
38. Moore, G. A. & Moore, B. C. J. Perception of the low pitch of frequency-shifted complexes. *J. Acoust. Soc. Am.* **113**, 977–985 (2003).
39. Jackson, H. M. & Moore, B. C. J. The role of excitation-pattern and temporal-fine-structure cues in the discrimination of harmonic and frequency-shifted complex tones. *J. Acoust. Soc. Am.* **135**, 1356–1570 (2014).
40. Marmel, F. *et al.* The role of excitation-pattern cues in the detection of frequency shifts in bandpass-filtered complex tones. *J. Acoust. Soc. Am.* **137**, 2687–2697 (2015).
41. Moore, B. C. J. Distribution of auditory-filter bandwidths at 2 kHz in young normal listeners. *J. Acoust. Soc. Am.* **81**, 1633–1635 (1987).
42. Wright, B. A. Auditory filter asymmetry at 2000 Hz in 80 normal-hearing ears. *J. Acoust. Soc. Am.* **100**, 1717–1721 (1996).
43. Marmel, F. *et al.* Subcortical neural synchrony and absolute thresholds predict frequency discrimination independently. *J. Assoc. Res. Otolaryngol.* **14**, 757–766 (2013).
44. Bones, O., Hopkins, K., Krishnan, A. & Plack, C. J. Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia* **58**, 23–32 (2014).
45. Şek, A. P. & Moore, B. C. J. *Guide to Psychoacoustics* 1–164 (Adam Mickiewicz University Press, 2021).
46. Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R. & Alcántara, J. I. A test for the diagnosis of dead regions in the cochlea. *Br. J. Audiol.* **34**, 205–224 (2000).
47. Moore, B. C. J., Glasberg, B. R. & Baer, T. A model for the prediction of thresholds, loudness and partial loudness. *J. Audio Eng. Soc.* **45**, 224–240 (1997).
48. Glasberg, B. R. & Moore, B. C. J. Derivation of auditory filter shapes from notched-noise data. *Hear. Res.* **47**, 103–138 (1990).
49. Moore, B. C. J. & Gockel, H. Resolvability of components in complex tones and implications for theories of pitch perception. *Hear. Res.* **276**, 88–97 (2011).
50. Steiger, J. H. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **87**, 245–251 (1980).
51. Metsalu, T. & Vilo, J. ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43**, W566–W570 (2015).
52. Whiteford, K. L. & Oxenham, A. J. Sensitivity to frequency modulation is limited centrally. *J. Neurosci.* **43**, 3687–3695 (2023).
53. Moore, B. C. J., Vickers, D. A. & Mehta, A. The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. *Int. J. Audiol.* **51**, 715–721 (2012).

54. Moore, B. C. J. & Glasberg, B. R. Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hear. Res.* **28**, 209–225 (1987).
55. Sumner, C. J. *et al.* Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans. *Proc. Natl. Acad. Sci. USA* **115**, 11322–11326 (2018).
56. Oxenham, A. J. & Shera, C. A. Estimates of human cochlear tuning at low levels using forward and simultaneous masking. *J. Assoc. Res. Otolaryngol.* **4**, 541–554 (2003).
57. Sheft, S. & Yost, W. A. Temporal integration in amplitude modulation detection. *J. Acoust. Soc. Am.* **88**, 796–805 (1990).
58. Ernst, S. M. & Moore, B. C. J. The role of time and place cues in the detection of frequency modulation by hearing-impaired listeners. *J. Acoust. Soc. Am.* **131**, 4722–4731 (2012).
59. Kidd, G. R., Watson, C. S. & Gygi, B. Individual differences in auditory abilities. *J. Acoust. Soc. Am.* **122**, 418–435 (2007).

## Acknowledgements

We thank Hedwig Gockel and two reviewers for helpful comments on an earlier version of this paper.

## Author contributions

B.C.J.M. designed the study. V. obtained project funding, supervised data collection, and collated the results. B.C.J.M. analysed the results and wrote the first manuscript draft. V. revised the manuscript draft. Both authors agreed upon the final version of the manuscript.

## Funding

Open access funding provided by Norwegian University of Science and Technology. This study was funded by Norges Teknisk-Naturvitenskapelige Universitet.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to V.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023