Frequency discrimination of complex tones; assessing the role of component resolvability and temporal fine structure[a)]

Brian C. J. Moore,[b)] Brian R. Glasberg, and Helen J. Flanagan
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

Joe Adams
Faculty of Music, University of Cambridge, West Road, Cambridge CB3 9DP, England

(Received 25 May 2005; revised 11 October 2005; accepted 14 October 2005)

Thresholds for discriminating the fundamental frequency (F0) of a complex tone, F0DLs, are small when low harmonics are present, but increase when the number of the lowest harmonic, N, is above eight. To assess whether the relatively small F0DLs for N in the range 8–10 are based on (partly) resolved harmonics or on temporal fine structure information, F0DLs were measured as a function of N for tones with three successive harmonics which were added either in cosine or alternating phase. The center frequency was 2000 Hz, and N was varied by changing the mean F0. A background noise was used to mask combination tones. The value of F0 was roved across trials to force subjects to make within-trial comparisons. N was roved by ±1 for every stimulus, to prevent subjects from using excitation pattern cues. F0DLs were not influenced by component phase for $N=6$ or 7, but were smaller for cosine than for alternating phase once N exceeded 7, suggesting that temporal fine structure plays a role in this range. When the center frequency was increased to 5000 Hz, performance was much worse for low N, suggesting that phase locking is important for obtaining low F0DLs with resolved harmonics. © *2006 Acoustical Society of America.* [DOI: 10.1121/1.2139070]

PACS number(s): 43.66.Hg, 43.66.Nm, 43.66.Fe [JHG] Pages: 480–490

I. INTRODUCTION

The pitch of a periodic complex tone usually corresponds to its fundamental frequency (F0), even when the fundamental component is not present (Schouten, 1940), which has been called the phenomenon of the "missing fundamental." The pitch evoked by a group of harmonics is called residue pitch, virtual pitch, or low pitch. A low pitch can be heard even when the fundamental component is masked by low-frequency noise (Licklider, 1956), which proves that the pitch is not produced by a distortion component at F0.

Several classes of model have been proposed to explain the perception of the low pitch of complex tones. In one class, the pattern-recognition models (de Boer, 1956; Thurlow, 1963; Goldstein, 1973; Terhardt, 1974), it is assumed that two stages are involved. The first stage is a frequency analysis to determine the frequencies (or "pitches" in Terhardt's terminology) of some of the individual sinusoidal components of the complex tone. This might depend on both place and temporal analysis (Srulovicz and Goldstein, 1983; Shamma, 1985; Shamma and Klein, 2000). The second stage is a pattern recognizer which determines the pitch of the complex tone from the frequencies of the resolved components. For this class of model, it should only be possible for the auditory system to derive a low pitch when there is at least one resolvable harmonic.

An alternative class of model is based on temporal analysis of the waveforms evoked on the basilar membrane (Schouten, 1940; de Boer, 1956). The lower harmonics in a complex tone are partially resolved by the basilar membrane, each leading to a peak in response at the appropriate place. The waveform evoked on the basilar membrane by a resolved low harmonic reflects the periodicity of that harmonic, and the pattern of phase locking evoked at the place of maximal response to the harmonic is similar to what would be evoked by a sine wave with the same frequency as the harmonic. However, higher harmonics interfere on the basilar membrane. The temporal pattern of the waveform evoked by high harmonics reflects the F0 of periodic sounds. Temporal theories assume that pitch is derived from the temporal pattern of neural spikes arising from one or more points on the basilar membrane where harmonics are interfering. According to these theories, interference of harmonics on the basilar membrane is required for a low pitch to be heard.

Research has shown that a low pitch can be perceived both when only low resolved harmonics are present and when only high unresolved harmonics are present (Ritsma, 1962; 1963; Moore and Rosen, 1979; Moore and Glasberg, 1988; Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994; Kaernbach and Bering, 2001). This has led to models in which information from both low and high harmonics is used to determine low pitch (Moore, 1982; van Noorden, 1982; Meddis and Hewitt, 1991; Meddis and O'Mard, 1997).

[a)] Part of this work was presented at the 28th Midwinter meeting of the Association for Research in Otolaryngology

[b)] Electronic mail: bcjm@cam.ac.uk

Accepting that a (weak) low pitch can be derived from the temporal waveform evoked on the basilar membrane by a group of high unresolved harmonics, the question arises as to whether the pitch is determined by the envelope or by the temporal fine structure. This has been assessed using amplitude-modulated (AM) tones. Consider an AM tone with carrier frequency fc and modulation frequency fm. The spectrum of this sound contains components with frequencies $fc-fm$, fc, and $fc+fm$. Consider a case where, initially, fc is an integer multiple of fm, for example $fc=2000$ Hz and $fm=200$ Hz. In this case the sound has a low pitch corresponding to 200 Hz. If now fc is shifted upwards (say to 2030 Hz) keeping fm fixed, the envelope repetition rate remains the same (200 Hz), but a slight upward shift in pitch is heard (de Boer, 1956; Schouten *et al.*, 1962; Patterson, 1973; Moore and Moore, 2003). This has been explained in terms of the temporal fine structure of the waveform (either of the stimulus itself, or the waveform evoked on the basilar membrane). The time interval between peaks in the temporal fine structure close to adjacent envelope maxima is slightly reduced by the upward shift in fc, and the shifted pitch corresponds to this time interval.

Moore and Moore (2003) modified the stimuli used to assess the role of temporal fine structure so as to eliminate possible cues related to shifts in the excitation pattern with shifts in fc. They used complex tones with more than three components and passed the tones through a fixed bandpass filter. They assessed the effect on pitch of shifting all components upwards in frequency by a fixed amount, keeping the spacing between components (and therefore the envelope repetition rate) the same. They showed that when the nominal harmonic number of the lowest component in the complex tone, N, was relatively high (N above about 14), there was no shift in pitch, whereas when N was 9, significant pitch shifts occurred. Moore and Moore argued that, when a complex tone contains only very high harmonics, the pitch is determined from the envelope periodicity rather than from the temporal fine structure. The pitch shift found when N was 9 could be explained either in terms of a sensitivity to the temporal fine structure or in terms of a pattern-matching process to (partially) resolved harmonics.

Hall *et al.* (2003) also studied the relative importance of envelope and temporal fine structure cues. They compared modulation-rate discrimination thresholds for amplitude-modulated (AM) and quasi-frequency-modulated (QFM) tones. Thresholds were similar for AM and QFM tones whose carrier fell in a low spectral region, but as the spectral region of the carrier frequency increased, performance worsened more rapidly for QFM than for AM tones. Hall *et al.* argued that when components occupy relatively low spectral regions, phase locking both to the fine structure and to the envelope can be used as a cue. However, as the spectral region occupied by the components increases, phase locking to the fine structure becomes less robust, whereas phase locking to the envelope remains as a potentially strong cue. In the study of Hall *et al.* the modulation rate was fixed at 100 or 200 Hz and performance was measured for several carrier frequencies over the range 1500 to 6000 Hz. Hence, in their study it was not possible to separate the effects of carrier frequency *per se* from the effects of harmonic number (determined by the ratio of carrier to modulation frequency).

The ability to detect changes in F0 for a harmonic complex tone is usually good when the tone contains low resolved harmonics and poor when the tone contains only high unresolved harmonics (Hoekstra and Ritsma, 1977; Moore and Glasberg, 1988; Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994; Kaernbach and Bering, 2001; Bernstein and Oxenham, 2003, 2005). It is usually assumed that performance is good for tones containing low harmonics because these harmonics are resolved in the peripheral auditory system, although Bernstein and Oxenham (2003) presented evidence suggesting that resolvability *per se* is not the critical factor. Performance may be poor for very high harmonics because only information about the temporal envelope of the stimulus is available, temporal fine structure information being lost, the evidence for which was described earlier (Moore and Moore, 2003).

If complex tones are filtered or synthesized so as to contain only a few harmonics, F0 discrimination is usually good when N is below 8. However, as N is increased above 8, performance starts to worsen, until it reaches a plateau of relatively poor performance when N is 12–13 (Hoekstra and Ritsma, 1977; Bernstein and Oxenham, 2003). The question addressed in this paper is: what mechanisms underlie the F0 discrimination of complex tones when N is in the range 8–11, where performance is still relatively good? The good performance in this range might occur because the harmonics are still partially resolved. Data on the audibility of individual partials in steady complex tones (Plomp, 1964; Plomp and Mimpen, 1968; Moore and Ohgushi, 1993) suggest that the lowest five harmonics can be "heard out" with high accuracy, but accuracy decreases for harmonics above the fifth, and harmonics above the eighth cannot be heard out at all. However, in studies in which the "target" harmonic was pulsed on and off, it has been reported that harmonics up to about the 11th could be heard out (Gibson, 1970; Bernstein and Oxenham, 2003). If this estimate is correct, then the good F0 discrimination for tones with lowest harmonics in the range 8–11 might depend on (partially) resolved harmonics. Also, the pitch-shift effect found for stimuli with components in this range (Schouten *et al.*, 1962; Moore and Moore, 2003) could be explained in terms of the best-fitting fundamental frequency obtained via a pattern-recognition process (Goldstein, 1973). An alternative possibility is that, in this range, listeners can extract information from the temporal fine structure of the sound and not just the envelope. Use of temporal fine structure information could lead to greater precision in the estimation of F0 than use of envelope information.

In the present experiment, we assessed the two possibilities discussed above by comparing F0 discrimination for complex tones with components added in two different phase relationships. One, cosine phase, was chosen to give waveforms on the basilar membrane with a high peak factor (ratio of peak to root-mean-square value), which should lead to optimal use of temporal fine structure information, since there are only a few high-amplitude peaks in the fine structure in each stimulus period. This is illustrated in the upper

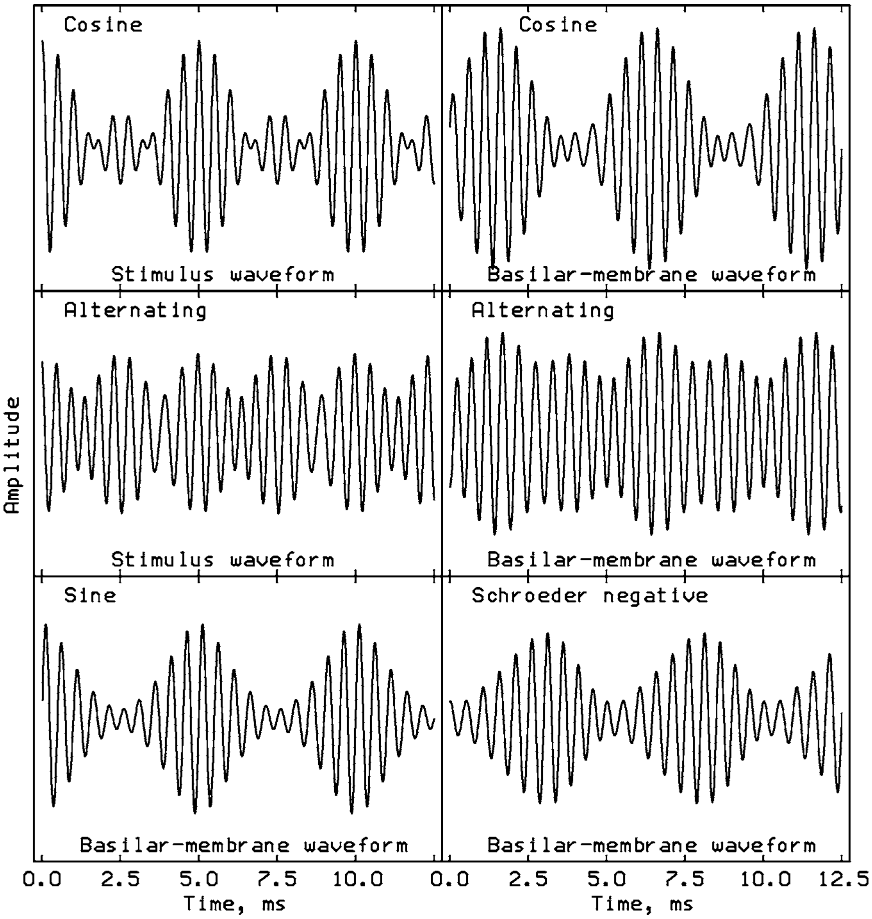

FIG. 1. The top and middle rows illustrate the effect of component phase on the waveforms of three-component complex tones. The top-left trace shows the waveform of a cosine-phase stimulus (the 9th, 10th, and 11th harmonics of a 200-Hz F0) and the top-right trace shows a simulation of the waveform evoked by that stimulus on the basilar membrane at the place tuned to 2000 Hz, calculated as described by Alcántara *et al.* (2003). The middle two traces show corresponding waveforms for an alternating-phase stimulus. The lower two traces simulate basilar-membrane waveforms evoked by 11-component stimuli similar to those used by Houtsma and Smurzynski (1990). The tones contained harmonics 9–19 of a 200-Hz F0. For the bottom-left trace, the components were added in sine phase. For the bottom-right trace, the components were added in Schroeder-negative phase.

traces of Fig. 1; the left trace shows the waveform of one of the stimuli (the 9th, 10th, and 11th harmonics of a 200-Hz F0) and the right trace shows a simulation of the waveform evoked by that stimulus on the basilar membrane, calculated as described by Alcántara *et al.* (2003) and assuming that the curvature of the phase response of the auditory filter centered at 2000 Hz is 2.5×10^{-5} rad/Hz^2 (Lentz and Leek, 2001; Oxenham and Dau, 2001). The other, alternating phase, was chosen to give waveforms with lower peak factors, with many fine structure peaks of similar magnitude, which should lead to poorer use of temporal fine structure information. This is illustrated in the middle traces of Fig. 1 (the lower traces are explained later). Note that the envelope of the alternating-phase stimulus has two peaks per period. For alternating-phase stimuli with many high harmonics, this can lead to a near-octave shift in pitch relative to a stimulus with sine-phase or cosine-phase components (Patterson, 1987; Shackleton and Carlyon, 1994). However, for our three-component stimuli, while the simulated basilar-membrane waveform still has two envelope peaks per period, one of these envelope peaks is markedly bigger than the other. Informal listening by the experimenters and some of the musically trained subjects indicated that, for a given F0, the pitch of our alternating-phase stimuli was the same as that of the cosine-phase stimuli (although the pitch of the former was often less clear).

It is generally assumed that the relative phase of the components in a complex tone does not affect pitch perception when the components are resolved, but can affect pitch perception when they are not resolved (Moore, 1977; Carlyon and Shackleton, 1994; Shackleton and Carlyon, 1994; Houtsma and Smurzynski, 1990; Bernstein and Oxenham, 2005). Hence, we expected to find no effect of component phase when the complex tones contained only low harmonics, but poorer F0 discrimination for alternating than for cosine phase when the tones contained only very high harmonics (say above the 14th). The question of interest was: would there be a phase effect when the lowest harmonic was in the range 8–11? If so, this would imply that the harmonics were unresolved, and this, combined with the pitch-shift effect described earlier, would provide evidence for a role of temporal fine structure derived from unresolved harmonics. If there was not a phase effect, this would support the idea that the good F0 discrimination for harmonics in that range is linked to partial resolution of the harmonics.

In experiment 1, all complex tones had frequency components centered at 2000 Hz, a frequency for which temporal fine structure is represented in the auditory nerve of the chinchilla (Javel, 1980) and the cat (Cariani and Delgutte, 1996a, b) and probably in humans (Moore, 1973). We chose this center frequency as pilot experiments had shown that many subjects did not clearly perceive a low pitch when the center frequency was lower (e.g., 1000 Hz) and the harmonic numbers were relatively high, whereas all subjects reported hearing a low pitch when the center frequency was 2000 Hz and the components were added in cosine phase. In experiment 2, the complex tones were centered at 5000 Hz, a frequency for which phase locking is very weak, and temporal fine structure information would probably be unusable. The comparison of results for the two frequencies was intended to give some insight into the relative effectiveness of temporal fine structure and temporal envelope information.

There are two previous experiments that have some resemblance to the ones reported here. Houtsma and Smurzynski (1990) measured F0DLs as a function of N for complex tones with 11 harmonics, which were added either in sine phase or Schroeder negative phase (Schroeder, 1970). F0DLs were smaller for the sine-phase than for the Schroeder-negative-phase stimuli for $N=13$ and higher (although the statistical significance of the phase effect was not assessed), but there was no phase effect for $N=7$ or 10. Values of N between 10 and 13 were not used. There are four ways in which the experiments of Houtsma and Smurzynski differ from our experiments. First, their F0 was always close to 200 Hz, so the spectral region occupied by the harmonics varied with N. This means that it is impossible to separate the effect of spectral region from the effect of harmonic number. In contrast, we used stimuli centered on two fixed spectral regions, one where phase locking is available and one where it is not. Second, their use of complex tones with many harmonics makes it difficult to determine which harmonics were most important in determining the F0DLs. We used stimuli with only three harmonics. Third, the use of Schroeder negative phase may not have been optimal for minimizing the use of temporal information; Houtsma and Smurzynski pointed out that the two phases used by them "were merely chosen as two rather convenient phase configurations for which some elementary masking data exist and on the basis of which one might expect different pitch behavior." This is illustrated by the two lower traces in Fig. 1. These show simulations of basilar membrane responses at a place tuned to 2000 Hz, calculated in the way described earlier, for stimuli similar to those used by Houtsma and Smurzynski. In both cases, the complex tones contained harmonics 9 to 19 of a 200-Hz F0. For the left trace, the components were added in sine phase, while for the right trace they were added in Schroeder-negative phase. Both waveforms have a single distinct envelope maximum per period, and the peak factors of the waveforms are similar for the two cases, being 2.37 for the sine-phase complex and 2.16 for the sine-phase complex. In contrast, the phases of the stimuli used by us were specifically chosen, based on simulations of the waveforms on the basilar membrane (Alcántara *et al.*, 2003), so as to give waveforms with high peak factors (cosine phase) and low peak factors (alternating phase) on the basilar membrane. Finally, we used values of N spaced in steps of 1 instead of the steps of 3 used by Houtsma and Smurzynski.

The other experiment that is comparable to ours was published recently by Bernstein and Oxenham (2005). They compared F0 discrimination for multiple-component complex tones filtered into two frequency regions, a "low" region, for which phase-locking information was thought to be available, and a "high" region, for which phase-locking information was thought to be "greatly reduced." Components in the complex tones were added in either sine phase or random phase. For $N>15$, F0DLs were larger for random-phase than for sine-phase stimuli. For $N\approx 10$, there was a small phase effect for stimuli in the high spectral region but no phase effect for stimuli in the low spectral region. No values of N between 10 and 15 were tested. Generally, the patterns of performance were similar for the two frequency regions, which led Bernstein and Oxenham to conclude "This implies that phase locking to the stimulus fine structure did not play a significant role overall in F0 discrimination for the stimuli used in this experiment." However, they did note that the stimuli used in the "high" spectral region had audible components extending down to 3.28 kHz "where phase-locking to the stimulus fine structure might still have been available." The results reported here show clear differences between the results for center frequencies of 2000 and 5000 Hz, consistent with a strong role of temporal fine structure information for the lower center frequency. Our experiments also differ from that of Bernstein and Oxenham in our use of closely spaced values of N, and of stimuli with only three harmonics; this allowed us to determine more precisely which harmonics were required to observe an effect of phase on F0DLs.

II. EXPERIMENT 1: CENTER FREQUENCY 2000 HZ

A. Method

Thresholds for the discrimination of F0 (F0DLs) were measured using a two-interval two-alternative forced-choice procedure. Subjects had to indicate, by pressing one of two buttons, which of the two intervals contained the stimulus with higher F0, and feedback was provided after each trial by lights on the response box. For a given center F0, $F0_c$, and a given difference, $\Delta F0$, one stimulus in a trial had an F0 of $F0_c+0.5\Delta F0$ and the other had an F0 of $F0_c-0.5\Delta F0$. A three-down, one-up geometric tracking procedure was used to estimate the 79% correct point on the psychometric function (Levitt, 1971). Twelve turnpoints were obtained. A run started with a relatively large value of $\Delta F0$, typically 20% of $F0_c$. In most cases, this was sufficiently large to ensure that the difference was easily discriminable. The value of $\Delta F0$ was increased or decreased by a factor of 1.414 until four turnpoints had occurred. Then, the factor was reduced to 1.189 and eight further turnpoints were obtained. The geometric mean of the frequency differences at the last eight turnpoints was taken as the estimate of threshold for that run. At least three estimates were obtained for each combination of stimulus condition and F0. In cases where the variability of the three estimates was high, up to five additional runs were obtained, and the final threshold was taken as the geometric mean of all estimates.

B. Stimuli

Each complex tone contained three successive equal-amplitude harmonics, each with a level of 60.2 dB SPL (overall level of 65 dB SPL). The nominal frequency of the center component was 2000 Hz. The nominal number, N, of the lowest harmonic ranged from 6 to 15. N was varied by changing the mean $F0_c$. To force subjects to make within-trial comparisons, rather than making comparisons with a long-term memory for F0, the value of $F0_c$ was roved by

±10% across trials. To reduce the ability of subjects to perform the task by comparing the pitches of individual resolved harmonics (for example the lowest harmonic or highest harmonic), or from using excitation pattern cues for unresolved harmonics (Moore and Moore, 2003), the value of N was roved by ±1 independently for each of the two intervals within each forced-choice trial (Houtsma and Smurzynski, 1990). For example, for a value of N of 8, the harmonic numbers in each interval were selected randomly from 7, 8, 9 or 8, 9, 10 or 9, 10, 11, independently of which interval had the higher F0.

The components in the complex tones were added with two different starting phases. In one condition, cosine phase, all components started at a positive peak in the waveform. In the other condition, alternating phase, the phase of the middle component was shifted by 90°. Each tone had rise/fall times of 20 ms and a steady-state duration of 480 ms. The silent interval between the two tones within a trial was 500 ms.

To prevent discrimination of F0 based on combination tones, a continuous background noise was presented during each run. The noise was "threshold equalizing noise" (Moore *et al.*, 2000) designed to give equal masked thresholds at all frequencies over the range 50 to 2800 Hz; the upper cutoff frequency was 3000 Hz. The noise was synthesized by adding sinusoidal components spaced at 0.1-Hz intervals with appropriate amplitudes and random phases. The resulting 10-s noise segment was seamlessly recycled to give continuous noise and recorded onto CDR. During the experiment the noise was replayed from the CDR. The noise spectrum level (*re* 20 μPa) was 6 dB at 2000 Hz, corresponding to a level per ERB_N of 30 dB SPL, where ERB_N stands for the average value of the equivalent rectangular bandwidth of the auditory filter for young normal-hearing listeners at moderate sound levels (Glasberg and Moore, 1990).

The complex tones were generated digitally on-line using a Tucker-Davis Technologies (TDT) system II. The tones were played through a 16-bit digital-to-analog converter (TDT, DD1) at a 50-kHz sampling rate, low-pass filtered at 8 kHz (Kemo VBF8/04), attenuated (TDT, PA4), mixed with the noise (TDT, SM3), and presented via a headphone buffer (TDT, HB6), a manual attenuator (Hatfield 2125), and one earpiece of a Sennheiser HD580 headset.

C. Subjects and training

Five normal hearing subjects were tested (three males and two females), four of whom had some musical training. Their ages ranged from 20 to 35 years. All subjects had absolute thresholds better than 20 dB HL over the range of audiometric frequencies from 250 to 8000 Hz. All subjects except author HF were paid for their services. Pilot experiments showed that subjects initially found the randomization of harmonic number to be very distracting, and they often performed poorly even when N was 6 or 7, because they based their judgments on the shift in frequency of the lowest harmonic rather than the shift in F0. To overcome this problem, subjects were initially trained using complex tones with seven harmonics (each with a level of 60.2 dB SPL), initially without and then with randomization of ±1 for the number of the lowest harmonic. For tones with more harmonics, the change in F0 was more salient, and the randomization of harmonic number was less distracting. Once subjects achieved stable performance for the complex tones with seven harmonics, even when N was randomized, the number of harmonics was reduced to three. Subjects were given at least 3 h of training using complex tones with three harmonics (i.e., the stimuli to be used in the main experiment) before data were accepted.

D. Results and discussion

Figure 2 shows the results for each subject and the geometric mean across subjects (bottom-right panel). Results for the cosine-phase stimuli are shown by open circles and results for the alternating-phase stimuli are shown by filled circles. The pattern of results for the cosine-phase stimuli is similar to that found by Hoekstra and Ritsma (1977) for discrimination of the rate of click trains passed through a $\frac{1}{3}$-oct filter centered at 2000 Hz. F0DLs are small for the lowest harmonic numbers ($N=6$ or 7) and increase progressively as N is increased to about 12. The F0DLs then flatten off. There is no clear effect of component phase for $N=6$ or 7. However, alternating phase led to higher F0DLs than cosine phase for harmonic numbers of 8 and above.

A within-subjects analysis of variance (ANOVA) was conducted on the logarithms of the F0DLs (as the standard deviation of the F0DLs was roughly proportional to the mean value of the F0DLs) with factors lowest harmonic number ($N=6$ to 15) and phase (cosine or alternating). Both main effects were significant: for harmonic number, $F(9,36)=24.59$, $p<0.001$; for phase, $F(1,4)=93.09$, $p<0.001$. The interaction of harmonic number and phase was also significant: $F(9,36)=2.64$, $p<0.02$. *Posthoc* tests of the effect of phase, using the least-significant-differences (LSD) test, showed no significant effect for $N=6$ or 7. The effect of phase was significant at $p<0.02$ for $N=10$ to 15 and significant at $p<0.001$ for $N=8$ and 9.

Bernstein and Oxenham (2003) also assessed the effect of component phase on F0 discrimination as a function of the number of the lowest harmonic in complex tones. In contrast to us, they did not find a significant effect of phase or a significant interaction of phase with harmonic number. This may be due to several differences between the experiments: (1) They compared sine phase and random phase rather than the cosine and alternating phase used by us. (2) They used complex tones with 12 components, which would have evoked activity over a range of places along the basilar membrane, with different waveforms at different places. (3) They tested different subjects for the sine-phase and random-phase conditions.

Our results are broadly consistent with those of Houtsma and Smurzynski (1990), although they found no phase effect (sine phase versus Schroeder negative phase) when N was 10 and a clear phase effect when N was 13 (no intermediate values were assessed). The lack of a phase effect for $N=10$ (while we found a clear effect) may have occurred because the waveform evoked on the basilar membrane by the com-

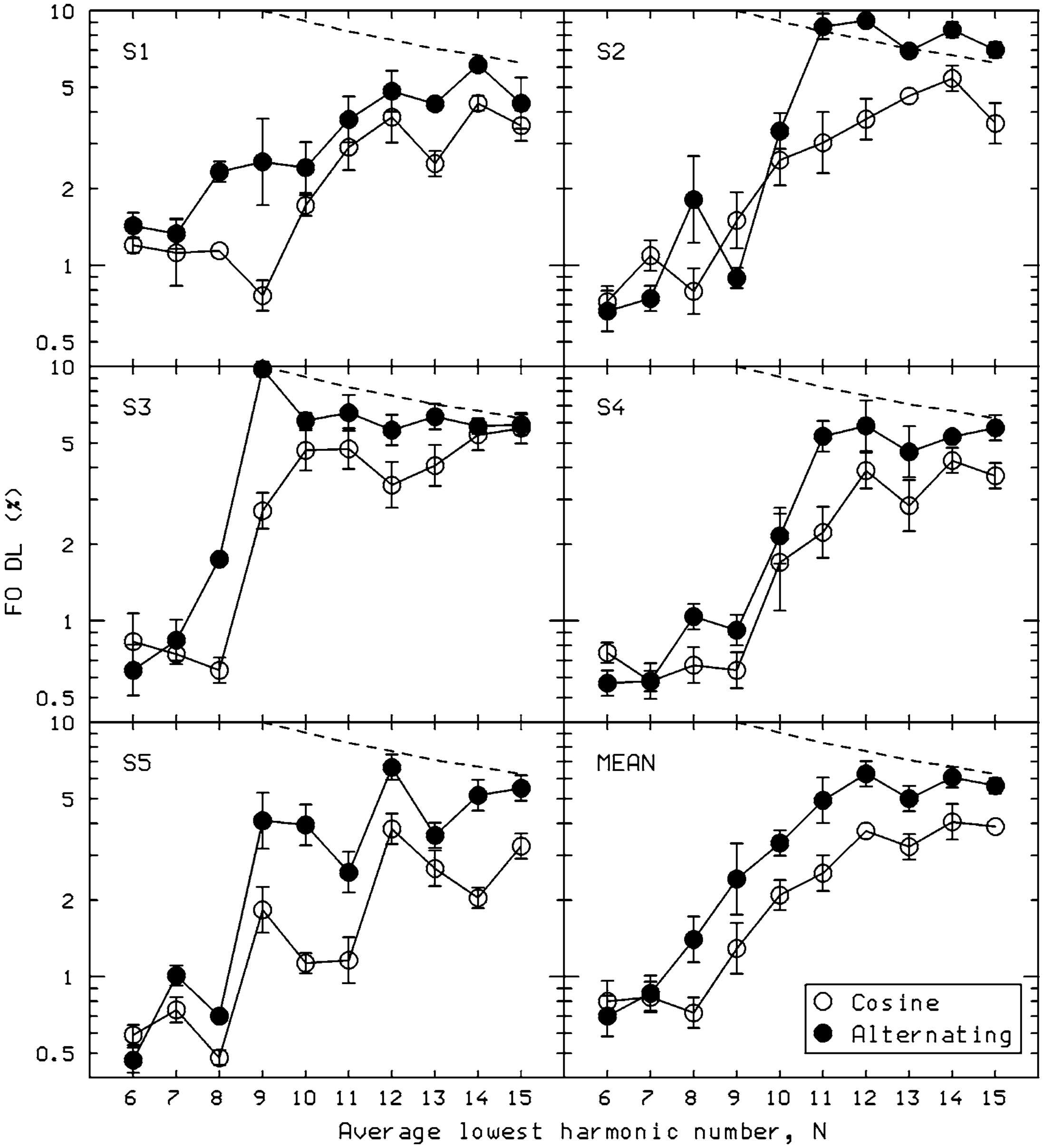

FIG. 2. Individual and mean (bottom-right) results for experiment 1 (harmonics centered at 2000 Hz). F0DLs, expressed as a percentage of $F0_c$, are plotted on a logarithmic scale as a function of the average number, N, of the lowest harmonic. Open and filled circles show results for components added in cosine phase and alternating phase, respectively. Error bars in the results for individual subjects show ± one standard error (SE) of the mean across repeated runs. Error bars in the mean results show ± one standard error (SE) of the mean across subjects. The dashed lines show the smallest F0DLs that could be achieved if performance were based on discrimination of the frequency of the lowest harmonic in the complex tones.

plex tone with components added in Schroeder negative phase probably still had distinct envelope peaks and dips, allowing effective use of temporal information (see the bottom-right trace in Fig. 1).

Our results are also broadly consistent with those of Bernstein and Oxenham (2005) for the "low" spectral region; they measured F0DLs for complex tones with many components added either in sine phase or random phase. However, they found a significant effect of phase only for $N>10$, whereas we found a significant effect for $N=8$ to 10. Again, the lack of a phase effect in their data for $N=10$ may have occurred because random phase was not optimal for minimizing the use of temporal information; with random phase and many harmonics there is a reasonable probability of a waveform with a high peak factor at some point along the basilar membrane.

For our data, the absence of a phase effect for low harmonic numbers ($N=6$ or 7) is as expected and is consistent with the assumption that harmonics up to about the fifth to eighth are resolved in the auditory system. However, F0DLs were smaller for cosine than for alternating phase once the number of the lowest harmonic was 8 or more, which suggests that the components were not resolved when N was 8 or more. The relatively small F0DLs for $N=8-11$, combined with the phase effect, and combined with earlier data showing a "pitch shift" effect for harmonic numbers in this range (Moore and Moore, 2003), lead to the interpretation that temporal fine structure plays a role in F0 discrimination for $N=8-11$.

One concern with these results is whether the subjects based their judgments on spectral differences between stimuli, rather than on the differences in F0. The randomization of N by ±1 was intended to reduce the usefulness of spectral cues. However, when $\Delta\mathrm{F0}$ was large, spectral cues might still have been used. Consider the F0DLs that would be expected if subjects performed the task by comparing the frequency of the lowest harmonic in the tones across the two intervals in a forced-choice trial. As a limiting case, assume that this comparison could be made perfectly. Since harmonic $N+1$ was centered at 2000 Hz, the mean frequency separation between adjacent harmonics, $\mathrm{F0}_c$, was equal to $2000/(N+1)$. For $\Delta\mathrm{F0} < \mathrm{F0}_c$, a strategy of comparing the frequency of the lowest harmonic in the two tones within a trial would lead to only 6/9 (66.7%) correct,[1] which is well below the "target" of 79.4% correct tracked by the psychophysical procedure. However, for $\Delta\mathrm{F0} = \mathrm{F0}_c$, the strategy would lead to 7/9 (77.8%) correct, and if $\Delta\mathrm{F0}$ were a little greater than $\mathrm{F0}_c$, the strategy would lead to 8/9 (88.9%) correct, which is greater than the "target" of 79.4% correct. The dashed lines in Fig. 2 are defined by $\Delta\mathrm{F0} = \mathrm{F0}_c$. If subjects based their decisions on the frequency of the lowest harmonic, and were perfectly able to discriminate the frequency of that harmonic, the measured F0DLs should lie slightly above the dashed line. Imperfect discrimination of the frequency of the lowest harmonic would lead to F0DLs further above the line.

For the cosine-phase stimuli, the F0DLs fall below the dashed line for all values of N for all subjects, indicating that subjects did not base their decisions on the frequency of the lowest harmonic. For the alternating-phase stimuli, the F0DLs fall well below the dashed line for low values of N, but for higher values of N the F0DLs approach and in one case (S2) lie above the dashed line. For alternating-phase stimuli with high N, subjects may have partly based decisions on the frequency of the lowest harmonic, possibly because the low pitch was very weak.

III. EXPERIMENT 2: CENTER FREQUENCY 5000 HZ

A. Rationale

The F0DLs measured in experiment 1 were smaller for cosine than for alternating phase even when N was high (14 or 15) and temporal fine structure information was probably not used (Hall *et al.*, 2003; Moore and Moore, 2003). Presumably, the phase effect for these high harmonics reflects the influence of phase on the envelopes of the waveforms on the basilar membrane. We have argued that the phase effect for N in the range 8–11 probably reflects the use of temporal fine structure cues. This argument is based on two facts: (1) performance was relatively good for $N=8-11$ and (2) pitch shift effects occur for harmonics in this range (Schouten *et al.*, 1962; Moore and Moore, 2003). However, it is difficult to be sure that temporal fine structure was involved. As a further way of assessing the role of temporal fine structure, as opposed to envelope cues, experiment 1 was repeated, but with the nominal center frequency shifted to 5000 Hz. At this frequency, phase locking is very weak (Palmer and Russell, 1986) and temporal fine structure information would probably be unusable.

B. Procedure, stimuli, and subjects

The procedure and stimuli were the same as for experiment 1, except that the nominal center frequency was 5000 Hz. The background threshold-equalizing noise used to mask combination tones had an upper cutoff frequency of 8000 Hz.

Five new subjects (four males and one female) were recruited for this experiment, all of whom were music students. Their ages ranged from 20 to 23 years. All subjects had absolute thresholds better than 20 dB over the range of audiometric frequencies from 250 to 8000 Hz. All subjects except author JA were paid for their services. Training was conducted in a similar way to that described for experiment 1.

C. Results and discussion

Figure 3 shows the results for each subject and the geometric mean across subjects (bottom-right panel). Results for the cosine-phase stimuli are shown by open circles and results for the alternating-phase stimuli are shown by filled circles. Note that the range of values on the ordinate is different for Figs. 2 and 3, reflecting the generally higher F0DLs when the center frequency was 5000 Hz. However, to facilitate comparison, the ratio of maximum to minimum values on the ordinate is the same for Figs. 2 and 3. The results of experiment 2 are somewhat more erratic than for experiment 1, although the mean results are fairly orderly. The pattern of results is very different from that found in experiment 1. It is also very different from that found by Bernstein and Oxenham (2005) for the "high" spectral region, probably because the high region in their experiment included the frequency range down to 3280 Hz, a frequency at which phase locking still occurs. For our data, F0DLs tended to decrease with increasing harmonic number, rather than to increase over the range $N=8-12$. The phase effect was smaller and less consistent than for experiment 1 and than found by Bernstein and Oxenham (2005) for the "high" spectral region. In the mean data for experiment 2 there is no phase effect for $N=6$ or 7, and a small but reasonably consistent phase effect for $N=9$ and above.

A within-subjects ANOVA was conducted on the logarithms of the F0DLs with factors lowest harmonic number ($N=6$ to 15) and phase (cosine or alternating). Both main effects were significant: for harmonic number, $F(9,36)=6.66$, $p<0.001$; for phase, $F(1,4)=29.82$, $p<0.01$. The interaction of harmonic number and phase was not significant: $F(9,36)=1.11$, $p>0.05$. *Posthoc* tests of the effect of phase, using the least-significant-differences (LSD), showed no significant effect for $N=6$, 7, 8, 12, 13, or 14. The effect of phase was significant at $p<0.05$ for $N=9$, 10, 11, and 15.

The dashed lines in Fig. 3 show the limit of performance that would be expected if subjects based decisions on the frequency of the lowest harmonic. For the cosine-phase stimuli, the F0DLs mostly fall below the dashed lines, indicating that performance was not based entirely on the frequency of the lowest harmonic; presumably information derived from the temporal envelope was used. However, for the alternating-phase stimuli and for $N>9$, the F0DLs often fell

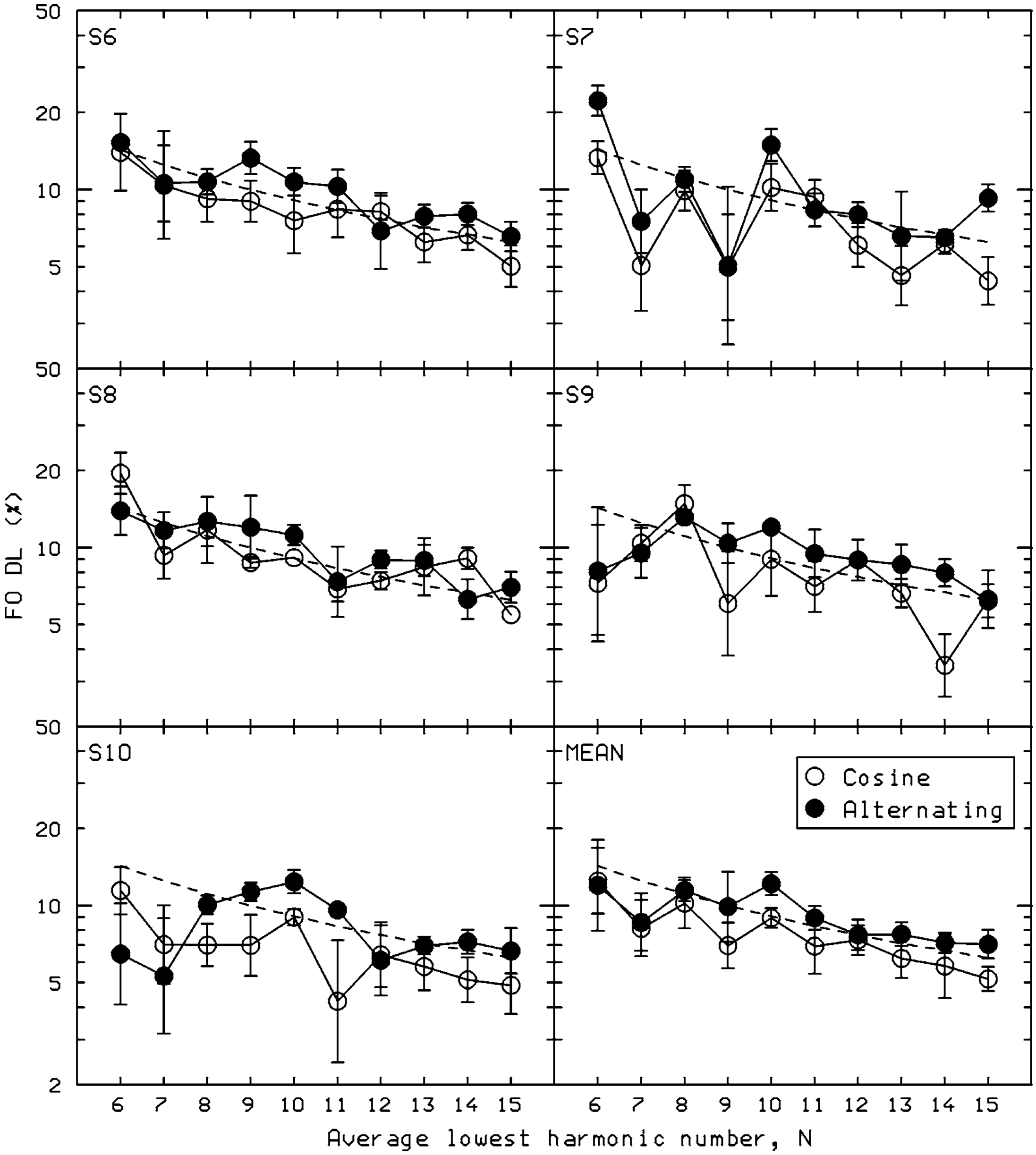

FIG. 3. As Fig. 2, but showing the results for experiment 2 (harmonics centered at 5000 Hz).

close to or a little above the dashed line, indicating that performance may have been based partly or exclusively on the frequency of the lowest harmonic. Possibly, the phase effect for high *N* would have been larger if this spectral cue had not been available.

The absence of a phase effect for $N=6$ or 7 suggests that the components in the complex tones were resolved for these values of *N*. Despite this, F0 discrimination was very poor, F0DLs being around 10%. This indicates that the presence of resolved harmonics is not sufficient to give good F0 discrimination. Rather, it appears that temporal information (phase locking) is required. This is consistent with previous work suggesting that the coding of the frequencies of individual components within complex tones (with harmonics below 5000 Hz) depends at least partly on temporal information (Moore *et al.*, 1984; Hartmann and Doty, 1996; de Cheveigné, 1999).

For $N=9$ or 10, there was a clear phase effect in experiment 2, but performance was much poorer than for experiment 1. This is consistent with our argument that the relatively good performance found in experiment 1 for intermediate harmonic numbers reflects the use of temporal fine structure information. When such information is not available (experiment 2), performance is much poorer. In experiment 2, the decrease in F0DLs with increasing *N* over the range 8–15 may reflect three effects. First, it may reflect the use of spectral cues, especially for the alternating-phase stimuli, as described above. Second, when spectral cues were not used, it may be connected with the salience of the distracting spectral changes produced by the randomization of *N*. As the mean value of *N* increased, the change in frequency of the lowest component produced by randomizing *N* by ±1 became smaller (for example a change by +1 for $N=8$ corresponds to a change in frequency of 555.6 Hz, whereas a change by +1 for $N=15$ corresponds to a change in frequency of 312.5 Hz). Hence, the distracting effect would have decreased with increasing N. The third effect is connected with the influence of auditory filtering. For the higher values of N, all harmonics would be well within the passband of the auditory filter centered on the harmonics, and so the waveform at the output of the filter (equivalent to a specific place on the basilar membrane) would be minimally affected by the amplitude and phase response of the filter. For the lower values of N, the outer harmonics would lie closer to the edge of the passband of the filter, and the nonlinear phase response of the filter (Lentz and Leek, 2001; Oxenham and Dau, 2001) might lead to a reduction of the peak factor of the waveform at the output of the filter, at least for the cosine-phase stimuli. The reduced peak factor would impair the ability to extract the periodicity of the temporal envelope.

When the frequencies of the harmonics of a complex tone fall within the range where phase-locking information is available (as in experiment 1), F0DLs may be relatively small when low harmonics are present because the temporal information is relatively independent in different frequency channels. Each component produces maximum excitation at a different place on the basilar membrane, and the temporal information derived from each place is different from that at the other places. This helps to resolve ambiguities in the temporal information and leads to a clear estimate of F0 (Moore, 2003). In contrast, when only a few intermediate or high harmonics are present, the information carried in the interspike intervals is ambiguous. The waveform on the basilar membrane is not greatly different from the stimulus waveform, and there are several peaks in the temporal fine structure close to each envelope peak. Each of the peaks in the temporal fine structure has the potential to evoke nerve spikes. For such a stimulus, many different interspike intervals are present, only one of which corresponds to F0 (Javel, 1980; Cariani and Delgutte, 1996a, b). This leads to an ambiguous and unclear pitch. The ambiguity can be decreased by increasing the number of harmonics, which leads to distinctly different waveforms at different points along the basilar membrane and results in a clearer pitch (Houtsma and Smurzynski, 1990; Kaernbach and Bering, 2001).

The finding of a phase effect for $N=8$ and above when the center frequency was 2000 Hz appears to be inconsistent with the conclusion of Gibson (1970) and Bernstein and Oxenham (2003) that harmonics up to about the 10th or 11th are resolved for complex tones with F0=100 or 200 Hz, when those tones contain equal-amplitude successive harmonics. However, it is possible that harmonics with numbers in the range 8–11 are partially resolved, but still interact sufficiently to convey information about F0 in the temporal fine structure of the waveform on the basilar membrane. Thus, our finding of an effect of phase on F0 discrimination for $N=8-12$ does not necessarily rule out the possibility that harmonics in that range are resolved to a small extent. However, it is worth noting that, for the center frequency of 2000 Hz, the effects of phase on the F0DLs were at least as large for $N=8$ and 9 as for higher values of N. This suggests that the components were already interacting strongly for $N=8$ and 9. It is worth considering, therefore, whether the task used by Bernstein and Oxenham (2003) provides a good measure of harmonic resolvability.

Bernstein and Oxenham used a task similar to that of Moore and Ohgushi (1993) in which subjects heard a pure tone followed by a complex tone and were asked to indicate whether the pure tone was higher or lower in frequency than a comparison tone contained within the complex. However, unlike the task used Moore and Ohgushi, Bernstein and Oxenham pulsed the comparison tone on and off, while the other components in the complex tone were presented as a single longer burst. The pulsing of the comparison tone was intended to promote the segregation of the comparison tone from the other components in the complex. However, it may have had other effects, for example producing a release from adaptation to the comparison tone [Viemeister and Bacon (1982); note, however, that Bernstein and Oxenham argued that release from adaptation should not affect peripheral resolvability]. In any case, pulsing a component within a complex tone on and off will produce a change in the basilar-membrane waveform at the place tuned to the pulsed harmonic, and this may lead to the percept of a change in that frequency region. This may have been sufficient to allow subjects to perform the task of Bernstein and Oxenham even when the pulsed harmonic was not actually resolved.

We have argued that, for N in the range 8–11, subjects are able to use information related to the temporal fine structure of the stimuli, whereas for very high N (14 or 15) only envelope information or spectral cues are available. A possible explanation for the loss of temporal fine structure information with increasing N is related to the precision with which interspike intervals in a given neuron can be measured by the auditory system. It seems reasonable to assume that this precision follows Weber's law, i.e., a given interspike interval, t, can be measured with a standard deviation, $\sigma(t)$, which is proportional to t. Consider the temporal pattern of neural spikes evoked by a complex tone with F0=200 Hz and with three harmonics centered at 2000 Hz. There will be many interspike intervals close to 5 ms (the true period), and there will be other intervals corresponding to 5 ± 0.5 ms, e.g., 4, 4.5, 5.5, and 6 ms (Javel, 1980; Cariani and Delgutte, 1996a, b). Provided that $\sigma(t)$ is less than about $0.05t$, it will be possible for the auditory system to extract the interval, t, corresponding to 5 ms. Assume now that the F0 is decreased to 125 Hz, keeping the center frequency the same. Now the true period is 8 ms, and there will be competing intervals such as 7, 7.5, 8.5, and 9 ms. If $\sigma(t)$ is only a little less than $0.05t$, the standard deviation $\sigma(t)$ associated with the estimation of the interval of 8 ms will exceed 0.25 ms. In this case, the auditory system will no longer reliably be able to distinguish the interval of 8 ms from the nearby intervals of 7.5 and 8.5 ms; temporal fine structure information will be lost and only envelope information will remain.

IV. CONCLUSIONS

(1) For complex tones whose lowest harmonic number, N, was 6 or 7, F0DLs were independent of component phase, consistent with the idea that the harmonics for

such tones are resolved in the auditory system. F0DLs for such tones were small (around 0.8%) when the harmonics were centered at 2000 Hz (experiment 1) but were relatively large (around 10%) when the harmonics were centered at 5000 Hz. This suggests that the presence of resolved harmonics is not sufficient to give good F0 discrimination. Rather, temporal information (phase locking) is required.

(2) For complex tones whose lowest harmonic number, N, was in the range 8–11, F0DLs were smaller when the components were added in cosine phase than when they were added in alternating phase. This is consistent with the idea that the harmonics were not resolved and that F0DLs depended on temporal information. When the harmonics were centered at 2000 Hz, F0DLs increased as N was increased from 8 to 12. This is consistent with a progressive loss of temporal fine structure information over this range. When the harmonics were centered at 5000 Hz, F0DLs decreased slightly as N was increased from 8 to 12. This is consistent with the absence of temporal fine structure information at this center frequency, forcing subjects to rely solely on envelope cues or spectral cues.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council (UK) and by the Grindley Fund. We thank Bob Carlyon and Hedwig Gockel for helpful discussions and Tom Baer, Hedwig Gockel, John Grose, Christian Kaernbach, and an anonymous reviewer for helpful comments on an earlier version of this paper.

---

[1] This footnote derives the proportion correct that would be achieved for different values of $\Delta F0$, assuming that judgments were based on the change in frequency of the lowest harmonic, and that discrimination of the frequency of that harmonic was perfect. In any given trial, there were three possible values of N for each stimulus (randomly drawn from $N_{\mathrm{nom}}-1$, N_{nom}, and $N_{\mathrm{nom}}+1$, where N_{nom} is the nominal value of N), and nine possible combinations of N for the two stimuli in a trial. For three of those combinations, N was the same for the two stimuli in a trial, so subjects would score perfectly. For another three combinations, the frequency shift produced by the change in N would be in the same direction as the frequency shift produced by the change in F0, so subjects would again score perfectly. We consider next the outcomes for the remaining three combinations, for three ranges of $\Delta F0$.

(1) $\Delta F0 < F0_c$: subjects would be consistently wrong. Hence, the overall proportion correct is $(3+3)/9=6/9$.

(2) $\Delta F0 = F0_c$: for two of the combinations, the frequency shift produced by the change in N would exactly cancel the frequency shift produced by the change in F0, and subjects would guess and score 50% correct on average. For the final combination (corresponding to a difference in N of 2 across stimuli), subjects would be consistently wrong. Hence, the overall proportion correct is $(3+3+(0.5\times 2))/9=7/9$.

(3) $\Delta F0 > F0_c$ but $< 2F0_c$: for two of the combinations, the frequency shift produced by the change in N would be smaller than the frequency shift produced by the change in F0, and subjects would score perfectly. For the final combination (corresponding to a difference in N of 2 across stimuli), subjects would be consistently wrong. Hence, the overall proportion correct is $(3+3+2)/9=8/9$.

Alcántara, J. I., Moore, B. C. J., Glasberg, B. R., Wilkinson, A. J. K., and Jorasz, U. (**2003**). "Phase effects in masking: Within-versus across-channel processes," J. Acoust. Soc. Am. **114**, 2158–2166.

Bernstein, J. G., and Oxenham, A. J. (**2003**). "Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number?" J. Acoust. Soc. Am. **113**, 3323–3334.

Bernstein, J. G., and Oxenham, A. J. (**2005**). "An autocorrelation model with place dependence to account for the effect of harmonic number on fundamental frequency discrimination," J. Acoust. Soc. Am. **117**, 3816–3831.

Cariani, P. A., and Delgutte, B. (**1996a**). "Neural correlates of the pitch of complex tones. I. Pitch and pitch salience," J. Neurophysiol. **76**, 1698–1716.

Cariani, P. A., and Delgutte, B. (**1996b**). "Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch and the dominance region for pitch," J. Neurophysiol. **76**, 1717–1734.

Carlyon, R. P., and Shackleton, T. M. (**1994**). "Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms?," J. Acoust. Soc. Am. **95**, 3541–3554.

de Boer, E. (1956). "On the "residue" in hearing," Ph.D. thesis, University of Amsterdam.

de Cheveigné, A. (**1999**). "Pitch shifts of mistuned partials: A time-domain model," J. Acoust. Soc. Am. **106**, 887–897.

Gibson, L. (1970). "The ear as an analyzer of musical tones," 80th Meeting of the Acoustical Society of America p. 63.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103–138.

Goldstein, J. L. (**1973**). "An optimum processor theory for the central formation of the pitch of complex tones," J. Acoust. Soc. Am. **54**, 1496–1516.

Hall, J. W., Buss, E., and Grose, J. H. (**2003**). "Modulation rate discrimination for unresolved components: temporal cues related to fine structure and envelope," J. Acoust. Soc. Am. **113**, 986–993.

Hartmann, W. M., and Doty, S. L. (**1996**). "On the pitches of the components of a complex tone," J. Acoust. Soc. Am. **99**, 567–578.

Hoekstra, A., and Ritsma, R. J. (**1977**). "Perceptive hearing loss and frequency selectivity," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London).

Houtsma, A. J. M., and Smurzynski, J. (**1990**). "Pitch identification and discrimination for complex tones with many harmonics," J. Acoust. Soc. Am. **87**, 304–310.

Javel, E. (**1980**). "Coding of AM tones in the chinchilla auditory nerve: Implications for the pitch of complex tones," J. Acoust. Soc. Am. **68**, 133–146.

Kaernbach, C., and Bering, C. (**2001**). "Exploring the temporal mechanism involved in the pitch of unresolved harmonics," J. Acoust. Soc. Am. **110**, 1039–1048.

Lentz, J. J., and Leek, M. R. (**2001**). "Psychophysical estimates of cochlear phase response: masking by harmonic complexes," J. Assoc. Res. Otolaryngol. **2**, 408–422.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Licklider, J. C. R. (**1956**). "Auditory frequency analysis," in *Information Theory*, edited by C. Cherry (Academic, New York).

Meddis, R., and Hewitt, M. (**1991**). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification," J. Acoust. Soc. Am. **89**, 2866–2882.

Meddis, R., and O'Mard, L. (**1997**). "A unitary model of pitch perception," J. Acoust. Soc. Am. **102**, 1811–1820.

Moore, B. C. J. (**1973**). "Frequency difference limens for short-duration tones," J. Acoust. Soc. Am. **54**, 610–619.

Moore, B. C. J. (**1977**). "Effects of relative phase of the components on the pitch of three-component complex tones," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London).

Moore, B. C. J. (**1982**). *An Introduction to the Psychology of Hearing*, *2nd Ed.* (Academic, London).

Moore, B. C. J. (**2003**). *An Introduction to the Psychology of Hearing*, *5th Ed.* (Academic, San Diego).

Moore, B. C. J., and Glasberg, B. R. (**1988**). "Effects of the relative phase of the components on the pitch discrimination of complex tones by subjects with unilateral and bilateral cochlear impairments," in *Basic Issues in Hearing*, edited by H. Duifhuis, H. Wit, and J. Horst (Academic, London).

Moore, B. C. J., and Ohgushi, K. (**1993**). "Audibility of partials in inharmonic complex tones," J. Acoust. Soc. Am. **93**, 452–461.

Moore, B. C. J., and Rosen, S. M. (**1979**). "Tune recognition with reduced pitch and interval information," Q. J. Exp. Psychol. **31**, 229–240.

Moore, B. C. J., Glasberg, B. R., and Shailer, M. J. (**1984**). "Frequency and intensity difference limens for harmonics within complex tones," J. Acoust. Soc. Am. **75**, 550–561.

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcántara, J. I. (**2000**). "A test for the diagnosis of dead regions in the cochlea," Br. J. Audiol. **34**, 205–224.

Moore, G. A., and Moore, B. C. J. (**2003**). "Perception of the low pitch of frequency-shifted complexes," J. Acoust. Soc. Am. **113**, 977–985.

Oxenham, A. J., and Dau, T. (**2001**). "Towards a measure of auditory-filter phase response," J. Acoust. Soc. Am. **110**, 3169–3178.

Palmer, A. R., and Russell, I. J. (**1986**). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner haircells," Hear. Res. **24**, 1–15.

Patterson, R. D. (**1973**). "The effects of relative phase and the number of components on residue pitch," J. Acoust. Soc. Am. **53**, 1565–1572.

Patterson, R. D. (**1987**). "A pulse ribbon model of monaural phase perception," J. Acoust. Soc. Am. **82**, 1560–1586.

Plomp, R. (**1964**). "The ear as a frequency analyzer," J. Acoust. Soc. Am. **36**, 1628–1636.

Plomp, R., and Mimpen, A. M. (**1968**). "The ear as a frequency analyzer II," J. Acoust. Soc. Am. **43**, 764–767.

Ritsma, R. J. (**1962**). "Existence region of the tonal residue. I.," J. Acoust. Soc. Am. **34**, 1224–1229.

Ritsma, R. J. (**1963**). "Existence region of the tonal residue. II.," J. Acoust. Soc. Am. **35**, 1241–1245.

Schouten, J. F. (**1940**). "The residue and the mechanism of hearing," Proc. K. Ned. Akad. Wet. **43**, 991–999.

Schouten, J. F., Ritsma, R. J., and Cardozo, B. L. (**1962**). "Pitch of the residue," J. Acoust. Soc. Am. **34**, 1418–1424.

Schroeder, M. R. (**1970**). "Synthesis of low peak-factor signals and binary sequences with low autocorrelation," IEEE Trans. Inf. Theory **IT-16**, 85–89.

Shackleton, T. M., and Carlyon, R. P. (**1994**). "The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination," J. Acoust. Soc. Am. **95**, 3529–3540.

Shamma, S., and Klein, D. (**2000**). "The case of the missing pitch templates: how harmonic templates emerge in the early auditory system," J. Acoust. Soc. Am. **107**, 2631–2644.

Shamma, S. A. (**1985**). "Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve," J. Acoust. Soc. Am. **78**, 1622–1632.

Srulovicz, P., and Goldstein, J. L. (**1983**). "A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum," J. Acoust. Soc. Am. **73**, 1266–1276.

Terhardt, E. (**1974**). "Pitch, consonance, and harmony," J. Acoust. Soc. Am. **55**, 1061–1069.

Thurlow, W. R. (**1963**). "Perception of low auditory pitch: a multicue mediation theory," Psychol. Rev. **70**, 461–470.

van Noorden, L. P. A. S. (**1982**). "Two-channel pitch perception," in *Music, Mind and Brain*, edited by M. Clynes (Plenum, New York).

Viemeister, N. F., and Bacon, S. P. (**1982**). "Forward masking by enhanced components in harmonic complexes," J. Acoust. Soc. Am. **71**, 1502–1507.