The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination

Trevor M. Shackleton and Robert P. Carlyon
Laboratory of Experimental Psychology, University of Sussex, Brighton, E. Sussex BN1 9QG, United Kingdom

(Received 11 May 1993; revised 23 December 1993; accepted 10 January 1994)

A series of experiments investigated the influence of harmonic resolvability on the pitch of, and the discriminability of differences in fundamental frequency ($F0$) between, frequency-modulated (FM) harmonic complexes. Both $F0$ (62.5 to 250 Hz) and spectral region (LOW: 125–625 Hz, MID: 1375–1875 Hz, and HIGH: 3900–5400 Hz) were varied orthogonally. The harmonics that comprised each complex could be summed in either sine (0°) phase (SINE) or alternating sine-cosine (0°–90°) phase (ALT). Stimuli were presented in a continuous pink-noise background. Pitch-matching experiments revealed that the pitch of ALT-phase stimuli, relative to SINE-phase stimuli, was increased by an octave in the HIGH region, for all $F0$'s, but was the same as that of SINE-phase stimuli when presented in the LOW region. In the MID region, the pitch of ALT-phase relative to SINE-phase stimuli depended on $F0$, being an octave higher at low $F0$'s, equal at high $F0$'s, and unclear at intermediate $F0$'s. The same stimuli were then used in three measures of discriminability: FM detection thresholds (FMTs), frequency difference limens (FDLs), and FM direction discrimination thresholds (FMDDTs, defined as the minimum FM depth necessary for listeners to discriminate between two complexes modulated 180° out of phase with each other). For all three measures, at all $F0$'s, thresholds were low ($<4\%$ for FMTs, $<5\%$ for FMDDTs, and $<1.5\%$ for FDLs) when stimuli were presented in the LOW region, and high ($>10\%$ for FMTs, $>7\%$ for FMDDTs, and $>2.5\%$ for FDLs) when presented in the HIGH region. When stimuli were presented in the MID region, thresholds were low for low $F0$'s, and high for high $F0$'s. Performance was not markedly affected by the phase relationship between the components of a complex, except for stimuli with intermediate $F0$'s in the MID spectral region, where FDLs and FMDDTs were much higher for ALT-phase stimuli than for SINE-phase stimuli, consistent with their unclear pitch. This difference was much smaller when FMTs were measured. The interaction between $F0$ and spectral region for both sets of experiments can be accounted for by a single definition of resolvability.

PACS numbers: 43.66.Hg, 43.66.Nm, 43.66.Ba [HSC]

INTRODUCTION

Many of the periodic sounds that we hear in everyday life contain both low-frequency harmonics, which are resolved by the peripheral auditory system, and high-frequency harmonics, which are not. The cues available for pitch perception differ between these two types of harmonic. For resolved harmonics, although the excitation pattern has pronounced peaks at each harmonic frequency, no single auditory filter has unambiguous information about the fundamental frequency ($F0$) of the stimulus, so some form of across-channel combination of information is required (Goldstein, 1973; Meddis and Hewitt, 1991a,b; Moore, 1989; Patterson, 1987; Piszczalski and Galler, 1979; Srulovicz and Goldstein, 1983; Terhardt, 1974; Terhardt *et al.*, 1982). In contrast, when several unresolved harmonics interact within a single auditory filter, the output of that filter repeats at a rate equal to $F0$, and so pitch can be directly determined from within-channel cues, even though the excitation pattern does not contain distinct peaks (Assmann and Summerfield, 1990; Licklider, 1951; Meddis and Hewitt, 1991a,b; Patterson, 1987; Schouten, 1940, 1970; Slaney and Lyon, 1990). On the basis of these differences it might seem that two different mechanisms would be required to extract the pitch of resolved and unresolved harmonics. However, models using a single mechanism have been shown to be effective in extracting pitch from both kinds of harmonic (Houtsma and Smurzynski, 1990; Meddis and Hewitt, 1991a; Patterson, 1987). This paper forms part of a project investigating whether a single, or double, mechanism is required for pitch perception. A companion paper (Carlyon and Shackleton, 1994) examines this problem directly by requiring listeners to make a simultaneous comparison of fundamental frequency ("$F0$") between two different spectral regions, each of which contains groups of harmonics which are either both resolved, both unresolved, or where the two differ in resolvability. It is argued that the existence of two pitch mechanisms would cause the discriminable $F0$ difference to be larger for the condition where the groups differ in resolvability. In the current paper, we examine the problem a little less directly through the study of pitch perception, and by obtaining three different measures of sequential $F0$ discrimination.

 0001-4966/94/95(6)/3529/12/$6.00

One way of studying pitch mechanisms is to perform discrimination experiments to measure the accuracy with which a harmonic, or group of harmonics, is encoded (e.g., Cullen *et al.*, 1986; Hoekstra, 1979; Houtsma and Smurzynski, 1990; Moore and Glasberg, 1988; Moore *et al.*, 1984). Thresholds for detecting changes in $F0$ when the stimulus comprises only low harmonics are much lower than when the stimulus comprises only high harmonics (Cullen *et al.*, 1986; Hoekstra, 1979; Houtsma and Smurzynski, 1990). With the exception of Hoekstra's (1979) study, however, increasing harmonic number (and hence decreasing resolvability) was also confounded with increasing harmonic frequency, so it is not possible to determine whether the increase in threshold was due to reducing resolvability or to increasing spectral frequency. Hoekstra's (1979) study orthogonally varied $F0$ and the spectral region in which the harmonics were presented, and found that harmonic number, and hence resolvability, was the critical variable. In addition to these studies, Moore and colleagues (Moore and Glasberg, 1986; Moore *et al.*, 1985a,b; Moore and Ohgushi, 1993) found that the ability to hear a harmonic out from a complex, or simply detect that it was mistuned, was better for resolved than for unresolved harmonics, although absolute frequency of the harmonic was also found to be important (cf. Hartmann *et al.*, 1990).

Another approach has been to perform pitch matching experiments for stimuli where the envelope repetition rate differs from $F0$, and for which different putative mechanisms would be expected to produce different pitches (Flanagan and Guttman, 1960a,b; Guttman and Flanagan, 1964; Lundeen and Small, 1984; Mathes and Miller, 1947; Moore, 1977; Ritsma and Engel, 1964; Rosenberg, 1965; Thurlow and Small, 1955; Warren and Wrightson, 1981). These experiments showed that, for low $F0$'s, the pitch reported corresponded to the envelope periodicity, whereas the pitch reported for high $F0$'s corresponded to the $F0$. However, the difficulty in interpreting these results is that they mainly use wideband stimuli, and so any analysis based upon the hypothesized resolution of components also requires an estimate of which spectral region dominates the pitch extraction process.

In this paper we obtain measures both of pitch perception, through pitch-matching experiments, and of $F0$ encoding accuracy, through three different discrimination paradigms. In all experiments we orthogonally vary $F0$ and the spectral region in which harmonics are presented, so that harmonic resolution and spectral frequency effects can be uncoupled. Measurements are made using both sine- and alternating-phase stimuli, so that mechanisms sensitive to temporal structure can be studied. Frequency-modulated ("FM") stimuli are used in the present study for two reasons. First, a companion study (Carlyon and Shackleton, 1994) required measures of the encoding accuracy of FM stimuli. Second, as it is not obvious what the best measure of encoding accuracy is for these stimuli, the opportunity is taken to compare three different measures.

I. GENERAL METHOD

All stimuli were harmonic series that were frequency modulated (before filtering) at a rate of 2.5 Hz, were of 400-ms duration (1 cycle of FM), and were gated on and off with 5-ms raised-cosine ramps. They were bandpass filtered using a pair of cascaded Kemo VBF25.03 filters (one high pass, one low pass, 48 dB/oct each), attenuated (Expt. 1: Tucker-Davis Technologies PA3, other expts: Wilsonics PATT), and fed into one input of a summing headphone amplifier. The levels of all components with frequencies in the filter passbands were 50 dB SPL. A 10-kHz-wide pink noise was presented continuously; its spectrum level in dB SPL was 22.8 at 500 Hz, 20.2 at 1000 Hz, 17.2 at 2000 Hz, and 13.8 at 4000 Hz. All stimuli were presented through the right earpiece of a Sennheiser HD414SL headset, and were monitored using an HP3561A spectrum analyzer.

The stimuli, before filtering, consisted of the fundamental and consecutive harmonics of a complex tone, summed either in sine phase (SINE) or in alternating sine and cosine phase (ALT: harmonics with frequencies which were odd multiples of the fundamental were in sine phase, and even multiples were in cosine phase). Three different spectral regions were used, a LOW region, obtained by setting the filter cutoffs to 125 and 625 Hz (3 dB down points), a MID region (1375 to 1875 Hz), and a HIGH region (3900 to 5400 Hz). The number of harmonics prior to filtering depended on $F0$, but was such that all harmonics up to 8125 Hz were present. In experiment 1, the signal source was a Macintosh II computer fitted with a Digidesign Audiomedia DSP card. The stimulus was generated in real time at a sampling rate of 44.1 kHz by interpolating from a look-up table waveform comprising 60 harmonics, and was played out through a 16-bit DAC via an on-card reconstruction filter (Russell and Darwin, 1991). In subsequent experiments the signal was played out through a CED 1401 laboratory interface (12-bit DAC) at a sampling rate of 20 kHz, before being low-pass filtered at 8.6 kHz (Kemo VBF25.01, 135 dB/oct). The stimuli were generated in advance and stored on hard disk.

A total of 14 listeners took part in the different experiments. Their absolute thresholds at octave frequencies between 250 and 8000 Hz were within 15 dB of the standard (1969 ANSI). Listener TS was the first author. Listeners were tested individually in an IAC single-walled sound-attenuating booth within a large single-walled sound-attenuating room.

II. PHASE EFFECTS IN PITCH PERCEPTION

The experiments reported in this section investigated the effect on the pitch of a stimulus of playing it in alternating phase, as opposed to sine phase, as a function of $F0$ and of spectral content. Waveform-based theories (Schouten, 1940, 1970) predict that the doubling of waveform peaks which occurs when a harmonic series is put in ALT phase should increase its pitch by an octave relative to that of an otherwise-identical SINE-phase stimulus. In contrast, "pattern recognition" theories (Goldstein, 1973;

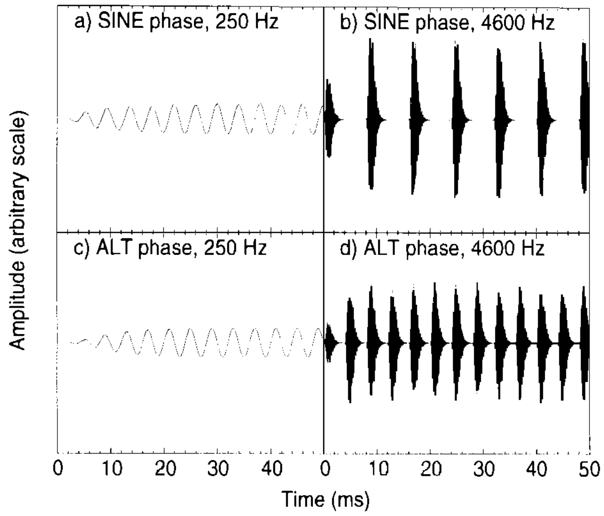

FIG. 1. Output of two simulated auditory filters (Patterson *et al.*, 1988) with center frequencies of 250 Hz (a), (c), and 4600 Hz (b), (d) to a wideband harmonic stimulus with an $F0$ of 125 Hz and all components in sine phase (a), (b), or alternating sine and cosine phase (c), (d).

Piszczalski and Galler, 1979; Terhardt, 1974; Terhardt *et al.*, 1982) state that pitch is derived from the power spectrum, and make predictions that are independent of phase. Our prediction was that the dependence of pitch on phase would be determined by the degree to which the components of the complex were resolved by the peripheral auditory system (Moore, 1977). Figure 1(a) and (b) show a wideband SINE-phase stimulus with a $F0$ of 125 Hz filtered through two simulated auditory filters (Patterson *et al.*, 1988) centered on 250 and 4600 Hz, respectively. The output of the low-frequency filter consists of a single, resolved, harmonic, whereas that of the high-frequency filter consists of several, unresolved, harmonics. Figure 1(c) and (d) show an ALT phase stimulus filtered in the same manner. There are clear secondary peaks observable in the high-frequency filter for the ALT stimulus compared with the SINE stimulus, whereas there is little difference (apart from a constant phase shift) between the waveforms passed through the low-frequency filter for ALT and SINE stimuli. It is on this basis that we would predict pitch increases of an octave for unresolved harmonics in ALT phase, but not for resolved harmonics. Experiment 1 used pitch matching to test these predictions, whereas experiment 2 used a more efficient method of pitch preference determination to obtain more precise information about the transition region between the two modes.

A. Experiment 1. Pitch matching

1. Method

Eight untrained listeners took part in a pitch-matching experiment. They had a wide range of musical experience. The only difference musical experience appeared to make was that musicians tended to find matches more quickly, but not more precisely, than the nonmusicians. Pitch matches were obtained using the method of adjustment. Listeners were presented with either a SINE or an ALT stimulus at a fixed $F0$, followed by a SINE stimulus whose

TABLE I. Geometric mean of pitch matches expressed as a ratio of the match frequency to target frequency. The figures in brackets are the standard deviations expressed as a percentage of the mean ratio.

		Target frequency (Hz)		
Stimulus phase	Spectral region	62.5	125	250
SINE	LOW	1.00 (2.9)	1.01 (8.8)	1.00 (1.6)
	MID	1.01 (4.0)	1.00 (5.0)	0.98 (4.0)
	HIGH	1.00 (3.8)	1.00 (7.7)	0.98 (6.0)
ALT	LOW	1.02 (4.6)	1.00 (4.6)	0.99 (1.7)
	MID	1.96 (6.6)	1.37 (26.2)	1.00 (4.2)
	HIGH	1.88 (6.9)	1.94 (8.0)	1.78 (15.3)

$F0$ they could adjust from half an octave below the fixed frequency to one and a half octaves above the fixed frequency. The matching SINE stimulus was within the same spectral region as the target stimulus. They were allowed to listen to the pair of stimuli as often as they liked, and were encouraged to approach a match from both above and below before signaling that they had found one. In each block, listeners only matched a single nominal $F0$ (62.5, 125, or 250 Hz) in a single spectral region (LOW, MID, or HIGH) for both SINE and ALT stimuli. To avoid stereotyped responses, three different target frequencies close to the nominal $F0$ were used (corresponding to 0.96, 1.00, and 1.04 times the nominal $F0$), and the initial $F0$ of the matching stimulus was randomly varied. Both the test and the matching stimuli were 400 ms long and were frequency modulated at a rate of 2.5 Hz and depth of 5%. The modulation was always started at its positive zero crossing. Three repeats of each $F0$ were presented per block, and two blocks were completed for each combination of $F0$ and spectral region, providing a total of 18 estimates of the pitch of a nominal $F0$ per listener. A block was completed in between 10 and 30 min.

2. Results

Table I shows the geometric average of the ratio of matched frequency to target frequency, averaged across listeners, with the standard deviation expressed as a percentage in brackets. It can be seen that all the matches to SINE-phase stimuli are clustered around $F0$ (ratio = 1), irrespective of the $F0$ or spectral region, and that the standard deviation is of the order of 2%–8%. This result is hardly surprising, since in this case target and matcher would be identical stimuli if the match were perfect. The results for ALT-phase stimuli are also shown in Table I and in the form of histograms averaged across all listeners in Fig. 2. The abscissa shows the ratio of matched frequency to target frequency. Histogram bin widths were 1% of the match ratio. The ordinate is the percentage of the total number of possible matches (144: 18 matches each for eight listeners,) which fall within a bin. The results for ALT-phase stimuli show an interaction between $F0$ and spectral region. In the LOW spectral region [upper row of panels: (a), (b), (c)] all matches are closely clustered around $F0$, and have a similar variance. In the HIGH spectral region [lower row of panels: (g), (h), (i)], virtu-

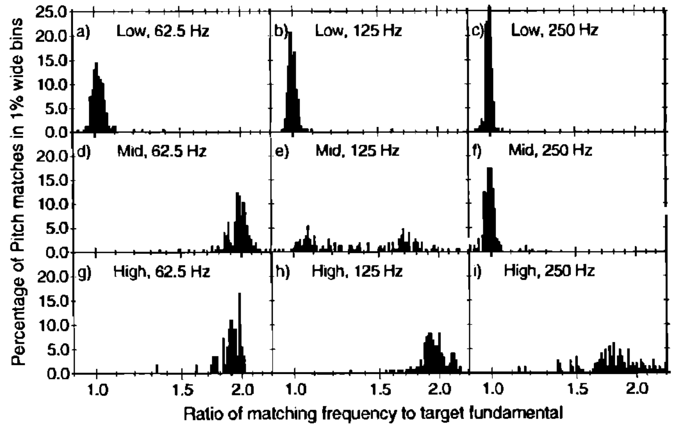

FIG. 2. Percentage of pitch matches to ALT-phase stimulus within 1% of the frequency ratio between match and target frequencies indicated on the abscissa. The fundamental frequency and spectral region of the target stimulus are shown within the panel.

ally all matches are closer to $2F0$ than $F0$, and only for an $F0$ of 250 Hz is the variance of matches much larger than that of SINE-phase matches to $F0$. A more complicated picture is shown by matches in the MID region [middle row of panels: (d), (e), (f)]. With an $F0$ of 62.5 Hz [Fig. 2(d)] the matches are clustered around $2F0$, whereas at an $F0$ of 250 Hz [Fig. 2(f)] the matches are clustered around $F0$. At an $F0$ of 125 Hz [Fig. 2(e)] matches are clustered near both $F0$ and $2F0$. This last condition was the only one where there were significant differences between listeners, three of whom produced matches which were exclusively near $F0$, one who mostly produced matches near $2F0$, and four who produced bimodal matches similar to the average distribution shown in Fig. 2(e). The detailed discussion of these results in terms of resolvability and mechanism will follow the next experiment. For the moment it is sufficient to say that the interaction between $F0$ and spectral region is consistent with resolvability being the primary variable, with harmonics of stimuli in the LOW region being resolved, those in the HIGH region being unresolved, and those in the MID region demonstrating a transition from resolved to unresolved as the $F0$ is decreased.

B. Experiment 2. Effect of phase on pitch preference

1. Rationale

Experiment 1 showed that pitch matches were grouped around $F0$ and $2F0$, and that the transition between these two matches was linked to harmonic resolvability. Experiment 2 examined the transition region between $F0$ and $2F0$ matching in more detail, and required listeners to state which of two SINE-phase stimuli played at $F0$ and $2F0$ had a pitch more similar to that of an ALT-phase stimulus played at $F0$. When the harmonics were unresolved it was predicted that listeners would choose the $2F0$ SINE stimulus, whereas when the harmonics were resolved it was predicted that the $F0$ SINE stimulus would be preferred. The new method had two advantages over that used in experiment 1. First, it was faster because we needed information only on which of two possible pitches was more similar to the reference, rather than requiring listeners to spend time making a more precise adjustment. This allowed us to obtain fine-grained information on the transition region, by studying a large number of $F0$'s, fairly quickly. Second, unlike the case where a listener matches the pitch of a complex to twice its fundamental, a pitch preference cannot be attributed to an "octave error," but demonstrates conclusively that the dominant pitch was near the octave.

2. Method

Each trial consisted of three stimuli. The first was always in ALT phase played at $F0$, the second and third were SINE-phase stimuli played at $F0$ and $2F0$ in a randomized order. Listeners were required to judge which of the second and third stimuli had the pitch more nearly equal to the first stimulus. Listeners were allowed to indicate that they were unable to make a decision. The FM imposed on the harmonic complexes had a rate of 2.5 Hz and depth of 5% and always began at a positive zero crossing of the modulation. Stimuli were computed beforehand, and consisted of the fundamental and all harmonics up to 8125 Hz, irrespective of the $F0$. The stimuli were then filtered into LOW, MID, and HIGH bands as described in Sec. I. Within a block, which lasted about 10 min, only a single spectral region was tested, but nine $F0$s (62.5, 74.3, 88.4, 105.1, 125.0, 148.7, 176.8, 210.2, 250.0 Hz) were randomly presented, ten times each. The order of presentation of spectral regions was randomly determined, in sessions which lasted 1 or 2 h. Ten blocks were run for each spectral region, leading to 100 pitch preference judgements for each combination of $F0$ and spectral region. Four listeners were used, two of whom, JS and TS (the first author), participated in experiment 1.

3. Results

The difference between the percentage of presentations upon which $2F0$ and $F0$ were the preferred pitches is shown in Fig. 3. The results for different spectral regions are shown in different panels, and for different listeners as separate lines within a panel. Positive values indicate that $2F0$ was the preferred pitch, whereas negative values indicate that $F0$ was the preferred pitch. In the LOW region, $F0$ was the preferred pitch for all $F0$'s, whereas in the HIGH region $2F0$ was the preferred pitch for all $F0$'s. In the MID region, at low $F0$'s, $2F0$ was the preferred pitch, whereas at high $F0$'s, $F0$ was the preferred pitch. The range of $F0$'s where there was a transition, and an unclear pitch, varied slightly between listeners, but always included the region around 125 Hz.

C. Discussion

The results of experiments 1 and 2 complement each other. Experiment 1 shows that pitch matches are primarily unimodal [except using ALT-phase stimuli with an $F0$ of 125 Hz presented in the MID region; Fig. 2(e)] and either clustered around $F0$ or near $2F0$. Experiment 2 expands upon these findings to show that there is a smooth,

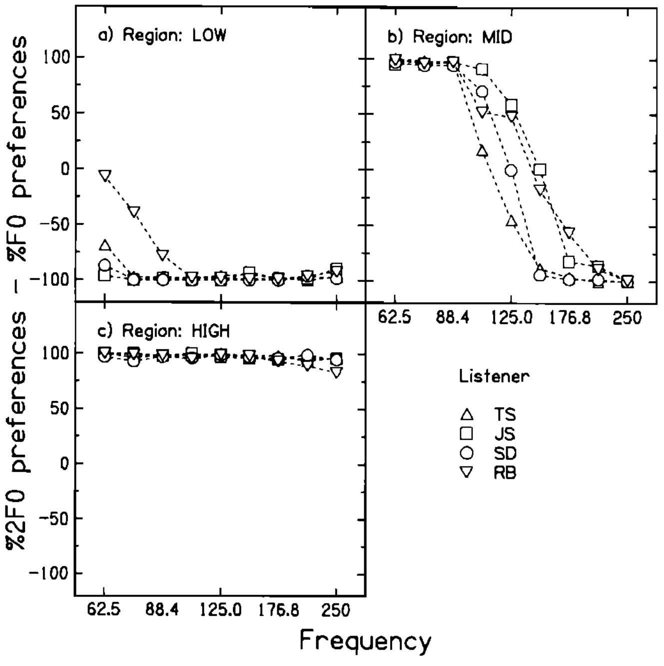

FIG. 3. Difference between the percentage of occasions on which listeners indicated which of two sine-phase stimuli of frequency $2F0$ or $F0$ sounded more similar in pitch to an alternating-phase stimulus of frequency $F0$. The different panels represent the LOW, MID, and HIGH spectral regions, and the different symbols represent individual listeners.

but sharp, transition between two pitch modes in the MID region as $F0$ is varied, and that the position of this transition varies between listeners [Fig. 3(b)]. It is likely that a similar transition would also occur in the LOW and HIGH regions if a sufficiently broad range of $F0$'s were used; there is evidence of this in the results for RB and TS which begin to turn toward preferences for $2F0$ at low $F0$'s in the LOW region [Fig. 3(a)]. The difference in transition region between different listeners could be due either to them having different auditory-filter bandwidths, or to them concentrating on different regions of the bandpass stimuli.

These results are in general agreement with those obtained by Flanagan and colleagues (Flanagan and Guttman, 1960a,b; Guttman and Flanagan, 1964; Rosenberg, 1965) who asked listeners to match the pitch of a pulse train in which all pulses had the same polarity to the pitch of a test stimulus, in which additional pulses of the opposite polarity were inserted between adjacent pulses of the original stimulus. This manipulation did not affect the $F0$ of the complex, but increased the pulse rate. At low $F0$'s (<100 Hz), listeners matched a pitch equal to the pulse rate, whereas at high $F0$'s (>200 Hz) listeners matched the $F0$ (Flanagan and Guttman, 1960a,b). In between these rates there was a region where the pitch was ambiguous. A similar result was obtained by Warren and Wrightson (1981), who used stimuli consisting of repeated segments of noise, where alternate segments could be either identical or phase inverted. They also found that, at high repetition rates (>100 Hz), phase did not affect the reported pitch, whereas at low rates (<50 Hz) it did. The region of ambiguous pitch found by Flanagan and Guttman (100 to 200 Hz) is similar to that found by us in the MID region (105 to 148 Hz), suggesting that their listeners were attending to pitch processes in a similar spectral region to that used by us in the MID region (1375 to 1875 Hz). This region is somewhat higher than the commonly accepted "dominance region" for pitch (e.g., Plomp, 1967; Ritsma and Engel, 1964), although it is close to the minimum of the absolute threshold curve (Dadson and King, 1952), consistent with the harmonics with the highest sensation level having been dominant in their task. The results of Warren and Wrightson (1981) imply a lower frequency listening region which is more in line with the "dominance region." The reasons for the differences between these studies are not obvious.

Lundeen and Small (1984) obtained pitch matches using a sinusoidal matcher and a harmonic target comprising the odd harmonics of an $F0$ of either 50 or 100 Hz. The harmonics were either all in phase (0°), in alternating phase (±45°), or in random phase. Their stimuli were all relatively broadband, and resolvability was not considered in the choice of stimulus conditions. They obtained matches near to $F0$, to $2F0$ (the harmonic separation), and to $4F0$ (the pitch predicted from temporal processing of the waveform in alternating phase) with about equal probability, independent of $F0$ or phase condition. These results do not show any clear similarity to ours, or to those of previous experiments (Flanagan and Guttman, 1960a,b; Guttman and Flanagan, 1964; Rosenberg, 1965; Warren and Wrightson, 1981).

Ritsma and Engel (1964) found, using QFM stimuli, that the pitch match to the octave broke into multiple (two or three) matches. This is a similar effect to the multiple pitches obtained using amplitude modulated stimuli (e.g., Schouten *et al.*, 1962). These pitches can be explained using temporal theory (e.g., Meddis and Hewitt, 1991a) and are due to calculating the time intervals between peaks in the fine structure of the waveform which are not exactly equal to the envelope period (Ritsma and Engel, 1964). There is tentative evidence for a similar effect in our data: For example, Fig. 2(g), which shows the distribution of matches to an ALT-phase stimulus with a 62.5-Hz $F0$ exhibits several peaks below the octave, and similar multiple peaks can be seen in other panels. However, we would not like to read too much into these data, since they are based on relatively low numbers of matches per histogram bin, so the troughs between peaks might be due to random fluctuations.

It is particularly apparent in Fig. 2(e) and (i), and to a lesser extent in Fig. 2(g) and (h) that matches near the octave tend to be flat. Such systematic deviations have been observed both for pure and, less strongly, for complex tones, but the deviation is in the opposite direction, with the higher-pitched tone being matched slightly sharp of the octave (e.g., Sundberg and Lindqvist, 1973). It is difficult to adapt the theories which have been suggested to account for this "octave stretch" phenomenon (Ohgushi, 1983; Terhardt, 1974) to also account for our own data.[1] One possible explanation for the deviations observed here is that they are an artifact of the method we used. The range

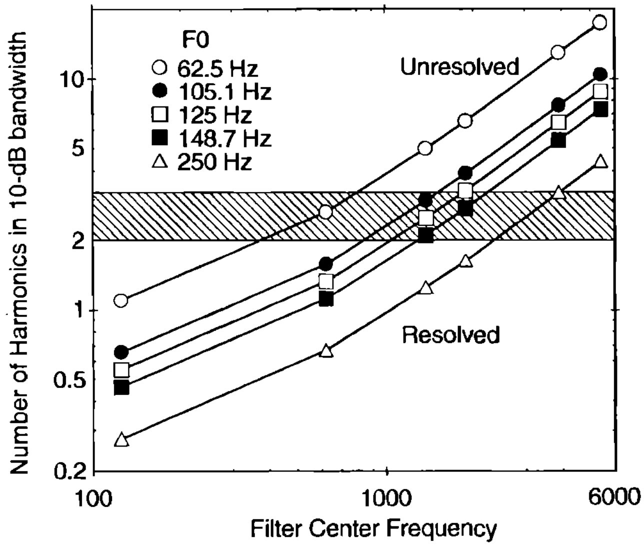

FIG. 4. Number of harmonics falling within the 10-dB-down bandwidth of an auditory filter as a function of filter center frequency for different $F0$'s. See Sec. II C for further details.

of permissible matches ranged from half an octave below the target frequency to one and a half octaves above the target frequency, and the starting frequency for the matcher was selected randomly from within this range. Listeners may have ignored the instruction to approach matches from both sides before signaling a match, and instead signaled a match as soon as it was sufficiently close to a true match. Thus since almost all trials would start with the matcher below the target such a strategy would result in flat matches. However, if this were the case, matches to the fundamental should also have been sharp, a finding which did not occur.

The interaction between $F0$ and spectral region found in experiments 1 and 2 suggests that harmonic resolution is important in determining the pitch mode. In the rest of this section we shall show that the results obtained are consistent with there being a transition from resolved to unresolved as the number of harmonics per auditory filter increases between frequency-independent criteria. Here, we define the resolvability of an harmonic in terms of the number of harmonics contained between the 10-dB-down points of an auditory filter centered on that harmonic. (The choice of 10-dB bandwidth is, to a certain extent, arbitrary.) The number of harmonics per filter is calculated by dividing the filter bandwidth (1.8 ERB) by $F0$ (Patterson *et al.*, 1988; Glasberg and Moore, 1990). Figure 4 shows, for several $F0$'s, the number of harmonics which, on average, fall within the 10-dB-down points of an auditory filter centered on the frequency shown on the abscissa. The two lines with solid symbols correspond to $F0$'s of 105.1 and 148.7 Hz, which Fig. 3(b) shows to be the limits of the region of unclear pitch when the stimulus is presented in the MID region. Lines are drawn across the graph at 2 and 3.25 harmonics per filter and demarcate areas where the harmonics are defined to be completely resolved, partially resolved, or unresolved. These cutoff points were selected so that no harmonics of an $F0$ of 148.7 Hz were resolved in the MID region, and only the upper harmonics of an $F0$ of 105.1 Hz were completely unresolved. According to this definition, for all the stimuli we have so far used, the harmonics have been unresolved in the LOW region, and resolved in the HIGH region. A transition from unresolved to resolved occurs in the MID region as $F0$ is increased. The upper few harmonics of a complex with an $F0$ of 62.5 Hz presented in the LOW region just fall within the partially resolved area. A closer examination of Fig. 3 shows that there is some evidence that pitch is becoming ambiguous (at least for one listener), and hence, according to our analysis, the harmonics need to be partially resolved.

To some extent, the finding that our definition can account for the pattern of results observed in experiment 2 is not surprising, as our definition of filter bandwidth was selected with these data in mind. However, we will show in Sec. IV A that the same definition can also account for the interaction between $F0$ and spectral region obtained in the three discrimination paradigms of experiments 3–5.

III. COMPARATIVE MEASURES OF PITCH ENCODING ACCURACY

Experiments 1 and 2 demonstrated that there is a difference between the perception evoked by resolved and by unresolved harmonics: Only complexes consisting of unresolved harmonics evoke a pitch which is sensitive to the periodicity of the envelope, and which increases by an octave when the harmonics are summed in ALT phase, compared to when they are summed in SINE phase. The experiments cannot reveal whether the pitch evoked by resolved harmonics is sensitive to the filtered envelope periodicities or to their spectral pattern, because envelope periodicity is not altered by the relative phases of harmonics until there are at least three harmonics per filter (Moore, 1977). All that can be concluded about the processing of resolved harmonics is that it ignores phase information between different auditory filters. In experiments 3 to 5 we compare the accuracy of the encoding of groups of resolved and unresolved harmonics, using stimuli identical to those in experiments 1 and 2. The different experiments do so by measuring FMTs, FMDDTs, and FDLs, respectively.

A. Experiment 3. Thresholds for detecting frequency modulation ("FMTs")

1. Method

Stimuli were presented using a 2I, 2AFC procedure with feedback. In one interval the stimulus was presented with no frequency modulation and in the other interval the stimulus was frequency modulated at 2.5-Hz rate and variable depth. Listeners were required to indicate which interval contained the frequency-modulated stimulus. Thresholds were obtained using Levitt's (1971) two-down, one-up adaptive procedure, which converged on the 70.7%-correct point on the psychometric function. The FM depth was multiplied by 1.07 after every incorrect response and divided by 1.07 after two consecutive correct responses, except for the trials before the first four turn-

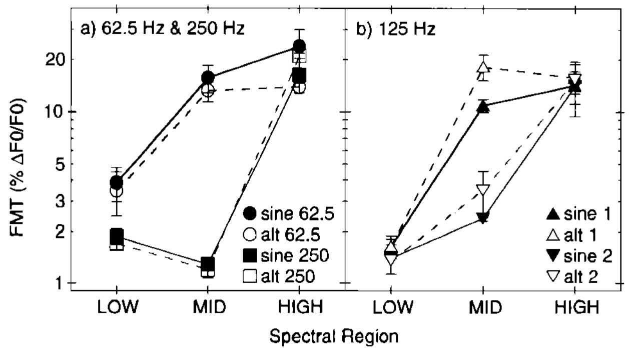

FIG. 5. Frequency modulation detection thresholds as a percentage of *F*0. Filled symbols represent sine-phase stimuli and open symbols alternating-phase stimuli. The left-hand panel shows results for *F*0's of 62.5 Hz (circles) and 250 Hz (squares) averaged across all listeners. The right-hand panel shows results for an *F*0 of 125 Hz, but with the results segregated into two groups of listeners (upward and downward triangles).

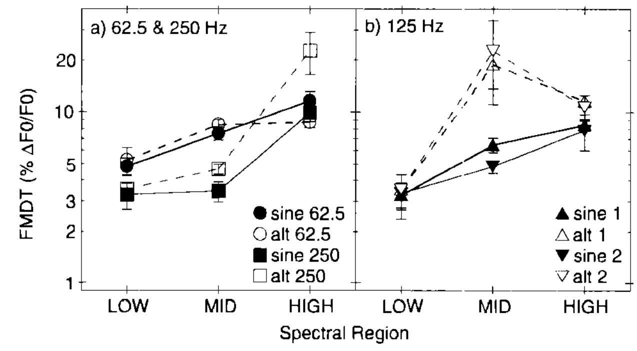

FIG. 6. As Fig. 5, but for FM direction discrimination thresholds.

points when a factor of 1.15 was used. Each run ended after 16 turnpoints, and the threshold for each run was obtained from the geometric mean of the modulation depths of the last 12 turnpoints. Listeners were tested at three nominal *F*0's (62.5, 125, 250 Hz), in the usual three spectral regions (LOW, MID, HIGH), and with ALT- and SINE-phase stimuli. Four listeners were tested at all three *F*0s, and thresholds were geometrically averaged from either two or three runs. An additional six listeners were tested at an *F*0 of 125 Hz and thresholds obtained from at least three runs. The FM was randomly selected on each trial to start on either a positive or negative zero crossing of the modulation, and the nominal *F*0 was varied by ±10% (by randomizing the playback rate) between intervals. These two forms of randomization were also used in experiment 4. They were intended to reduce the number of reliable detection cues available to the listener. For example, without these randomizations, the listener could make their decision based on the frequency $\frac{1}{4}$ way through both the reference and signal stimuli without needing to be able to detect an ongoing frequency change at all.

2. Results

FMTs, geometrically averaged across listeners, are shown in Fig. 5. The error bars indicate the single standard error of the average, with the variation associated with differences in overall performance between listeners removed.[2] Figure 5(a) shows the thresholds for stimuli in both SINE phase (solid symbols) and ALT phase (open symbols) at *F*0's of 62.5 and 250 Hz. For both *F*0s, the thresholds for stimuli in ALT phase are very similar to those for stimuli in SINE phase. For an *F*0 of 62.5 Hz (circles), thresholds are low in the LOW region, and high in the MID and HIGH regions; in contrast, thresholds at an *F*0 of 250 Hz (squares) are low in both the LOW and MID regions, and high in only the HIGH region. Figure 5(b) shows thresholds for an *F*0 of 125 Hz, but with the averages partitioned between two groups of listeners who showed different patterns of results. Listeners were divided according to the value of their threshold in the MID, SINE phase condition. If this threshold was closer to their LOW, SINE phase threshold than their HIGH, SINE phase threshold, they were placed in group 1; otherwise they were placed in group 2. In all cases this division was clear cut since all thresholds were well away from the dividing point. The data at *F*0's of 62.5 and 250 Hz did not support such a division. The thresholds for listeners in group 1 were similar to those for all listeners with an *F*0 of 62.5 Hz, whereas those for group 2 were more similar to those for all listeners at an *F*0 of 250 Hz. The results are consistent with the definitions of resolved and unresolved developed in Sec. II C. Low thresholds are obtained for all listeners for stimuli with any *F*0 presented in the LOW region and an *F*0 of 250 Hz in the MID region in agreement with our analysis that these stimuli consist of resolved harmonics. Similarly, high thresholds are obtained for all listeners and all stimuli with any *F*0 presented in the HIGH region, and for an *F*0 of 62.5 Hz in the MID region in agreement with our analysis that these stimuli consist of unresolved harmonics. The fact that listeners needed to be divided into two groups for an *F*0 of 125 Hz in the MID region is consistent with this being a transition region between resolved and unresolved harmonics.

B. Experiment 4. Thresholds for the discrimination of frequency modulation direction ("FMDDTs")

1. Method

A 3I, 2AFC paradigm was used. FM stimuli were presented in all three intervals. In the first interval and one of the second or third intervals the FM phase was the same (randomly chosen on each trial to be either 0° or 180°), whereas in the other interval the FM phase was different by 180°. The task was to identify which interval contained the different modulation. This is termed a modulation direction discrimination task because in order to perform the task listeners needed to determine the direction of the FM contour (i.e., up–down–up for 0° FM phase versus down–up–down for 180° FM phase). All other details were identical to experiment 3.

2. Results

Thresholds are plotted in Fig. 6 in exactly the same way as for the FM detection results of experiment 3 in Fig. 5. The same division between listeners determined in experiment 3 was used for an $F0$ of 125 Hz. The division between groups of listeners was less clear cut in this experiment, possibly because the difference between thresholds in the LOW and HIGH conditions was smaller. The results are largely similar to those obtained in experiment 3, except for the 125-Hz-$F0$ ALT-phase stimuli, presented in the MID region. For these stimuli, the FMDDTs are much larger than for the corresponding SINE stimuli [compare open and filled symbols in Fig. 6(b)], a difference which was much larger than for the corresponding FMT data of experiment 3. The difference in FMDDT between ALT- and SINE-phase stimuli is consistent with the interpretation that this combination produces an unclear pitch for ALT-phase stimuli. The fact that it was observed to a far less extent for the FMTs of experiment 3 supports Carlyon *et al.*'s (1992) conclusion that some feature other than pitch can be used in FM detection to maintain good performance.

C. Experiment 5. Frequency difference limens ("FDLs")

1. Method

Three listeners took part, all of whom had taken part in experiment 2, but only one of whom had participated in experiments 3 and 4. Stimuli were presented using a 3I, 2AFC procedure with feedback. The $F0$ was the same in the first and one of the second or third intervals, and either higher or lower by $\Delta F0$ in the other interval. The task was to determine which interval contained the different $F0$. The FM imposed on the stimuli always began at a positive zero crossing and had a constant depth of 5%. The baseline $F0$ used in each trial was randomly selected from a rectangular distribution of width $\pm 10\%$ centered on the nominal $F0$. The $F0$'s used in the standard and target intervals were obtained by adding $\pm \Delta F0/2$ to the base $F0$ in the standard intervals and subtracting $\pm \Delta F0/2$ from the base $F0$ in the target interval with the sign of $\pm \Delta F0/2$ being randomly selected on each trial. After every incorrect response $\Delta F0$ was multiplied by 1.19, and after two consecutive correct responses it was divided by 1.19, except for the trials before the first four turnpoints when a factor of 1.41 was used. Each run ended after ten turnpoints and the threshold for each run was obtained from the geometric mean of the $\Delta F0$'s of the last six turnpoints. Most of the thresholds reported are based on the geometric means of at least six such runs (although the data for JS at 148.1 Hz are only based on three runs). Listeners were tested at three nominal $F0$'s, in the usual three spectral regions, and with ALT and SINE-phase stimuli. The same low and high nominal $F0$'s were used for all three listeners (88.4 and 250 Hz), but the middle nominal $F0$ was chosen separately for each listener, and was 125 for SD and TS, and 148.1 for JS. The choice of $F0$ was based on the results of experiment 2, so that with low and high $F0$'s the stimulus harmonics would always be unresolved and resolved, respectively, and with the middle $F0$ they would be partially resolved in the MID region.

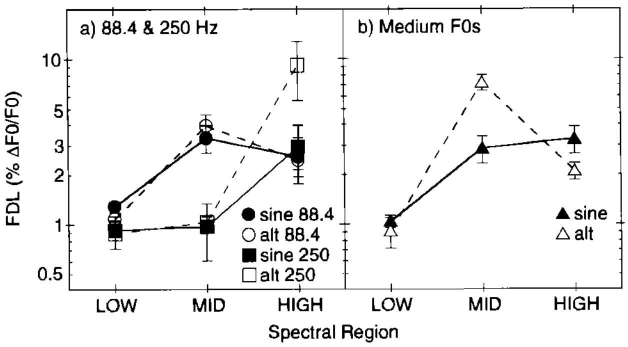

FIG. 7. Frequency difference limen averaged across three listeners. The left-hand panel shows results for $F0$'s of 88.4 Hz (circles) and 250 Hz (squares) averaged across all listeners. The right-hand panel shows results for the intermediate $F0$, which was selected individually for each listener (see text for details).

2. Results

The thresholds are shown as a percentage of the nominal $F0$ in Fig. 7. The left-hand panel shows thresholds for $F0$'s of 88.4 and 250 Hz, whereas the right-hand panel shows thresholds obtained at the intermediate $F0$. The results are not divided at the intermediate $F0$ because they all clearly fell within the same group (note that a different set of listeners are used from experiments 3 and 4). The results show a very similar interaction between spectral region and $F0$ to that obtained in experiment 4. It is also worth noting that, as in experiment 4, thresholds in the MID region at an $F0$ of 125 Hz were much higher for ALT-phase than for SINE-phase stimuli, consistent with the ALT-phase stimuli having a weak or ambiguous pitch.

IV. DISCUSSION

A. Can a single measure of resolvability predict performance in different tasks?

In Sec. II C we showed that whether ALT-phase stimuli were perceived as having a pitch of $F0$ or $2F0$ depended upon whether fewer than 2, or more than 3.25 harmonics, respectively, interacted within a single auditory filter (defined by its 10-dB-down bandwidth). We used these cutoff points as criteria for whether stimulus harmonics were resolved, or unresolved, respectively. The same rule can also be applied to the discriminability of SINE-phase stimuli, where low thresholds occur for stimuli with $F0$ and spectral region combinations which our analysis calculates to result in resolved harmonics, whereas high thresholds occur under conditions where we calculate there are unresolved or partially resolved harmonics. For example, FDLs (solid symbols in Fig. 7) are all below 1.5% for resolved stimuli, as defined by our criterion, with all other thresholds above 2.5%. Similar dichotomies can be applied to the FMTs (Fig. 5) and to the FMDDTs (Fig. 6). Several others (Cullen *et al.*, 1986; Hoekstra, 1979; Houtsma and Smurzynski, 1990) have also shown that frequency discrimination thresholds for high harmonics are larger than for low harmonics. Houtsma and Smurzynski (1990) found that, when the resolution of the lowest harmonics in the stimulus was decreased from 1.8 harmonics/filter to 2.9 harmonics/filter, FDLs increased from about 0.5 Hz to above 5 Hz for an $F0$ of 200 Hz. Similarly, Cullen and Long (1986) found an increase in rate jnds for pulse trains with an $F0$ of 200 Hz from 0.6 Hz when the stimulus was unfiltered to 4.5 Hz when the trains were high-pass filtered at a cutoff of 2.5 kHz (resolution of 2.7 harmonics/filter just above cutoff). Hoekstra (1979) measured frequency difference limens for $\frac{1}{3}$-oct-wide stimuli; he found that thresholds increased markedly when the resolvability of the lower components was increased from approximately 2.0 to 3.7 harmonics/filter. The position of the transition region in these experiments is in fairly good agreement with our criterion.

None of the above tasks require listeners to separate individual harmonics from the complex in order to perform the task, and all give a similar measure of the harmonic separation, or resolvability, at which there is a transition from good to poor performance. We have used the terms resolved and unresolved to describe the harmonic separation on either side of this transition. However, some might argue that harmonics can only be truly described as resolved if it is demonstrated that they can be processed independently of each other. For example, Moore and Ohgushi (1993) found that the ability to hear a single component out from an inharmonic complex was very poor when the resolution was 2.6 harmonics/filter and increased to near perfect when the resolution was improved to 1.2 harmonics/filter. These values are smaller than those we obtained as cutoff values in our definitions for unresolved and resolved, and are comparable with a transition from good performance to poor performance at resolutions of between 1.0 and 1.8 harmonics/filter in tasks which require listeners to detect the mistuning of a single harmonic within a complex (Moore *et al.*, 1984, 1985b). It would therefore appear that to be able to perceptually segregate a harmonic from its companion harmonics requires a greater harmonic separation than we estimate from our tasks. This does not invalidate our findings, but does suggest that the critical separation at which harmonics appear to interact within individual auditory filters (i.e., become unresolved) depends upon the paradigm and therefore the definition of resolvability may be task dependent.

Patterson (1987) asked listeners to discriminate between a harmonic complex in cosine phase and one in which the phase of alternate harmonics was changed between cosine phase and a variable phase. The variable phase was adjusted until listeners were correct 90% of the time. Fundamental frequency and the lowest harmonic used were varied orthogonally, so performance can be linked directly to harmonic resolution. It was found that the phase threshold was lower the less resolved the stimulus harmonics were, although thresholds tended to be larger for higher $F0$'s, even when comparing complexes with similar resolvability. Our criterion does not work so well for these stimuli. Although it is true that the thresholds obtained by Patterson were higher for stimuli with a resolution better than two harmonics/filter than for stimuli with a resolution poorer than three harmonics/filter, equal performance did not always follow from equal resolution. We have no convincing explanation for this failure. However, Patterson argued that the detection of the phase change was mediated by small perturbations in the minima of the waveform, so it is possible that a model which takes into account the attenuation of the outer harmonics for resolutions of around three harmonics/filter may be more successful.

B. Comparison of three different measures of discriminability

All three measures of discrimination show a similar pattern of thresholds for harmonics that were either completely resolved or completely unresolved, as defined above: Thresholds are consistently lower for resolved than for unresolved harmonics. Thus in this respect, the discrimination results obtained with static stimuli (Hoekstra, 1979; Houtsma and Smurzynski, 1990; Moore *et al.*, 1984) extend to FM stimuli, and to a novel form of discrimination measure, the FM direction discrimination threshold (FMDDT). However, there are differences between the measures when the stimulus was partially resolved, which for ALT-phase stimuli led to an unclear pitch. Carlyon *et al.* (1992) showed, for stimuli with an $F0$ of 125 Hz presented in the MID region, FMTs did not differ markedly between stimuli in ALT and SINE phase. This result is largely confirmed here: Although there is some evidence in Fig. 4 that the ALT-phase thresholds are higher than those obtained for SINE-phase stimuli, this effect is much smaller than that seen in Figs. 5 and 6 for FMDDTs and FDLs, respectively. Experiments 1 and 2 showed that an $F0$ of 125 Hz in the MID region produced an unclear pitch for ALT-phase stimuli, so it is reasonable to argue that the poor performance in the FDL and FMDDT tasks is because they require pitch to be extracted, whereas the FMT can proceed without explicit pitch extraction. Note in particular that FDL and FMDDT values at 88 Hz in the HIGH region are not generally higher for ALT than for SINE phase, even though the ALT-phase stimuli have a perceived pitch an octave above their true $F0$ (experiment 2). Thus although playing the stimulus in ALT phase caused its perceived $F0$ to double, this did not affect its discriminability: Listeners could encode the pitch accurately, even though it did not correspond to $F0$. In other words these experiments show that phase affects the discriminability of stimuli when it affects the clarity of the perceived pitch, but not when it simply shifts the mean perceived value.

C. Implications for theories of pitch perception

It has already been shown that the traditional theories which can only operate on resolved harmonics (Goldstein, 1973; Piszczalski and Galler, 1979; Terhardt, 1974; Terhardt *et al.*, 1982), or unresolved harmonics (Schouten, 1940, 1970) cannot, individually, explain the perception of pitch for both resolved and unresolved harmonics (e.g., Hoekstra, 1979; Houtsma and Smurzynski, 1990). For this reason, among others, theories were suggested which used

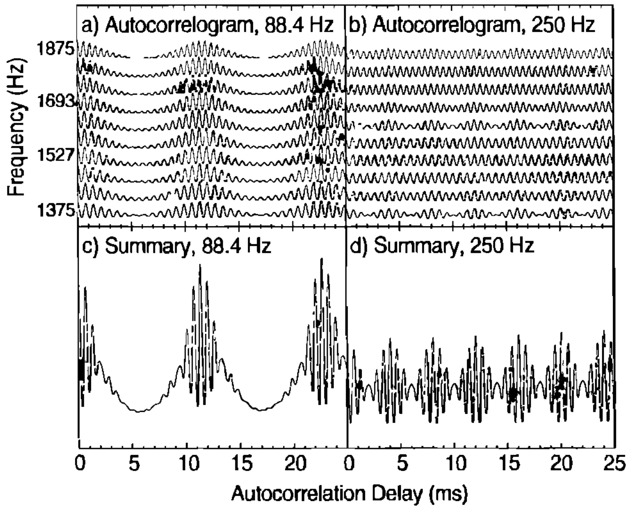

FIG. 8. Autocorrelograms and summary autocorrelations (Meddis and Hewitt, 1991a) for stimuli presented in the MID region with $F0$'s of 88.4 and 250 Hz (see text for further details).

a common mechanism across the entire frequency spectrum (Assmann and Summerfield, 1990; Houtsma and Smurzynski, 1990; Meddis and Hewitt, 1991a,b; Moore, 1989; Slaney and Lyon, 1990). An alternative approach would be to assume that the two mechanisms described by the traditional theories jointly operated, depending upon whether the harmonics were resolved or not. This form of dual-mechanism theory obviously has no problem in explaining why thresholds for resolved and unresolved harmonics are so different, but the single-mechanism theories are posed more problems. Houtsma and Smurzynski (1990) argue that the Srulovicz and Goldstein (1983) model, which was initially proposed to account for the pitch of resolved harmonics only, can also predict increased thresholds for unresolved harmonics; however this argument does not hold for the other common-mechanism theories (Assmann and Summerfield, 1990; Meddis and Hewitt, 1991a; Moore, 1989; Slaney and Lyon, 1990). These perform a temporal analysis of the output of each channel of an auditory filter bank, and then combine the periodicity information across channel frequency. The most explicitly defined mechanism for doing this is proposed by Meddis and Hewitt (1991a) who suggest that each hair-cell output is autocorrelated and these autocorrelations summed across frequency to form a summary autocorrelation. Figure 8 shows the autocorrelations for individual channels in the MID region for $F0$'s of 88.4 and 250 Hz [Fig. 8(a), (b)] and their associated summary autocorrelations [Fig. 8(c), (d)]. These $F0$'s were chosen because they produce high and low FDL thresholds, being unresolved and resolved, respectively, according to our definition. The most obvious change between Fig. 8(c) and (d) is the larger number of cycles of "envelope" in the 250-Hz case. This should be ignored because it also occurs when stimuli with these $F0$'s are both presented in either the LOW or the HIGH regions (both resolved or unresolved, respectively), and FDLs for 250-Hz stimuli are not markedly lower than those for 88.4 Hz in those cases [Fig. 7(a)]. The point to notice is that the fine structure of the summary autocorrelation has the same period for both $F0$'s (this is characteristic of the highest frequency component in the signal), and that the ratio of the amplitudes of the main peak to the secondary peaks is approximately the same for both $F0$'s (or slightly larger at 88.4 Hz). Moore (1977), followed by Meddis and Hewitt (1991a), has argued that the number of candidate pitch peaks in the region of the "true" pitch peak indicates how salient the pitch percept will be; according to this argument the pitch encoding for 88.4 and 250 Hz should be equally accurate (or the 88.4 Hz better). In fact, thresholds are much higher for the 88.4-Hz stimulus, suggesting that the summary autocorrelation model is inadequate to fully explain performance for these stimuli. If the pitch extraction mechanism did not form summary autocorrelations but operated directly upon the autocorrelation display, then the dichotomy between resolved and unresolved thresholds would be easier to explain since the individual channel autocorrelations [Fig. 8(a) and (b)] are very different for the different $F0$'s. In its current form, therefore, the Meddis and Hewitt (1991a) model has difficulties explaining our data. However, it is possible that future developments may solve this problem. The question of whether the pitch perception of resolved and unresolved harmonics can be explained by a single mechanism is examined in more detail in our companion paper (Carlyon and Shackleton, 1994).

D. Summary

The results presented here demonstrate two different modes of pitch perception. One mode occurs with stimuli comprising unresolved harmonics, whereas the other occurs with stimuli comprising resolved harmonics. Pitch-preference and pitch-matching experiments revealed that summing unresolved harmonics in alternating phase caused their pitch to increase by an octave relative to that of otherwise-identical sine-phase stimuli, whereas this was not the case for resolved harmonics. A simple rule, defining whether or not a given group of harmonics was resolved, could account for the results of the pitch experiments. The rule was that, when fewer than two harmonics occurred, on average, within the 10-dB bandwidth of an auditory filter then the stimulus harmonics were effectively resolved; whereas when there were more than 3.25 harmonics/filter then they were effectively unresolved. In between there was a transition region. The same definition of resolvability could also account for the results of three different experiments which aimed to measure the discriminability of $F0$ differences. For each of the three measures (FM detection thresholds, FMTs; FM direction discrimination thresholds, FMDDTs; and frequency difference limens, FDLs), thresholds were higher with unresolved than with resolved harmonics. For stimuli of borderline resolvability, summing harmonics in alternating phase led to an unclear pitch, and greatly increased FDLs and FMDDTs, but only slightly increased FMTs. It therefore appears that, when detecting FM, listeners are able to use some cue, other than pitch changes, which is not available when detecting static

$F0$ differences (FDL) or when discriminating the direction of FM (FMDDT). It is, however, not clear exactly what this cue is.

ACKNOWLEDGMENTS

The research reported here was supported by a project grant from the Image Interpretation Initiative of the SERC (U.K.), by a Wellcome Trust Postdoctoral Fellowship to the first author, and by a Royal Society University Research Fellowship to the second author. We are grateful to Brian Moore, Ray Meddis, Armin Kohlrausch, and Chris Plack for helpful comments on a previous version of the manuscript.

[1]Ohgushi (1983) pointed out that the shortest interspike interval (ISI) observed in auditory-nerve fibers in response to a pure tone was slightly longer than the period of that tone, due to the refractory nature of nerve-fiber firing patterns. Furthermore, he noted that this effect increased with increasing frequency, so that, in terms of ISIs, high-frequency tones would sound "flatter" than low-frequency tones, relative to their genuine frequencies. Therefore, in order to match to the octave, listeners would have to counteract this by adjusting the high-frequency tone slightly "sharp." This explanation cannot account for our results in which listeners matched the envelope repetition rates of two complex tones in the same frequency region to be nearly equal. Under these circumstances the effects of auditory-nerve refactory periods would have been unaffected by the phase of the stimuli, and so listeners should not have shown any systematic bias.

According to Terhardt (1974), pitch is derived exclusively from the spectral properties of the stimulus, and so his theory is inconsistent with the observed effect of phase on pitch unless one assumes that, for some reason, listeners were making "octave errors" when components were unresolved. Our experiment 2 showed that this was not the case, so, like Ohgushi's theory, Terhardt's model cannot account for the observed pitch matches.

[2]Each listener's mean score across all conditions was subtracted from his/her thresholds for a given condition before the standard error across listeners was calculated for that condition. This technique emphasizes the similarity between the *pattern* of results between subjects, but conceals the differences between the overall performance level of subjects. In experiment 3 the ratio between best and worst listener's thresholds was always less than three, whereas in experiment 4 it was always less than ten. The technique was applied only to the calculation of standard errors, and did not affect the mean values plotted.

ANSI (**1969**). ANSI S3.6-1969, "Specifications for audiometers" (American National Standards Institute, New York).

Assmann, P. F., and Summerfield, Q. (**1990**). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," J. Acoust. Soc. Am. **88**, 680–697.

Carlyon, R. P., Demany, L., and Semal, C. (**1992**). "Detection of across-frequency differences in fundamental frequency," J. Acoust. Soc. Am. **91**, 279–292.

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.

Cullen, J. K., Jr., and Long, G. (**1986**). "Rate discrimination of high-pass filtered pulse trains," J. Acoust. Soc. Am. **79**, 114–119.

Dadson, R. S., and King, J. H. (**1952**). "A determination of the normal threshold of hearing and its relation to the standardisation of audiometers," J. Laryngol. Otol. **66**, 366–378.

Flanagan, J. L., and Guttman, N. (**1960a**). "On the pitch of periodic pulses," J. Acoust. Soc. Am. **32**, 1308–1319.

Flanagan, J. L., and Guttman, N. (**1960b**). "Pitch of periodic pulses without fundamental component," J. Acoust. Soc. Am. **32**, 1319–1328.

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.

Guttman, N., and Flanagan, J. L. (**1964**). "Pitch of high-pass-filtered pulse trains," J. Acoust. Soc. Am. **36**, 757–765.

Hartmann, W. M., McAdams, S., and Smith, B. K. (**1990**). "Hearing a mistuned harmonic in an otherwise periodic complex tone" J. Acoust. Soc. Am. **88**, 1712–1724.

Hoekstra, A. (**1979**). "Frequency discrimination and frequency analysis in hearing," Ph.D., Institute of Audiology, University Hospital, Groningen, The Netherlands.

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.

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

Licklider, J. C. R. (**1951**). "A duplex theory of pitch perception," Experientia **7**, 128–133.

Lundeen, C., and Small, A. M. (**1984**). "The influence of temporal cues on the strength of periodicity pitches," J. Acoust. Soc. Am. **75**, 1578–1587.

Mathes, R. C., and Miller, R. L. (**1947**). "Phase effects in monaural perception," J. Acoust. Soc. Am. **19**, 780–797.

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

Meddis, R., and Hewitt, M. (**1991b**). "Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery: II. Phase sensitivity," J. Acoust. Soc. Am. **89**, 2883–2894.

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), pp. 349–358.

Moore, B. C. J. (**1989**). *An Introduction to the Psychology of Hearing* (Academic, New York).

Moore, B. C. J., and Glasberg, B. R. (**1986**). "Thresholds for hearing mistuned partials as separate tones in harmonic complexes," J. Acoust. Soc. Am. **80**, 479–483.

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 cochlear impairments," in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London), pp. 421–430.

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

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., Glasberg, B. R., and Peters, R. W. (**1985a**). "Relative dominance of individual partials in determining the pitch of complex tones," J. Acoust. Soc. Am. **77**, 1853–1860.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1985b**). "Thresholds for the detection of inharmonicity in complex tones," J. Acoust. Soc. Am. **77**, 1861–1867.

Ohgushi, K. (**1983**). "The origin of tonality and a possible explanation of the octave enlargement phenomenon," J. Acoust. Soc. Am. **73**, 1694–1700.

Patterson, R., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1988**). *Spiral VOS Final Report, Part A: The Auditory Filterbank* (Applied Psychology Unit, Cambridge, England).

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

Piszczalski, M., and Galler, B. A. (**1979**). "Predicting musical pitch from component frequency ratios," J. Acoust. Soc. Am. **66**, 710–720.

Plomp, R. (**1967**). "Pitch of complex tones," J. Acoust. Soc. Am. **41**, 1526–1533.

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

Ritsma, R. J., and Engel, F. L. (**1964**). "Pitch of frequency modulated signals," J. Acoust. Soc. Am. **36**, 1637–1655.

Rosenberg, A. E. (**1965**). "Effect of masking on the pitch of periodic pulses," J. Acoust. Soc. Am. **38**, 747–758.

Russell, P., and Darwin, C. J. (**1991**). "Real-time synthesis of complex sounds on a Mac II with 56001 DSP chip," Br. J. Audiol. **25**, 59–60.

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

Schouten, J. F. (**1970**). "The residue revisited," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Lieden), pp. 41–58.

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

Slaney, M., and Lyon, R. F. (**1990**). "A perceptual pitch detector," Proc. Int. Conf. Acoust. Speech, Signal Process. (IEEE, New York), Vol. 5, pp. 357–360.

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.

Sundberg, J. E. F., and Lindqvist, J. (**1973**). "Musical octaves and pitch," J. Acoust. Soc. Am. **54**, 922–929.

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

Terhardt, E., Stoll, G., and Seewann, M. (**1982**). "Algorithm for extraction of pitch salience from complex tonal signals," J. Acoust. Soc. Am. **71**, 679–688.

Thurlow, W. R., and Small, A. M. (**1955**). "Pitch perception for certain periodic auditory stimuli," J. Acoust. Soc. Am. **27**, 132–137.

Warren, R. M., and Wrightson, J. M. (**1981**). "Stimuli producing conflicting temporal and spectral cues to frequency," J. Acoust. Soc. Am. **70**, 1020–1024.