Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number?

Joshua G. Bernstein[a)] and Andrew J. Oxenham[b)]
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and Harvard—MIT Division of Health Sciences and Technology, Speech and Hearing Bioscience and Technology Program, Cambridge, Massachusetts 02139

(Received 31 October 2002; revised 16 March 2003; accepted 17 March 2003)

Three experiments investigated the relationship between harmonic number, harmonic resolvability, and the perception of harmonic complexes. Complexes with successive equal-amplitude sine- or random-phase harmonic components of a 100- or 200-Hz fundamental frequency (f_0) were presented dichotically, with even and odd components to opposite ears, or diotically, with all harmonics presented to both ears. Experiment 1 measured performance in discriminating a 3.5%–5% frequency difference between a component of a harmonic complex and a pure tone in isolation. Listeners achieved at least 75% correct for approximately the first 10 and 20 individual harmonics in the diotic and dichotic conditions, respectively, verifying that only processes before the binaural combination of information limit frequency selectivity. Experiment 2 measured fundamental frequency difference limens (f_0 DLs) as a function of the average lowest harmonic number. Similar results at both f_0's provide further evidence that harmonic number, not absolute frequency, underlies the order-of-magnitude increase observed in f_0 DLs when only harmonics above about the 10th are presented. Similar results under diotic and dichotic conditions indicate that the auditory system, in performing f_0 discrimination, is unable to utilize the additional peripherally resolved harmonics in the dichotic case. In experiment 3, dichotic complexes containing harmonics below the 12th, or only above the 15th, elicited pitches of the f_0 and twice the f_0, respectively. Together, experiments 2 and 3 suggest that harmonic number, regardless of peripheral resolvability, governs the transition between two different pitch percepts, one based on the frequencies of individual resolved harmonics and the other based on the periodicity of the temporal envelope. © *2003 American Institute of Physics.* [DOI: 10.1121/1.1572146]

PACS numbers: 43.66.Hg, 43.66.Fe, 43.66.Ba [NFV]

I. INTRODUCTION

The mechanisms underlying pitch perception have been a matter of intense debate ever since Ohm (1843) disputed Seebeck's (1841) description of the phenomenon of the missing fundamental frequency (f_0). More recently, one aspect of this debate has been concerned with the mechanisms underlying the different contributions that low- and high-frequency harmonics make to the overall perceived pitch of a harmonic complex. Early work showed a dominant frequency region for pitch that was determined by both relative and absolute frequency relations. Ritsma (1967) demonstrated that the third through fifth harmonics dominated the perceived pitch for various f_0's, such that the dominant frequency region for pitch was relative to the complex's f_0. Investigating a wider range of f_0's, Plomp (1967) found that the harmonics that dominated the perceived pitch also depended on the f_0 of the complex, suggesting that absolute frequency also influenced the dominance region.

Most models of pitch perception can account qualitatively for the dominance of low harmonics in determining the overall pitch and for the greatly reduced pitch salience observed when only high harmonics are presented. However, the mechanisms by which they do so differ considerably. For instance, models that rely on the spatial separation of frequency components along the cochlear partition (e.g., Goldstein, 1973; Wightman, 1973; Terhardt, 1974, 1979) predict that pitch salience will deteriorate as the spacing between the individual components within a complex becomes so small that the individual peaks in the cochlear representation are no longer resolved. Because the components of a harmonic complex are equally spaced on a linear frequency scale, but the absolute bandwidths of auditory filters increase with increasing center frequency (CF), the density of harmonics per auditory filter increases with increasing harmonic number. As a result, low-order harmonics are resolved from one another, but higher-order harmonics begin to interact within single auditory filters and eventually become unresolved. In contrast, models based on the autocorrelation of auditory-nerve fiber activity, pooled across the total population of fibers (e.g., Meddis and Hewitt, 1991a,b; Cariani and Delgutte, 1996; Meddis and O'Mard, 1997), predict poorer resolution within the model (and hence reduced performance in f_0 discrimination) as the *absolute* frequency of components increases (Cariani and Delgutte, 1996; Carlyon, 1998), due primarily to the roll-off in the phase-locking properties of auditory-nerve fibers above about 1.5 kHz (Weiss and Rose, 1988). These two categories of models are often referred to as "place" and "temporal" models, respectively.

[a)]Corresponding author electronic mail: jgbern@mit.edu
[b)]Electronic mail: oxenham@mit.edu

However, it should be noted that the term ‘‘place model’’ does not necessarily imply that the frequencies of individual harmonics are encoded via a place mechanism. Instead it is possible that the frequency information at each place is encoded via a temporal mechanism (Srulovicz and Goldstein, 1983; Shamma and Klein, 2000). Nevertheless, it is important for these place models that the components are sufficiently well resolved for the frequency of each to be estimated individually.

The defining role of absolute frequency and phase locking, implied by temporal models based on the pooled autocorrelation function, has been called into question by various psychophysical experiments indicating that relative frequency relationships play an important role in the deterioration of pitch salience for high-order harmonics. Houtsma and Smurzynski (1990) estimated pitch salience, in terms of melodic interval recognition and fundamental frequency difference limens (f_0 DLs), for complex tones comprising 11 successive harmonics as a function of the lowest harmonic present. They found that for both measures, performance was much poorer when only harmonics above the 10th were presented than when at least some harmonics below the 10th were present. Although they carried out their experiment at only one f_0 (200 Hz), meaning that the respective influences of absolute and relative frequencies could not be distinguished, earlier research with two harmonics (Houtsma and Goldstein, 1972), and later research with many harmonics (Carlyon and Shackleton, 1994; Shackleton and Carlyon, 1994; Kaernbach and Bering, 2001), strongly support the idea that performance in such tasks is limited primarily by the lowest harmonic number present, and not by the lowest absolute frequency present.

While it has been generally assumed that pitch discrimination deteriorates when only high harmonics are present because the harmonics are peripherally unresolved (Houtsma and Smurzynski, 1990; Carlyon and Shackleton, 1994; Shackleton and Carlyon, 1994), certain results in the literature cast some doubt on this interpretation. Houtsma and Goldstein (1972) estimated the pitch strength of harmonic complexes consisting of two successive components by measuring performance in musical interval identification. Harmonics that are unresolved when both are presented to the same ear (monotic) become resolved when presented to opposite ears (dichotic). If strong pitch salience required the presence of resolved harmonics, we might expect stronger pitch salience when two normally unresolved harmonics (i.e., unresolved under monotic presentation) are presented dichotically. However, the decrease in performance with increasing harmonic number was the same under monotic and dichotic presentations, suggesting that the decrease in pitch salience with increasing harmonic number may not be due to the harmonics becoming unresolved *per se*. Arehart and Burns (1999) reported similar results using three musically trained hearing-impaired listeners.

This paper further investigates the transition in f_0 DLs found in the data of Houtsma and Smurzynski (1990), to determine whether the frequency at which it occurs is defined by harmonic resolvability, harmonic number regardless of resolvability, or absolute frequency. An f_0 DL paradigm (Houtsma and Smurzynski, 1990) was used to test whether presenting normally unresolved components to opposite ears improves performance. Under diotic presentation, all components were presented to both ears, such that the peripheral spacing between components was the f_0. Under dichotic presentation, even and odd components were presented to opposite ears, such that peripheral spacing between components was twice the $f_0 (2 f_0)$. The approach differs from those of two earlier studies addressing this issue (Houtsma and Goldstein, 1972; Arehart and Burns, 1999) in two principal ways. First, the f_0 discrimination task does not require the musical training that is necessary for a musical interval identification task. Second, 12-component complexes yield a much stronger pitch salience than the relatively weak pitch elicited by two-tone complexes, even with low-order harmonics.

Underlying this study was the important assumption that approximately twice as many harmonics should be resolved in the dichotic conditions, where the peripheral frequency spacing between components is twice that of the diotic conditions. The first experiment was designed to test the validity of this assumption. In addition, experiment 1 addressed the discrepancy in the literature between direct and indirect estimates of harmonic resolvability, as described below.

II. EXPERIMENT 1: RESOLVABILITY OF INDIVIDUAL HARMONICS

A. Rationale

The existing studies on pitch perception show very good consistency in terms of the locus of the transition region between good and poor f_0 discrimination (Cullen and Long, 1986; Houtsma and Smurzynski, 1990). However, as pointed out by Shackleton and Carlyon (1994), while these data sets show a transition that occurs between harmonic numbers 10 and 13, direct measures of individual component resolvability have shown that listeners are generally only able to hear out the first five to eight harmonics of a harmonic complex (Plomp, 1964; Plomp and Mimpen, 1968). Similarly, Shackleton and Carlyon (1994) concluded that the limits of the resolvability of individual components within an inharmonic tone complex, as measured by Moore and Ohgushi (1993), were also lower than those estimated indirectly using f_0 DLS for harmonic tone complexes.

One reason for this discrepancy might be the nature of the respective tasks. Musicians have been shown to have better performance than nonmusicians in ‘‘hearing out’’ harmonics (Soderquist, 1970; Fine and Moore, 1993), while their auditory filter bandwidths are not significantly different (Fine and Moore, 1993). The difference between direct and indirect estimates of peripheral resolvability may be attributable to attentional limitations, whereby, in hearing out individual partials, subjects may have difficulty overcoming their perceptual fusion of the complex into a single auditory object. The difference could also be due to other nonperipheral limitations. In contrast to the Plomp (1964) and Moore and Ohgushi (1993) studies, which required subjects to hear out an individual partial presented simultaneously with a complex, this study gated the target harmonic on and off repeatedly within the presentation interval. This strategy was designed to help overcome any nonperipheral limitations and to encourage perceptual segregation, while not affecting peripheral resolvability.[1] If good f_0 discrimination depends on the presence of peripherally resolved harmonics, we expect that listeners should be able to hear out approximately ten harmonics—more than the five to eight measured by Plomp (1964).

B. Methods

In this and subsequent experiments, all subjects had some degree of musical training. The least musically trained subject had 4 years of instruction in middle school, while the most musically trained were two professional musicians with more than 18 years formal training. All subjects had normal hearing (15 dB HL or less *re* ANSI-1969 at octave frequencies between 250 Hz and 8 kHz). Four subjects (ages 18–26, two female) participated in this experiment.

All stimuli were presented in a background noise, uncorrelated between the two ears which we will call modified uniform masking noise (UMN_m). This noise is similar to uniform masking noise (UMN) (Schmidt and Zwicker, 1991), in that it is intended to yield pure-tone masked thresholds at a constant sound pressure level (SPL) across frequency, but the spectrum is somewhat different; UMN_m has a long-term spectrum level that is flat (15 dB/Hz SPL in our study) for frequencies below 600 Hz, and rolls off at 2 dB/oct above 600 Hz. The noise was low-pass filtered with a cutoff at 10 kHz. Thresholds for pure tones at 200, 500, 1500, and 4000 Hz in UMN_m in the left ear were estimated via a three-alternative forced-choice, two-down, one-up adaptive algorithm (Levitt, 1971). For each subject, pure tone thresholds in UMN_m fell within a 5-dB range at all four frequencies tested, such that harmonic components presented at equal SPL had nearly equal sensation level (SL). As an approximation, we defined 0 dB SL for each subject as the highest of the thresholds across the four frequencies tested, which ranged from 29.7 to 33 dB SPL across all subjects in this and subsequent experiments.

The stimuli were generated digitally and played out via a soundcard (LynxStudio LynxOne) with 24-bit resolution and a sampling frequency of 32 kHz. The stimuli were then passed through a programmable attenuator (TDT PA4) and headphone buffer (TDT HB6) before being presented to the subject via Sennheiser HD 580 headphones. Subjects were seated in a double-walled sound-attenuating chamber.

Each trial in the experiment consisted of two intervals, each with a 1-s duration, separated by 375 ms. The first interval contained three bursts of a 300-ms sinusoid (referred to as the comparison tone), including 20-ms Hanning window onset and offset ramps, separated by 50-ms silent gaps. The second interval consisted of a harmonic complex with the first 40 successive harmonics of the f_0 with duration 1000 ms, including 20-ms Hanning window onset and offset ramps. Components were presented in random phase to ensure that the frequency of the target component was detectable only if the component was spectrally resolved.[2] The target component was gated on and off in the same manner as in the first interval, while all the other components were on continuously throughout the interval. Each component was presented at a nominal 15 dB SL (adjusted for each

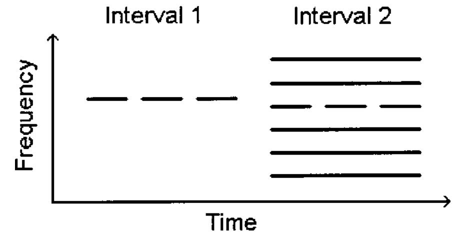

FIG. 1. Schematic of the stimuli used in experiment 1. Interval 2 contains a 40-component harmonic complex, with the target harmonic gated on and off to perceptually segregate it from the complex. Interval 1 contains a pure-tone probe, higher or lower in frequency than the target harmonic in interval 2, gated in the same way as the target harmonic.

subject), such that the stimuli in this experiment were similar in level to those used in experiment 2. The task was a two-alternative forced-choice task, where the listener was required to discriminate which of the comparison tone (interval 1) or target tone (interval 2) was higher in frequency. A schematic of the stimuli is shown in Fig. 1.

Four conditions were presented, for all combinations of the harmonic complex in interval 2 presented diotically or dichotically, with a 100- or 200-Hz average f_0 ($\overline{f_0}$). Fifty trials for each of ten target harmonic numbers in each condition were presented (diotic: 5 through 14, inclusive; dichotic: 11, 12, 13, 14, 16, 18, 20, 22, 25, and 28), for a total of 500 trials per condition. The trials were presented in runs, each consisting of five trials for each of the ten harmonics for one condition, presented in random order. In the dichotic conditions, the comparison and target harmonics were always presented to the same ear throughout a run, and the distribution of the even and odd harmonics of the complex in interval 2 to the left and right ears was varied accordingly. For example, for a trial where the target 14th harmonic and comparison tone were presented to the right ear, the even harmonics in interval 2 were also presented to the right ear. In the dichotic conditions, five runs were presented with the target in the left ear, and five runs were presented with the target in the right ear.

The difference (Δf) between the frequency of the comparison tone (f_{comp}) and that of the target tone (f_{targ}) was set as a proportion of f_{targ}. This is different from Plomp's (1964) experiment, where he required listeners to identify which of two pure tones was in fact a component of the complex. One comparison tone was at the frequency of one of the components, and the other was halfway between the frequencies two successive components, such that it always fell at the same place relative to the target tone on a *linear* scale. In our experiment, the comparison tone was adjusted relative to the target tone on a *logarithmic* scale, ensuring that any decrease in performance with increasing harmonic number reflects a reduction in resolvability, and not the increase in linear pure tone DLs with increasing frequency (Moore, 1973).

In each trial, f_{comp} was either higher or lower than f_{targ}, each with probability 0.5, with $\Delta f = |f_{targ} - f_{comp}|$ chosen from a uniform distribution of 3.5 to 5.0% of the f_{targ}. The value of Δf was always at least 3.5% of the f_{targ}, which is well above the frequency discrimination threshold for tones in quiet (Moore, 1973). The f_0 of the complex was randomly chosen from a uniform distribution between $0.935\,\overline{f_0}$ and

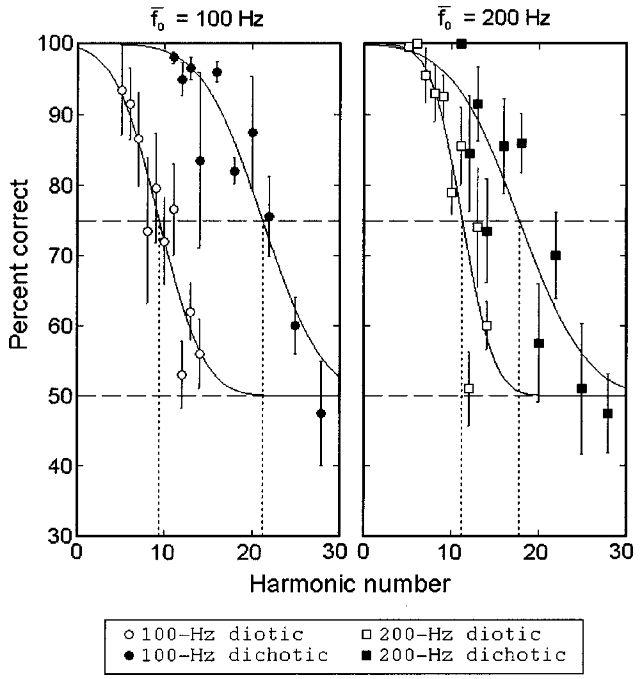

FIG. 2. Mean results of experiment 1, showing percent correct in identifying the probe tone as higher or lower than the target tone as a function of harmonic number. Error bars represent plus and minus one standard error across the individual scores for the four subjects. Open symbols indicate diotic conditions, with all harmonics presented to both ears; filled symbols indicate dichotic conditions, with odd and even harmonics presented to opposite ears. The left and right panels show results with f_0's of 100 and 200 Hz, respectively. Solid lines represent the best fits of the erfc function [Eq. (1), footnote 4] to the pooled data. The limit of harmonic resolvability, defined as the harmonic that yields 75% correct performance, is depicted by a vertical dotted line. The upper and lower horizontal dashed lines indicate 75% correct (limit of harmonic resolvability) and 50% correct (chance), respectively.

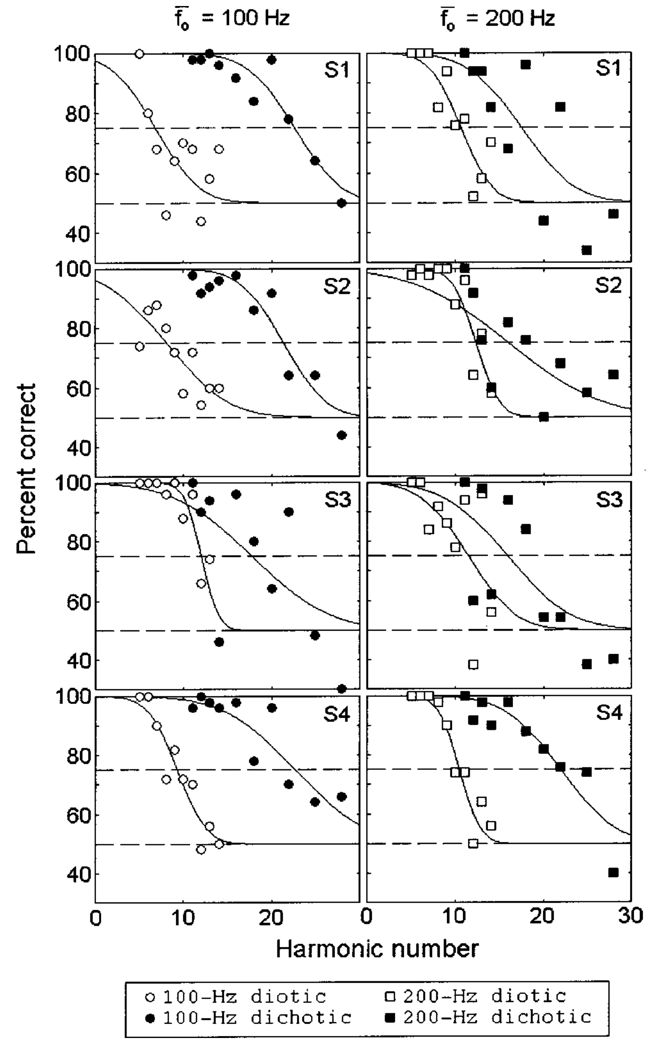

FIG. 3. Results from the individual subjects in experiment 1, showing percent correct in identifying the probe tone as higher or lower than the target tone as a function of harmonic number. Each data point represents performance over 50 stimulus trials. Each row represents results from one subject. The left column (circles) and right column (squares) show results with f_0's of 100 and 200 Hz, respectively. The solid curves respresent best fits of the erfc function [Eq. (1)] to the individual data. The upper and lower dashed lines in each plot represent 75% and 50% correct, respectively.

$1.065\,\overline{f_0}$. Randomizing Δf was intended to prevent the listener from correctly identifying the frequency relationship without actually hearing out the target tone by memorizing the frequency relationship between the comparison tone and the complex's f_0. Testing a large number of target harmonics (ten per condition) and randomizing f_0 further prevented this type of alternative cue.[3]

Each subject began with a training phase, where runs rotated through the four conditions, during which feedback was provided. Training continued until a subject was reliably obtaining nearly 100% correct for the lowest harmonic tested in each condition. The training period varied across subjects from 15 min to 2 h. During the data collection phase, feedback was not provided.

C. Results

Figure 2 shows the mean data. The error bars denote ± 1 standard error of the mean performance across all listeners. Although there was significant variability in performance across subjects, a systematic trend is clear in the data. Percent correct generally decreases with increasing harmonic number, with the 75% correct point corresponding roughly to the 10th harmonic in the diotic conditions, and to the 20th harmonic in the dichotic conditions. For each condition, the pooled data from all subjects were fit (solid lines in Fig. 2) to a complementary error function (erfc) bound to 50% and 100% correct at the extremes.[4] The nonlinear least squares Gauss–Newton method was used to fit the data to Eq. (1) with two free parameters (n_0 and w). The estimated n_0 was taken to be the estimated limit of harmonic resolvability, in accordance with the methods of Plomp (1964). Judgments of the goodness of fit were based on a 95% confidence interval ($\pm 2\sigma$) measure of uncertainty in the n_0 estimate. The values obtained for the estimated limits of resolvability and 95% confidence interval, $n_0 \pm 2\sigma$, for the pooled data were: 9.34 $\pm$1.03 (diotic 100 Hz), 21.18$\pm$1.65 (dichotic 100 Hz), 11.20$\pm$0.74 (diotic 200 Hz), and 17.73$\pm$1.91 (dichotic 200 Hz).

Figure 3 shows the individual data. The left column shows data from the 100-Hz $\overline{f_0}$ and the right column shows data from the 200-Hz $\overline{f_0}$. There was considerable intersubject variability in performance, as well as certain nonmonotonic trends within individual subjects. One subject (S2) had difficulty hearing out even the lowest harmonics in the 100-Hz diotic condition. Two subjects (S1 and S3) showed nonmonotonicities in the diotic conditions near the 12th harmonic. In the dichotic conditions, large nonmonotonicities were exhibited by one subject (S3) at the 100- and 200-Hz $\overline{f_0}$'s, and by two others (S1 and S2) at the 200-Hz $\overline{f_0}$. For these subjects, performance decreased below 75% in the vicinity of the 12th to 16th harmonics, and then increased before once again dropping below 75% for higher harmonics. The nonmonotonicities in the diotic and dichotic conditions in the vicinity of the 12th and 14th harmonics are also present in the mean data (Fig. 2).

Individual subject data in each condition were fit to the erfc function [Eq. (1)]. Fits ranged from good for subjects and conditions where the psychometric function exhibited few nonmonotonicities (e.g., subject S4, diotic 200 Hz, $2\sigma = 0.71$ harmonics), to extremely poor for subjects and conditions where the psychometric function exhibited many nonmonotonicities (e.g., subject S3, dichotic 100 Hz, $2\sigma = 6.67$ harmonics).

D. Discussion

Five aspects of the results merit attention. First, roughly twice as many harmonics can be heard out in the dichotic conditions as in the diotic conditions. This is the most important result of the experiment, as it verifies the central assumption for experiment 2, that only processes before the combination of binaural information limit harmonic resolvability.

Second, our estimates of the limits of harmonic resolvability in the diotic conditions are greater than those reported by Plomp (1964). Our results indicate that the first 9 to 11 harmonics of a complex for $\overline{f_0}$'s of 100 and 200 Hz are peripherally resolved. This estimate closely matches the indirect estimate of the limits of harmonic resolvability (Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994), where the lowest harmonic present must be the 10th or below in order to yield small f_0 DLs. This indicates that enough harmonics are peripherally resolved to account for the limits of good f_0 discrimination, thereby resolving the apparent discrepancy between direct and indirect measures of resolvability (Shackleton and Carlyon, 1994). A caveat to this conclusion is that the "enhancement" effect (see footnote 1) may have helped to overcome some nonperipheral limitation to harmonic resolvability that occurs before the detection of pitch. Therefore, in the absence of "enhancement," all of these peripherally resolved harmonics might not be available to the pitch detector. Also, this is an operational definition of resolvability, which depends on the 3.5%–5.0% Δf used in this experiment. A smaller Δf may have yielded a lower estimate of the number of resolved harmonics.

Third, there was some indication of more resolved harmonics for the 200-Hz than the 100-Hz f_0, consistent with results of Shera *et al.* (2002) indicating that the cochlear filter bandwidths relative to CF decrease with increasing absolute frequency at low signal levels. Nevertheless, this difference was small, indicating that harmonic number largely determines resolvability. The limited range of f_0's used in this study prevents a comparison with the effects of f_0 reported by Plomp (1964), where for f_0's greater than 200 Hz, the number of resolved harmonics decreased with increasing f_0.

Fourth, some subjects experienced difficulties with even low-frequency harmonics, or displayed nonmonotonic psychometric functions. For example, for subject S2 at the 200-Hz f_0 and subject S3 at both f_0's, the initial drop below 75% correct performance in the dichotic conditions occurred at a similar harmonic number as in the diotic conditions. This suggests that there may be some central limitation on resolution for these subjects and conditions that operates on both diotic and dichotic complexes. However, for all subjects, harmonics above the 14th are well resolved under dichotic presentation, and any central limitation of harmonic resolvability seems to appear only near the 14th harmonic.

Fifth, the estimate of n_0 in the dichotic 200-Hz condition had a large 95% confidence interval ($\pm 10.8\%$), consistent with the poor fit apparent in a visual inspection of the data. Given the high range of pure tone frequencies presented in this condition, this large uncertainty may reflect absolute frequency effects. However, even at the highest frequencies tested (5.6 kHz), the minimum Δf we used (3.5%) is still greater than the 0.5% obtained for similar frequency long-duration tones in quiet (Moore, 1973). Although the 60 dB SPL tones used in the Moore (1973) study are not comparable to the 15 dB SL tones used in this study, Hoekstra (1979) showed that a reduction from moderate (40 dB) to low (15 dB) SLs increased DLs for a 2-kHz pure tone by less than a factor of 2. This suggests that the variable results found at these very high frequencies cannot be ascribed solely to the coding limitations of individual components.

III. EXPERIMENT 2: FUNDAMENTAL FREQUENCY DIFFERENCE LIMENS

A. Rationale

In experiment 2 we measured f_0 DLs as a function of the lowest harmonic number present for diotic and dichotic harmonic complexes. If good discrimination ability were dependent on the presence of resolved harmonics *per se*, the auditory system should be able to utilize the information provided by the additional resolved harmonics available under dichotic presentation, such that the order of magnitude increase in f_0 DLs (Houtsma and Smurzynski, 1990) would occur at twice the harmonic number as compared to diotic presentation. Alternatively, if good discrimination ability were dependent only on the presence of low-numbered harmonics, regardless of resolvability, the additional resolved harmonics should provide no benefit, such that the increase in f_0 DLs would occur at the same lowest harmonic number in both dichotic and diotic conditions.

In order to determine if the increase in f_0 DLs is due to absolute or relative frequency effects, we performed the measurements at two different f_0's (100 and 200 Hz). Based on the results of Shackleton and Carlyon (1994), suggesting that the DL shift is due to relative frequency effects (i.e., the presence or absence of resolved harmonics), we expect that the DL shift should occur at approximately the same harmonic number for both f_0's. Alternatively, if the DL shift were mainly due to absolute frequency effects as implied by many temporal pitch models, then the DL shift should occur at about the same absolute frequency, or twice the harmonic number for the 100-Hz f_0 as compared to the 200-Hz f_0 conditions. While we measured f_0 DLs with harmonics in random phase in order to allow a direct comparison with the harmonic resolvability data of experiment 1, we also performed the measurements with harmonics in sine phase to allow a more direct comparison with earlier data.

B. Methods

Stimuli were 500-ms (including 30-ms Hanning window rise and fall) harmonic complexes with 12 successive components. Each component was presented at 10 dB SL in $\mathrm{UMN_m}$ background noise (see experiment 1). This low level was used to prevent the detection of combination tones. Stimuli were presented diotically and dichotically with f_0's of 100 and 200 Hz, in sine phase and random phase, for a total of eight conditions. Discrimination thresholds were estimated for eight normal-hearing subjects. Four subjects (ages 18–24, two female), including the first author, participated in the sine-phase conditions. Two had also participated in experiment 1. Four new subjects (ages 18–24, one female) participated in the random-phase conditions. The setup for stimulus delivery was the same as in experiment 1.

Fundamental frequency DLs as a function of the complex's average lowest harmonic number ($\overline{N}$) were estimated via a three-alternative forced-choice, two-down, one-up adaptive algorithm tracking the 70.7% correct point (Levitt, 1971). The f_0 difference (Δf_0) was initially set to 10% of the f_0. The starting step size was 2% of the f_0, decreasing to 0.5% after the first two reversals, and then to 0.2% after the next two reversals. The f_0 DL was estimated as the average of the Δf_0's at the remaining six reversal points.

Two of the intervals contained harmonic complexes with a base $f_0(f_{0,\mathrm{base}})$, while one interval contained a complex with a higher $f_0(f_{0,\mathrm{base}}+\Delta f_0)$. The task was to identify the interval with the higher f_0. Subjects were informed that two of the intervals had the same pitch, and one had a higher pitch, and were asked to identify the interval with the higher pitch. In order to prevent subjects from basing their judgments on the frequency of the lowest harmonic, the lowest harmonic number (N) was roved from interval to interval, such that in the three intervals it was $\overline{N}-1$, $\overline{N}$, and $\overline{N}+1$, in random order. The highest harmonic number was also roved, such that 12 components were presented in each stimulus interval. For the dichotic conditions, odd and even components were presented randomly to the left or right ear on a trial-by-trial basis. Feedback was provided after each trial. Subjects were informed that there were different sound qualities that varied from interval to interval. They were told to ignore the timbre ("treble/bass quality") of the sounds, as responses based on timbre would result in incorrect answers, and to respond based solely on the pitch. Fundamental frequency discrimination was tested for $\overline{N}=3$, 6, 9, 12, 15, 18 and 24 in all eight conditions.

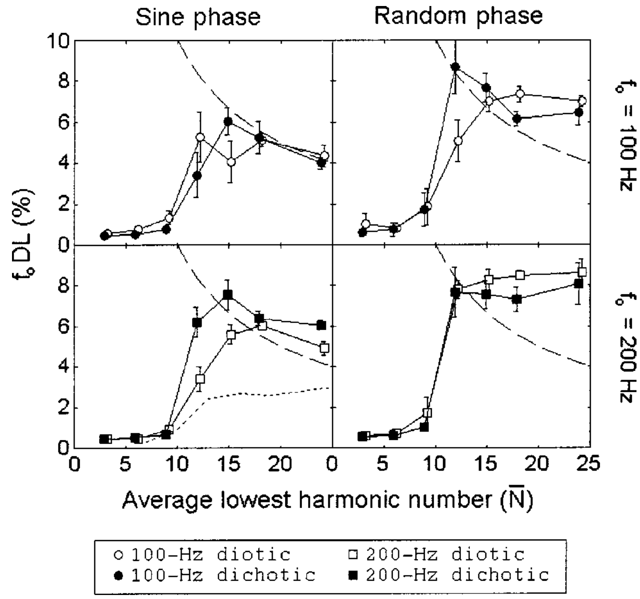

FIG. 4. Mean results from experiment 2. Each data point represents the mean f_0 DL (%) across four subjects; error bars denote plus and minus one standard error of the mean. The long-dashed curves show the limit of performance based only on the lower spectral edge of the complexes (see text for details). The short-dashed curve in the lower left panel shows data from Houtmsa and Smurzynski (1990) for a monotic complex with a 200-Hz f_0.

Each subject went through a training phase of at least 1 h, which continued until performance was no longer showing consistent improvement. During the measurement phase, four adaptive runs were made per subject, for each value of $\overline{N}$ in each condition, and the estimated f_0 DL for a subject was taken as the mean of these four estimates. If the standard deviation across the last six reversals points in any one run was greater than 0.8%, the data for that run was excluded and the run was repeated at the end of the experiment.

C. Results

Figure 4 shows the estimated f_0 DLs (expressed as a percentage of the f_0) as a function of $\overline{N}$. Each data point represents the arithmetic mean and the error bars represent $\pm$ the standard error across the mean f_0 DLs measured for four subjects. The central finding of this study is that the dramatic increase in f_0 DLs occurs at the same $\overline{N}$ under diotic and dichotic presentation. Furthermore, this increase occurs at the same $\overline{N}$ at both f_0's.

To investigate other trends in the data, an analysis of variance (ANOVA) with three within-subject factors [f_0, $\overline{N}$ and mode of presentation (diotic or dichotic)] and one between-subject factor (phase) was conducted. While the f_0 DL measurement used f_0 steps on a linear frequency scale in accordance with the methods of Houtsma and Smurzynski (1990), the statistical analysis was performed with logarithmically (log) transformed data, in an attempt to satisfy the uniform variance assumption. Only the following main effects and interactions were found to be significant ($p<0.05$). There was a main effect of $\overline{N}$ [$F(1,6)=179.5$, $p<0.0001$], two-way interactions between f_0 and $\overline{N}$ [$F(1,6)=5.60$, $p<0.0005$] and between f_0 and mode of presenta-

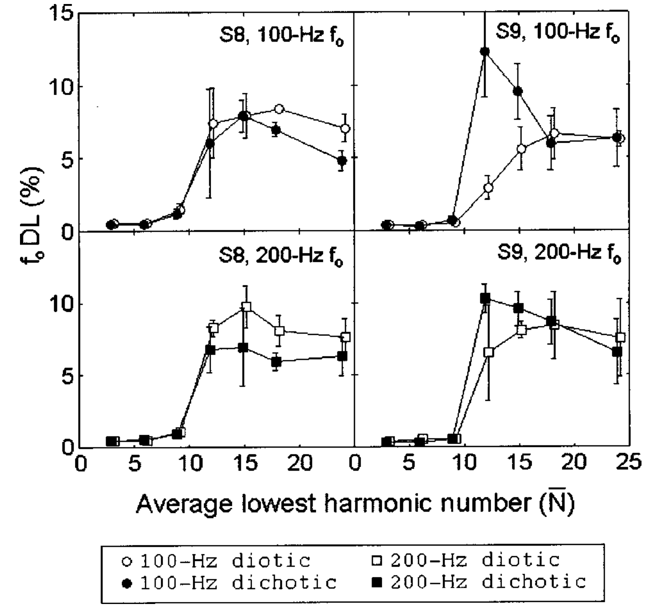

FIG. 5. Individual results from experiment 2 for two sample subjects. Error bars show plus and minus one standard deviation across four stimulus trials. Subject S9 (right column) shows larger f_0 DLs under dichotic presentation for $\bar{N}=12$ and 15. Subject S8 (left column) shows smaller f_0 DLs under dichotic presentation for $\bar{N}=18$ and 24.

tion $[F(1,6)=6.8, p<0.05]$, a three-way interaction between mode of presentation, phase and f_0 $[F(1,6)=8.32, p<0.05]$ and a four-way interaction between all factors $[F(6,36)=5.84, p<0.0005]$.

The significant four-way interaction suggests caution in interpreting main effects and low-order interactions. Nevertheless, the ANOVA supports two trends in the data concerning $\bar{N}$ and f_0. First, the main effect of $\bar{N}$ clearly reflects the result that good performance in f_0 discrimination requires $\bar{N}\leqslant 9$. Second, the two-way interaction between f_0 and $\bar{N}$ does not reflect an absolute frequency effect on the frequency of the increase in f_0 DLs, since this transition occurs at the same $\bar{N}$ for both f_0's. Rather, this interaction probably reflects an absolute frequency effect for complexes with $\bar{N}>9$, where larger f_0 DLs are seen for the 200-Hz f_0 as compared to the 100-Hz f_0. Interpreting the effects of mode of presentation and phase requires a closer examination of the data.

The significant higher-order interactions probably reflect the result that dichotic f_0 DLs were somewhat higher or lower than diotic f_0 DLs depending on f_0, phase and $\bar{N}$. Two trends in the difference between f_0 DLs measured under dichotic versus diotic presentation were apparent in the data. The first trend was that dichotic f_0 DLs were larger than diotic f_0 DLs presentation at $\bar{N}=12$ or $\bar{N}=15$ for all combinations of f_0 and phase except for the 200-Hz random phase case. This trend will be addressed further in conjunction with results of experiment 3. The second, less apparent, trend was that dichotic f_0 DLs were slightly smaller for $\bar{N}=18$ and $\bar{N}=24$ at both f_0's. Differences between diotic and dichotic f_0 DLs were seen in some subjects, but not in others. Figure 5 shows mean DLs for two sample subjects who participated in the random-phase conditions. At one extreme, subject S9 (right column) shows larger DLs under dichotic presentation near $\bar{N}=12$ for both f_0's. Of the 16 combinations of subject and f_0, seven showed larger dichotic DLs near $\bar{N}=12$ (four of eight in sine-phase, three of eight in random-phase). At the other extreme, subject S8 (left column) shows larger DLs under diotic presentation at $\bar{N}=18$ and 24. While no subjects showed larger diotic DLs near $\bar{N}=18$ for sine-phase stimuli, two did for random-phase stimuli.

The results of Houtsma and Smurzynski (1990) suggested that the phase relationship between harmonics affected the f_0 DLs for high-order, but not low-order harmonics. While this trend also appears in our data, the ANOVA indicated no significant main effect of phase or two-way interactions between phase and any other factor $(p>0.05)$. Although the f_0 DLs appear larger in the random phase conditions for $\bar{N}>9$, this difference is not statistically significant for the logarithmically transformed data. The lack of a significant phase effect in our data may be due to the fact that phase was a between-subjects factor, giving the test less statistical power than if random and sine phase complexes had been tested in the same subjects.

Another possibility is that even though the lowest harmonic number was roved from interval to interval, for large Δf_0's it is possible for listeners to achieve above chance performance without extracting f_0 information. Phase effects may not be present in any condition where f_0 information was not used to perform the task. Both in our 3IFC study and in the Houtsma and Smurzynski (1990) study, if the listener were to base their answer on the frequency of the lowest harmonic present in each interval (or the low-frequency edge of the excitation pattern for unresolved complexes), they would achieve 66.7% correct (near the 70.7% correct point approximated by the two-up, one-down adaptive procedure) if $\Delta f_0/f_0>1/\bar{N}$. Any data point falling above the DL $=(100/\bar{N})\%$ dashed line in Fig. 4 could reflect responses based on the "lowest harmonic" cue, rather than f_0 extraction. Performance is slightly worse than the "lowest harmonic" prediction for $\bar{N}=18$ and 24 in the sine-phase conditions, and much worse for $\bar{N}>12$ in all random-phase conditions. Thus in this study, listeners may be using lowest harmonic cues, rather than f_0 pitch cues, to perform f_0 discrimination for complexes with high $\bar{N}$, especially when the components are in random phase. In the Houtsma and Smurzynski (1990) study, f_0 DLs are much smaller than the lowest harmonic cue prediction, and therefore most likely reflect actual f_0 discrimination performance.

To look for possible phase effects, Scheffe *post-hoc* tests compared sine- and random-phase data for the four combinations of f_0 and mode of presentation for $\bar{N}=12$, which is above the resolved harmonic region, but below the region where the "lowest harmonic" cue may have influenced the results. Results indicate that f_0 DLs were significantly different $(p<0.05)$ in the 100-Hz dichotic and 200-Hz diotic conditions, providing some weak evidence for the presence of phase effects in these conditions.

D. Discussion

The fact that the transition from small to large f_0 DLs occurs at the same $\overline{N}$ under diotic and dichotic presentation indicates that the auditory system is unable to utilize the information provided by the additional resolved harmonics in the dichotic case for f_0 discrimination. While two subjects did show slightly smaller f_0 DLs for dichotic complexes than for diotic complexes when $\overline{N} \geqslant 18$, the f_0 DLs (around 6%) are still much larger than those found for lower numbered harmonics. This supports the hypothesis that good f_0 discrimination is not limited by harmonic resolvability, but by harmonic number, regardless of resolvability. This result also indicates that subjects cannot ignore the input from one ear in performing the f_0 discrimination task. Remember that the ear with the even harmonics contains consecutive harmonics of 2 f_0, with a lowest harmonic around $\overline{N}/2$. If subjects were able to ignore the ear with odd harmonics, we would expect the transition between good and poor f_0 discrimination to occur at twice the average lowest harmonic number, i.e., around $\overline{N}=20$. Thus, this result is consistent with the idea that pitch is derived from a combined ''central spectrum'' (Zurek, 1979) that prevents an independent pitch percept derived from the input to one ear. Note, however, that the odd and even harmonics were presented to left and right ears at random in each trial, making it impossible for the listener to know which ear to ignore. It is possible that if odd and even harmonics were presented consistently to the same ears, subjects may have been able to learn to ignore the input from the ear with odd harmonics.

The transition from small to large f_0 DLs occurs at the same $\overline{N}$ at both f_0's, consistent with the results of Kaernbach and Bering (2001). This confirms our expectation (Shackleton and Carlyon, 1994) that the dramatic increase in f_0 DLs is due to a relative frequency effect that depends more on harmonic number than on an absolute frequency effect, such as the roll-off of phase-locking with increasing absolute frequency. Nevertheless, effects of absolute frequency are also present, in that the f_0 DLs for $\overline{N}>9$ are greater for the 200-Hz f_0 than for the 100-Hz f_0. These absolute frequency effects may be related to phase locking, where the additional information available from phase locking to the fine structure at a lower absolute frequency region in the 100-Hz condition aided f_0 discrimination. Also, because we tested only f_0's of 100 and 200 Hz, we did not observe the absolute frequency effects reported in other studies where the f_0 DL transition occurs at a lower $\overline{N}$ for f_0's below 100 Hz and above 200 Hz (Ritsma, 1962; Krumbholz *et al.*, 2000; Pressnitzer *et al.*, 2001).

For the diotic 200-Hz sine-phase condition, f_0 DLs for complexes with $\overline{N}>10$ are approximately twice as large as those of the monotic 200-Hz sine-phase results of Houtsma and Smurzynski (1990), depicted as a dashed line in the lower left panel of Fig. 4, although the transition from small to large f_0 DLs occurs at the same $\overline{N}$ in both studies. The difference in DL between this and the earlier study can be probably explained in terms of differences in sensation level. Hoekstra (1979) showed that an increase in sensation level from the 10 dB used in our study to the 20 dB used in the Houtsma and Smurzynski (1990) study decreased f_0 DLs for harmonic complexes by a factor of 2 to 4, depending on f_0 and subject.

IV. EXPERIMENT 3: PERCEIVED PITCH OF DICHOTIC STIMULI

A. Rationale

Flanagan and Guttman (1960) investigated the pitch of same- and alternating-polarity click trains. A same-polarity click train has a click rate equal to the f_0, and a spectrum consisting of all the harmonics of f_0, whereas an alternating-polarity click train has a click rate that is $2f_0$, and a spectrum consisting of only the odd harmonics of the f_0. According to Flanagan and Guttman (1960), stimuli with $f_0<150$ Hz elicit a pitch corresponding to the click rate, regardless of polarity, while stimuli with $f_0>150$ Hz elicit a pitch corresponding to the f_0. This result is consistent with a two-mechanism model of pitch perception. Click trains with a high f_0 that contain resolved components in the absolute frequency dominance region for pitch (Plomp, 1967) yield a pitch at the f_0, consistent with a mechanism that extracts pitch from spectral cues. Click trains with a low f_0 that contain only unresolved components in the dominance region yield a pitch consistent with a mechanism that extracts pitch from peaks in the temporal envelope of the waveform. The temporal envelope of the alternating polarity click train repeats at the difference frequency between components of 2 f_0, whereas the waveform of the same polarity click train repeats at the f_0.

Experiment 3 estimated the perceived pitch of the dichotic stimuli used in experiment 2. If, as suggested by the results of experiment 2, the individual resolved components above the 10th harmonic are not used in f_0 discrimination, then the pitch of dichotic complexes with $\overline{N}>10$ may be derived from the repetition rate of the temporal envelope. If so, these complexes should yield a perceived pitch at 2 f_0, consistent with the peripheral difference frequency between adjacent components. Alternatively, the central pitch mechanism may be able to make some, but poor, use of the higher-order resolved harmonics. If so, these dichotic stimuli should yield a pitch at the f_0 derived from the combined central spectrum, but with the poor f_0 discrimination performance seen in experiment 2.

B. Methods

Assuming that listeners would only perceive a pitch at the f_0 or at 2 f_0 for their alternating-phase stimuli, Shackleton and Carlyon (1994) asked listeners to identify which of two sine-phase stimuli, with fundamental frequencies equal to the f_0 or to 2 f_0 of the alternating-phase stimulus, most closely matched each alternating-phase stimulus. Similarly, we assumed that our dichotic stimuli would yield perceived pitches corresponding to either the f_0, consistent with spectral cues, or to 2 f_0, consistent with monaural temporal envelope cues. However, we used a different experimental paradigm. Subjects compared the pitch of a dichotic stimulus with that of a diotic stimulus, where the f_0 of the diotic

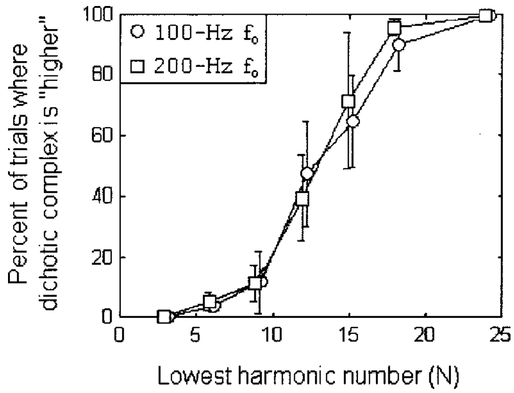

FIG. 6. Mean results from experiment 3, showing the percentage of trials where subjects reported a dichotic complex to have a higher pitch than a diotic complex with f_0 a factor of $\sqrt{2}$ higher. Error bars indicate plus or minus one standard error of the mean across the individual subjects. For lowest harmonic number $N<12$, subjects nearly always identified the diotic complex as "higher;" for $N>15$ subjects nearly always identified the dichotic complex as "higher." The transition from 0% to 100% occurs at approximately the same harmonic number for both f_0's tested.

stimulus was a half-octave (a factor of $\sqrt{2}$) higher than that of the dichotic stimulus. We assumed that the diotic sine-phase stimulus yielded a perceived pitch near its f_0. Thus, if the dichotic stimulus was judged as "higher" we assumed that the subject perceived its pitch to be $2f_0$. Similarly, if the dichotic stimulus was judged "lower," we assumed the subject perceived its pitch to be the f_0.

The dichotic stimuli were sine-phase complexes identical to those described in experiment 2. The diotic stimuli were sine-phase harmonic complexes with f_0 half an octave above the f_0 of the dichotic stimulus in the same trial, with harmonics chosen such that the bandwidth was limited to that of the dichotic stimulus. The diotic and dichotic stimuli were presented randomly in the first and second intervals and the subject was asked to identify the "higher" interval. Lowest harmonic number was not roved from interval to interval. Each run consisted of seven trials for each of the seven average lowest harmonic numbers tested in experiment 2, for a total of 49 trials. Twelve runs were performed at both the 100- and 200-Hz f_0's, such that each dichotic complex was presented a total 84 times per subject. To acquaint subjects with the task, they underwent a short (15 min) training session during which they were required to identify the higher of two pure tones separated by $\frac{1}{2}$ octave. Four subjects (ages 18–24, one female) took part in this experiment. Three had already participated in experiment 1 or 2. The setup for stimulus delivery was identical to that described in experiment 1.

C. Results

The results shown in Fig. 6 clearly indicate that subjects perceived a pitch lower than $\sqrt{2}$ times the f_0 for dichotic stimuli with a low lowest harmonic number and a pitch higher than $\sqrt{2}$ times the f_0 for dichotic stimuli with a high lowest harmonic number. The transition between the two pitch percepts occurred between lowest harmonic numbers 9 and 18, roughly the same region as was seen for the f_0 DL shift in experiment 2. If our assumption that listeners always perceive a pitch corresponding to either the f_0 or $2f_0$ holds, then listeners are perceiving a pitch corresponding to the f_0 for complexes containing harmonics lower than the 10th, consistent with spectral cues, and a pitch corresponding to $2f_0$ for complexes containing only harmonics above the 15th, consistent with temporal envelope cues. In between, the pitch appears to be ambiguous. Further testing would be necessary to determine if this ambiguity reflects two simultaneous pitches (at the f_0 and at $2f_0$) deriving from different mechanisms.

D. Discussion

The values of N that yielded an ambiguous pitch in this experiment correspond well to the values of $\overline{N}$ that yielded elevated f_0 DLs under dichotic presentation in experiment 2. This suggests that neither a mechanism that operates on resolved harmonics nor a mechanism that extracts the pitch from the temporal envelope responds well to dichotic stimuli in this region. Since approximately 20 harmonics are resolved under dichotic presentation (experiment 1), listeners had difficulty extracting the f_0 from these high-order, but resolved harmonics. Since dichotic complexes have fewer components falling within an auditory filter, the resulting temporal envelopes will be less modulated than the envelopes associated with the diotic stimuli, reducing the effectiveness of the envelope as a pitch cue.

The data show that listeners nearly always perceived a pitch near the f_0 for $N<12$. This result is in conflict with the results of Hall and Soderquist (1975), where subjects reported two pitches, one at each f_0, when three successive components each of a 200- and a 400-Hz f_0 were presented to opposite ears. The larger number of harmonics presented in the current study (six to each ear) may have encouraged the fusion of binaural information in processing pitch.

V. GENERAL DISCUSSION

A. Absolute or relative frequency?

The transitions from strong to weak pitch salience in experiment 2, and from a perceived pitch of the f_0 to $2f_0$ in experiment 3, occur at approximately the same lowest harmonic number for both the 100- and 200-Hz f_0's. These results are consistent with the idea that relative frequency relationships, such as those that govern harmonic resolvability, underlie the different pitch percepts associated with complexes containing low- and high-order harmonics (Houtsma and Smurzynski, 1990; Carlyon and Shackleton, 1994; Shackleton and Carlyon, 1994; Kaernbach and Bering, 2001). If the change in pitch salience were due to absolute frequency effects, as suggested by autocorrelation models (Cariani and Delgutte, 1996; Carlyon, 1998), the transition should have occurred at the same absolute frequency, and not the same harmonic number, for the two f_0's.

B. Resolvability or harmonic number?

Taken together, the results from the experiments demonstrate an interrelationship between harmonic number, resolvability, and pitch. Specifically, the region around the 10th harmonic appears to be important in defining transitions in harmonic resolvability, f_0 discrimination, and pitch height, at least for the f_0 range between 100 and 200 Hz. First, experiment 1 showed that for diotic stimuli approximately the first ten harmonics are resolved, while higher harmonics are unresolved. Second, consistent with Houtsma and Smurzynski (1990), experiment 2 showed that small f_0 DLs require the presence of harmonics below the 10th. Third, experiment 3 showed that a perceived pitch associated with the f_0 of the combined binaural spectrum requires the presence of harmonics below the 10th. Taken together, these three observations are consistent with the idea that complexes containing resolved harmonics below the 10th yield fundamentally different pitch percepts from those containing only harmonics above the 10th.

Consistent with earlier data from two-tone complexes (Houtsma and Goldstein, 1972; Arehart and Burns, 1999), the interpretation that harmonic resolution *per se* is responsible for the changes in pitch perception is not supported by the comparison of the diotic and dichotic results in experiments 2 and 3. The additional resolved harmonics in the dichotic case yield neither small f_0 DLs in experiment 2, nor a pitch match consistent with extraction of cues from a centrally combined spectrum in experiment 3, both of which would be expected if the shift from a salient spectral pitch to a weak temporal pitch were due to harmonics becoming unresolved. For example, although a dichotic stimulus with $N = 15$ contains resolved components, it yields poor f_0 discrimination performance and an ambiguous pitch percept. Thus, harmonic number, regardless of resolvability, seems to underlie the changes in pitch perception.

C. Implications for pitch theories

''Harmonic template'' pitch theories propose that a pitch detection mechanism codes the individual frequencies of the peripherally resolved partials and compares them to an internally stored template to derive a pitch at the f_0 (e.g., Goldstein, 1973; Terhardt, 1974, 1979). The failure of these models to explain how periodicity information is extracted from complexes containing only high-order harmonics has driven an opposing view that f_0 extraction is performed by a single autocorrelation or similar mechanism that operates on all harmonics, regardless of resolvability (Licklider, 1951, 1959; Meddis and Hewitt, 1991a, b; Meddis and O'Mard, 1997; de Cheveigné, 1998). Meddis and O'Mard (1997) have claimed that their model accounts for the different pitch percepts associated with resolved and unresolved harmonic complexes, due to the inherent differences in the result of the autocorrelation calculation for resolved versus unresolved harmonics, although the validity of this claim has been put into doubt by further analysis of their model (Carlyon, 1998). Alternatively, several studies have suggested that pitch may be processed via two different mechanisms, a harmonic template mechanism operating on resolved harmonics, and a separate mechanism operating on the temporal envelope resulting from unresolved harmonics (Houtsma and Smurzynski, 1990; Carlyon and Shackleton, 1994; Shackleton and Carlyon, 1994; Steinschneider *et al.*, 1998; Grimault *et al.*, 2002).

The results of experiment 2 argue against a pitch processing mechanism that responds inherently differently to unresolved versus resolved harmonics. With such a mechanism, we would expect f_0 discrimination performance to improve when normally unresolved harmonics are artificially resolved under dichotic presentation, whereas experiment 2 showed that f_0 discrimination performance was the same or worse in the dichotic conditions. Therefore, any theory of pitch perception must account for relative frequency effects without relying on harmonic resolvability.

''Temporal'' theories could account for this relative frequency effect if the autocorrelation in each channel were constrained to be sensitive to a limited range of periodicities relative to the inverse of the channel's CF, thereby limiting the range of harmonic numbers contributing to the pitch percept (Moore, 1982). This modification would also need to somehow account for a pitch derived from the temporal envelope for complexes containing only high-order components. If this requirement could be met, the modified theory would be consistent with the ambiguous pitch and elevated f_0 DLs seen for dichotic complexes with $N = 12$ and $N = 15$, which have relatively ineffective envelope cues (see Sec. IV D).

''Place'' theories could account for this relative frequency effect if the templates that derive the pitch from low-order harmonics were constrained to consist of only those harmonics that are *normally* resolved. This is consistent both with the idea of harmonic templates learned from exposure to harmonic sounds (Terhardt, 1974) and the more recent proposal that templates for low-order harmonics may emerge from any form of wideband stimulation (Shamma and Klein, 2000). With this constraint, even though artificially resolved harmonics (above the 10th and up to the 20th partial) are available under dichotic presentation, the pitch processing mechanism will be unable to utilize these additional resolved harmonics since no template will have developed to match them.

Even with this constraint, ''harmonic template'' theories do not fully explain the results for dichotic complexes containing artificially resolved harmonics. For these stimuli, we would expect that the even harmonics in one ear would match a template corresponding to $2f_0$, yielding f_0 DLs on the order of those measured for complexes containing low-order harmonics. While ambiguous pitch matches suggest that listeners may sometimes perceive a pitch corresponding to $2f_0$ for these dichotic complexes, f_0 DLs are *larger* than those for diotic complexes with the same N. Apparently, the presence of the odd harmonics in the opposite ear has a substantial detrimental effect on f_0 discrimination.

One possible explanation for these results postulates the existence of inhibitory inputs to harmonic templates, tuned to partials of subharmonics of the f_0. Under normal circumstances, where all harmonics of a complex are present, such inhibition might be useful in preventing erroneous pitch percepts at multiples of the f_0. According to this scheme, while the resolved (mn)th partials of a complex (where m and n are integers) would facilitate a template for a pitch corresponding to n times the $f_0(nf_0)$ of the complex, the remaining resolved partials of the complex would inhibit this template. Thus, only the template for a pitch at the f_0 would respond to the stimulus, yielding a pitch percept corresponding to the f_0 and good f_0 discrimination. For dichotic complexes with $N>10$, templates for pitches corresponding to nf_0 would still be inhibited, but in this case the template for a pitch corresponding to the f_0, with a limited number of harmonics represented, would not respond to the high-order harmonics. With no template available, the pitch could only be derived from temporal cues.

Another interpretation of the results is that the pitch is extracted from a combined ''central spectrum'' representation (Zurek, 1979) that prevents an independent pitch percept derived from the input to one ear. The additional *peripherally* resolved components might not be available in the central spectrum representation used to derive pitch. Listeners may have been able to overcome this central fusion in hearing out individual harmonics in experiment 1, but not when deriving a pitch from the sum of components in experiments 2 and 3. The nonmonotonic psychometric functions seen in some subjects in experiment 1 may reflect an inability to overcome the binaural fusion even in the ''hearing out'' task.

VI. SUMMARY AND CONCLUSIONS

In experiment 1 approximately twice as many harmonics are resolved under dichotic as compared to diotic presentation, verifying that harmonic resolvability is not limited by binaural interactions. A direct estimate of the limits of harmonic resolvability indicated that approximately 9 and 11 harmonics are resolved for 100- and 200-Hz f_0's, respectively. The results from our direct measure, which minimizes nonperipheral limitations by gating the target component on and off, resolve the discrepancy between previous direct estimates that only five to eight harmonics are resolved (Plomp, 1964), and indirect estimates suggesting that approximately ten harmonics are resolved (Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994).

In experiment 2, listeners were unable to utilize the additional resolved harmonics available under dichotic presentation for f_0 discrimination. This implies that the deterioration in f_0 DLs with increasing lower cutoff frequency is due not to harmonics becoming unresolved *per se*, but instead to the increasing lowest harmonic number, regardless of resolvability. This result suggests constraints to both ''place'' and ''temporal'' models of pitch perception. For a ''harmonic template'' theory to account for the data, only those harmonics that are *normally* resolved should be represented in the templates. For an ''autocorrelation'' theory to do so, the range of periodicities to which the autocorrelation in each channel is sensitive should be CF-dependent (Moore, 1982).

The results of experiments 2 and 3 are consistent with a two-mechanism model of pitch perception (e.g., Carlyon and Shackleton, 1994). When harmonics below the 10th are present, a harmonic template mechanism is able to extract pitch from the resolved components, yielding small f_0 DLs and a pitch consistent with spectral cues. When only harmonics above the 10th are present, the auditory system relies on temporal envelope cues for pitch extraction, regardless of resolvability, yielding some ambiguous pitch percepts for dichotic complexes, and poor f_0 discrimination performance in all cases. A temporal model, constrained as described above, may nevertheless be able to account for these results within the framework of a single autocorrelation mechanism.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R01 DC 05216 and by NIH Training Grant No. 5T32 DC 00038. We thank Christophe Micheyl, Peninah Rosengard, Chris Plack, Neal Viemeister, and two anonymous reviewers for their helpful comments on previous versions of this manuscript. We also thank Chris Plack for suggesting the existence of inhibitory inputs to a 2 f_0 template, tuned to the frequencies of the odd harmonics of a complex with fundamental frequency f_0.

[1]Viemeister and Bacon (1982) found that a component whose onset is delayed relative to the remaining component produced more forward masking than when the entire complex is gated synchronously. Even if this ''enhancement'' effect can be thought of as ''amplifying'' the representation in a subset of auditory nerve fibers, this should not have any effect on peripheral resolvability, as the signal-to-noise ratio within that population would be unaffected. In fact, physiological enhancement of the response to a component of a harmonic complex with delayed onset time has been found in the cochlear nucleus (Scutt *et al.*, 1997) but not in the auditory nerve (Palmer *et al.*, 1995) of the guinea pig.

[2]The temporal waveform for several harmonics of a sine-phase complex that fall in one auditory filter is click-like, with brief peaks occurring at intervals of the f_0, separated by low-level epochs. Eliminating a spectrally unresolved harmonic component (i.e., adding it out of phase) will result in that component appearing during the low-level epochs, thereby allowing the detection of the subtracted component's frequency by ''listening in the valleys,'' or ''dip listening'' (Duifhuis, 1970). Since random-phase complexes generally have much flatter temporal envelopes and are not conducive to listening in the valleys (Alcántara and Moore, 1995), this greatly reduced the possibility of dip listening.

[3]In fact, the combined randomizations ensured that, for a given f_{comp}, the probability that the frequency of the target harmonic was higher than f_{comp} was approximately equal to the probability of it being lower (except when $f_{comp}<f_{targ}$ when the lowest target component was tested or $f_{comp}>f_{targ}$ when the highest target component was tested), so that subjects were prevented from answering correctly based only on the frequency of the comparison tone.

[4]Percent correct $$(n)=100\left[\frac{1}{2}+\frac{1}{2\sqrt{\pi}}\int_n^{\infty} e^{-[w(n'-n_0)]^2}\,dn'\right], \quad (1)$$

where n is harmonic number, n' is the harmonic number integration variable, w is a factor describing the slope of the psychometric function, and n_0 is the harmonic number that yields 75% correct.

Alcántara, J. I., and Moore, B. C. J. (**1995**). ''The identification of vowel-like harmonic complexes: Effects of component phase, level, and fundamental frequency,'' J. Acoust. Soc. Am. **97**, 3813–3824.

Arehart, K. H., and Burns, E. M. (**1999**). ''A comparison of monotic and dichotic complex-tone pitch perception in listeners with hearing loss,'' J. Acoust. Soc. Am. **106**, 993–997.

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

Carlyon, R. P. (**1998**). ''Comments on 'A unitary model of pitch perception [J. Acoust. Soc. Am. **102**, 1811–1820 (1997)],' '' J. Acoust. Soc. Am. **104**, 1118–1121.

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

de Cheveigné, A. (**1998**). ''Cancellation model of pitch perception,'' J. Acoust. Soc. Am. **103**, 1261–1271.

Duifhuis, H. (**1970**). ‘‘Audibility of high harmonics in a periodic pulse,’’ J. Acoust. Soc. Am. **48**, 888–893.
Fine, P. A., and Moore, B. C. J. (**1993**). ‘‘Frequency analysis and musical ability,’’ Music Percept. **11**, 39–53.
Flanagan, J. L., and Guttman, N. (**1960**). ‘‘On the pitch of periodic pulses,’’ J. Acoust. Soc. Am. **32**, 1308.
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.
Grimault, N., Micheyl, C., Carlyon, R. P., and Collet, L. (**2002**). ‘‘Evidence for two pitch encoding mechanisms using a selective auditory training paradigm,’’ Percept. Psychophys. **64**, 189–197.
Hall, J. W., and Soderquist, D. R. (**1975**). ‘‘Encoding and pitch strength of complex tones,’’ J. Acoust. Soc. Am. **58**, 1257–1261.
Hoekstra, A. (**1979**). ‘‘Frequency discrimination and frequency analysis in hearing,’’ Ph.D. Thesis, Institute of Audiology, University Hospital, Groningen, Netherlands.
Houtsma, A. J. M., and Goldstein, J. L. (**1972**). ‘‘The central origin of the pitch of pure tones: Evidence from musical interval recognition,’’ J. Acoust. Soc. Am. **51**, 520–529.
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.
Kaernbach, C., and Bering, C. (**2001**). ‘‘Exploring the temporal mechanism involved in the pitch of unresolved harmonics,’’ J. Acoust. Soc. Am. **110**, 1039–1048.
Krumbholz, K., Patterson, R. D., and Pressnitzer, D. (**2000**). ‘‘The lower limit of pitch as determined by rate discrimination,’’ J. Acoust. Soc. Am. **108**, 1170–1180.
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.
Licklider, J. C. R. (**1959**). ‘‘Three auditory theories,’’ in *Psychology, a Study of Science*, edited by S. Koch (McGraw–Hill, New York).
Meddis, R., and Hewitt, M. (**1991a**). ‘‘Virtual pitch and phase sensitivity studied of 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 of a computer model of the auditory periphery. II: Phase sensitivity,’’ J. Acoust. Soc. Am. **89**, 2883–2894.
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. (**1982**). *An Introduction to the Psychology of Hearing*, 2nd ed. (Academic, London), pp. 140–144.
Moore, B. C. J., and Ohgushi, K. (**1993**). ‘‘Audibility of partials in inharmonic complex tones,’’ J. Acoust. Soc. Am. **93**, 452–461.
Ohm, G. S. (**1843**). ‘‘Über die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen [On the definition of the tone and the related theory of the siren and similar tone-producing devices],’’ Ann. Phys. Chem. **59**, 513–565.
Palmer, A. R., Summerfield, Q., and Fantini, D. A. (**1995**). ‘‘Responses of auditory-nerve fibers to stimuli producing psychophysical enhancement,’’ J. Acoust. Soc. Am. **97**, 1786–1799.
Plomp, R. (**1964**). ‘‘The ear as a frequency analyzer,’’ J. Acoust. Soc. Am. **36**, 1628–1636.
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.
Pressnitzer, D., Patterson, R. D., and Krumbholz, K. (**2001**). ‘‘The lower limit of melodic pitch,’’ J. Acoust. Soc. Am. **109**, 2074–2084.
Ritsma, R. J. (**1962**). ‘‘Existence region of the tonal residue. I.,’’ J. Acoust. Soc. Am. **34**, 1224–1229.
Ritsma, R. J. (**1967**). ‘‘Frequencies dominant in the perception of the pitch of complex sounds,’’ J. Acoust. Soc. Am. **42**, 191–198.
Schmidt, S., and Zwicker, E. (**1991**). ‘‘The effect of masker spectral asymmetry on overshoot in simultaneous masking,’’ J. Acoust. Soc. Am. **89**, 1324–1330.
Scutt, M. J., Palmer, A. R., and Summerfield, A. Q. (**1997**). ‘‘Psychophysical and physiological responses to signals which are enhanced by temporal context,’’ Abstr., Assoc. Res. Otolaryngol. MidWinter Meeting.
Seebeck, A. (**1841**). ‘‘Beobachtungen über einige Bedingungen der Entstehung von Tönen [Observations on some conditions for the creation of tones],’’ Ann. Phys. Chem. **53**, 417–436.
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.
Shera, C. A., Guinan, J. J., and Oxenham, A. J. (**2002**). ‘‘Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements,’’ Proc. Natl. Acad. Sci. U.S.A. **99**, 3318–3323.
Soderquist, D. R. (**1970**). ‘‘Frequency analysis and the critical band,’’ Psychonomic Sci. **21**, 117–119.
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.
Steinschneider, M., Reser, D. H., Fishman, Y. I., Schroeder, C. E., and Arezo, J. C. (**1998**). ‘‘Click train encoding in primary auditory cortex of the awake monkey: Evidence for two mechanisms subserving pitch perception,’’ J. Acoust. Soc. Am. **104**, 2935–2955.
Terhardt, E. (**1974**). ‘‘Pitch, consonance, and harmony,’’ J. Acoust. Soc. Am. **55**, 1061–1069.
Terhardt, E. (**1979**). ‘‘Calculating virtual pitch,’’ Hear. Res. **1**, 155–182.
Viemeister, N. F., and Bacon, S. P. (**1982**). ‘‘Forward masking by enhanced components in harmonic complexes,’’ J. Acoust. Soc. Am. **71**, 1502–1507.
Weiss, T. F., and Rose, C. (**1988**). ‘‘Stages of degradation of timing information in the cochlea—a comparison of hair-cell and nerve fiber responses in the alligator lizard,’’ Hear. Res. **33**, 167–174.
Wightman, F. L. (**1973**). ‘‘The pattern-transformation model of pitch,’’ J. Acoust. Soc. Am. **54**, 407–416.
Zurek, P. M. (**1979**). ‘‘Measurements of binaural echo suppression,’’ J. Acoust. Soc. Am. **66**, 1750–1757.