The lower limit of melodic pitch

Daniel Pressnitzer,[a] Roy D. Patterson, and Katrin Krumbholz
Centre for the Neural Basis of Hearing, Department of Physiology, Downing Street, Cambridge CB2 3EG, United Kingdom

(Received 10 February 2000; revised 7 August 2000; accepted 6 February 2001)

An objective melody task was used to determine the lower limit of melodic pitch (LLMP) for harmonic complex tones. The LLMP was defined operationally as the repetition rate below which listeners could no longer recognize that one of the notes in a four-note, chromatic melody had changed by a semitone. In the first experiment, the stimuli were broadband tones with all their components in cosine phase, and the LLMP was found to be around 30 Hz. In the second experiment, the tones were filtered into bands about 1 kHz in width to determine the influence of frequency region on the LLMP. The results showed that whenever there was energy present below 800 Hz, the LLMP was still around 30 Hz. When the energy was limited to higher-frequency regions, however, the LLMP increased progressively, up to 270 Hz when the energy was restricted to the region above 3.2 kHz. In the third experiment, the phase relationship between spectral components was altered to determine whether the shape of the waveform affects the LLMP. When the envelope peak factor was reduced using the Schroeder phase relationship, the LLMP was not affected. When a secondary peak was introduced into the envelope of the stimuli by alternating the phase of successive components between two fixed values, there was a substantial reduction in the LLMP, for stimuli containing low-frequency energy. A computational auditory model that extracts pitch information with autocorrelation can reproduce all of the observed effects, provided the contribution of longer time intervals is progressively reduced by a linear weighting function that limits the mechanism to time intervals of less than about 33 ms. [DOI: 10.1121/1.1359797]

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

I. INTRODUCTION

A periodic click train produces a strong pitch when the click repetition rate is 100 Hz; however, when the rate is 10 Hz or less, there is no pitch and the individual clicks are heard as separate events. In the transition region, as the pitch percept fades away, the periodicity can still be detected as roughness, pulsation, or flutter (Guttman and Julesz, 1963; Terhardt, 1970; Warren and Bashford, 1981). Pitch differs from the other percepts inasmuch as it alone can convey information about musical intervals and thus, melodies (Plomp, 1976; Moore and Rosen, 1979; Dowling and Harwood, 1986; Houtsma, 1995; Griffiths *et al.*, 1998). The purpose of this study is to delineate the lower boundary of the region where pitch will support melodic patterns similar to those used in Western music.

In an influential study, Ritsma (1962) investigated the existence region of pitch for three-component harmonic complexes, specifically sinusoidally amplitude-modulated (SAM) tones. His data suggest that the lowest repetition rate[1] that produces a pitch is around 40 Hz for a carrier frequency of 150 Hz. In Ritsma's experiment, as the carrier frequency increases, the lower limit of pitch rises to a value of 350 Hz for a carrier frequency of 4.7 kHz. A replication of the study by Moore (1973) confirmed the basic findings. Both studies employed subjective judgments about the presence or absence of a pitch cue.

Ritsma (1971) and Ritsma and Hoekstra (1974) introduced an objective method to investigate the existence region of pitch. They measured rate discrimination threshold (RDT) across the lower boundary of the existence region defined by Ritsma (1962). They reported small RDTs for sounds that were inside the pitch region and large RDTs for sounds outside the pitch region. They concluded that the transition from a small to a large RDT revealed the limit of pitch (Ritsma and Hoekstra, 1974). Recently, in a companion paper, the correspondence between rate discrimination and the lower limit of pitch was re-examined with newer techniques (Krumbholz *et al.*, 2000). The results broadly confirmed the correspondence reported previously. Houtsma and Smurszynski (1990) have questioned the link between RDT and pitch; they showed that harmonic sounds in the region where the RDT is large, nevertheless support musical interval recognition (albeit with reduced accuracy). They suggest that the increase in RDT reflects the transition between the regions of resolved and unresolved harmonics (Shackleton and Carlyon, 1994; Carlyon and Shackleton, 1994) rather than the boundary of pitch perception.

Goldstein (2000) has reviewed evidence showing that the pitch of harmonic complex tones may, in fact, have different components or modes. Schouten (1940) introduced the term ''residue'' to characterize the pitch sensation produced by unresolved frequency components. de Boer (1976) proposed to generalize the use of the term residue to unresolved and resolved components, as the latter seemed to dominate the former in pitch perception (Plomp, 1967). Other terms to describe the pitch sensation corresponding to the missing fundamental of harmonic complex tones include ''low

[a]Author to whom correspondence should be addressed; IRCAM-CNRS, 1 place Stravinsky, 75004 Paris, France; electronic mail: pressnit@ircam.fr

pitch'' (Smoorenburg, 1970) or ''periodicity pitch'' (Terhardt, 1970). Guttman and Pruzansky (1962) suggested that in the case of complex harmonic tones, we should further distinguish between ''pitch'' as described in the American National Standard (the sensation that enables us to order notes on a scale from low to high), and ''musical pitch.'' By ''musical pitch'' they meant a sensation that can be used to convey musical values like diatonic intervals. They reported lower limits of 19 Hz using a subjective criterion like that of Ritsma (1962), and 60 Hz using an objective, octave-matching task. Unfortunately, their octave-matching task requires judgments that are difficult for listeners that are not musically trained. Patterson *et al.* (1983) introduced a melody-change task that involves pitch in a musical context but is much easier to perform. They used the technique to investigate the duration that complex tones need to support pitch.

In this paper, the melody-change task is adapted to determine the lower limit of melodic pitch (LLMP). The term ''melodic'' is introduced to emphasize that the experimental task provides an operational definition of pitch. Links between pitch, musical pitch, and melodic pitch will be discussed in the latter sections of the paper. In the LLMP task, listeners are required to detect a semitone change in a four-note random melody based on the chromatic scale. The range is restricted to 4 semitones; this enables the production of a sufficient number of random melodies while focusing on a limited range of repetition rates. The LLMP task has several advantages. First, it has face validity; melodies are the most fundamental elements of Western music and the semitone is the basic pitch interval of the Western chromatic scale. Second, the task is easy to perform; both musical and nonmusical listeners can perform the task whenever the notes produce a clear pitch. Parncutt and Cohen (1995) have shown that with a semitone change and an eight-note melody task, listeners reach asymptotic performance irrespective of musical education. The same is not the case for interval recognition and labelling tasks (Guttman and Pruzansky, 1962; Houtsma and Goldstein, 1972). Finally, the fact that the melodies are chosen at random minimizes the potential to use contour and knowledge-based cues that can play a part in the recognition of familiar melodies (Dowling and Fujitani, 1971; Patterson *et al.*, 1983).

II. EXPERIMENT I: THE LLMP FOR A BROADBAND, HARMONIC COMPLEX TONE

A. Rationale

The aim of the first experiment was to measure the LLMP for click trains; that is, the lowest repetition rate that enabled listeners to perform a melody task using broadband, harmonic complex tones. Click trains produce strong pitch percepts when the rate is as low as 100 Hz, the pitch of a low male voice. Below this, as the rate decreases, the pitch becomes weaker and eventually disappears to give way to the perception of isolated clicks. At this point, the melody task should become impossible because it requires the perception of pitch. By using a click train, the repetition rate can be lowered continuously while presenting energy across a large frequency range. This avoids confounding factors associated with the steep rise of the audiogram at low frequencies, which are problematic when studying pitch with low-frequency sinusoids. Also, the pitch strength of the three-component tones used by Ritsma (1962) is rather weak. Adding spectral components makes the task easier to perform (Patterson, 1973).

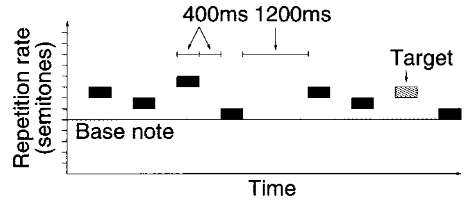

FIG. 1. Schematic of the melody task used in all experiments.

B. Method

1. Stimuli

The repetition rate of the harmonic complex tones was varied from 16 to 512 Hz in semitone steps (6%). Components that fell in the range 10 Hz–10 kHz were included. The components all started in cosine phase and so the sounds were essentially broadband click trains. Each tone was 400 ms long, and included 5-ms, squared-cosine on and off ramps. The silent interval between the tones within a melody was also 400 ms long. The stimuli were generated off-line by additive synthesis in the time domain. The overall presentation level of the broadband harmonic complex was 55 dB SPL.

The stimuli were generated with a 25-kHz sampling rate and presented using a TDT system II. The sound files containing the stimuli were stored on a PC disk. They were played back through a DD1 16-bit digital-to-analog converter, an FT-6 anti-aliasing filter with a 10-kHz cutoff, a PA4 attenuator, and a HB6 headphone buffer. The stimuli were presented diotically through a set of AKG K-240-DF headphones. The experiments took place in a double-walled, sound-insulated booth.

2. Procedure and listeners

A 4-alternative forced-choice (4AFC) task was used (Fig. 1). Each trial began with a short melody of four notes. The melody was characterized by the repetition rate, R_{rep}, of its base note. Given the base note, the melody was produced by drawing four values of R_{rep} randomly, with replacement, from the four semitones above the base note. This means that the melody could contain any note with a repetition rate equal to that of the base note, or 6%, 12%, 18%, or 24% higher than that of the base note. No other constraints were placed on the melody. In particular, any number of repetitions of the same note could occur randomly. In musical terms, the melodies were drawn from the chromatic scale rather than from a diatonic scale.

After a 1200-ms pause, the original melody was then repeated, but with one note changed plus or minus one semitone at random. The position of this target note was chosen at random. This means that the change could occur on the

TABLE I. Results for experiment I. The stimuli were broadband, harmonic complex tones. For individual subjects, the mean and standard deviation are based on the last six reversals of the adaptive runs.

	L1	L2	L3	Mean
Mean LLMP (Hz)	33.5	32.4	29.2	31.7
Standard deviation (Hz)	4.2	2.4	2.5	3.6

lowest or highest note of the melody and that there could be as many as five different values of R_{rep} in the two melodies. The change in repetition rate of the target note was the only difference between the two presentations of the melody. The listeners' task was to indicate the position of the target note by pressing one of four buttons on a response box. No time limit was imposed on listeners to give their response. Visual feedback was provided during a 1-s pause before proceeding to the next trial. If listeners were inattentive for a trial, they had the option of repeating the trial once with the same base note, but with a different melody and a new, random position of the target note.

A 3-down, 1-up adaptive threshold technique was used to track the LLMP (Levitt, 1971). Twelve reversals were measured. After three successive correct identifications of the target note, the repetition rate of the base note was lowered by 4 semitones for the first four reversals and 2 semitones for the last eight reversals. After each mistake, the repetition rate of the base note was increased by 3 semitones for the first four reversals and 1 semitone for the last eight reversals. The last six reversals were averaged to produce the threshold estimate for that run. Theoretically, this adaptive method converges during the last two-thirds of the track (which was the part that was analyzed) toward a probability for correct of $(1/3)^{1/3}$, i.e., the 69% correct point of the psychometric function. The starting base note for the adaptive run was 150 Hz. One complete practice run followed by two experimental runs were performed by each listener.

Three listeners aged 24 to 35 participated in the experiment. Listeners 1 and 2 had moderate musical training, listener 3 had no musical training whatsoever. Listener 1 was the first author. Listener 3 was paid for her participation. All had normal hearing thresholds ($<$15 dB HL) at standard audiometric frequencies.

C. Results and discussion

Results for the three listeners are shown in Table I. The LLMP for broadband harmonic tones is found to be between 30 Hz and 35 Hz.

These values are specific to the criterion used to define the LLMP, namely the 69% correct point when comparing two chromatic, four-note melodies. It is possible that listeners might have been able to use different cues on different trials. When the alteration of the target note changed the contour of the melody, the comparison could have been facilitated (Dowling and Fujitani, 1971). Primacy and recency effects in memory could also facilitate the task when the target note was at the beginning or the end of the melody (Crowder and Morton, 1969; McFarland and Cacace, 1992). Finally, the target note was sometimes well above the base-note used to define threshold. Nevertheless, since a 3-down, 1-up adaptive procedure was used, it is likely that these randomly occurring cues could not be used consistently, and that the only reliable strategy for the listeners was to store the pitch of the four notes in memory whenever possible.

It is also the case that the LLMP values compare well with previous values reported in the literature, even when obtained with very different experimental procedures. The LLMP is slightly lower than the lowest value reported by Ritsma (1962) for subjective perception of residue pitch (40 Hz). This is perhaps because broadband sounds were used instead of three-component complexes. Note also that if the middle note of the melody is chosen to define the LLMP, this small discrepancy vanishes. Guttman and Pruzansky (1962) found that listeners reported a pitch sensation for click trains with repetition rates as low as 19 Hz. However, when the same listeners performed an objective, octave-matching task, the limit was found to be 60 Hz. The LLMP value from the current study falls in between these two values. When analyzing their data with a fixed accuracy criterion (a semitone), Guttman and Pruzansky (1962) found that the limit was between 38 and 45 Hz, which is even closer to the value from the current study.

III. EXPERIMENT II: THE EFFECT OF FREQUENCY REGION

A. Rationale

The lower limit of the existence region of pitch increases with frequency region for SAM tones (Ritsma, 1962). Similarly, rate discrimination performance deteriorates when stimuli are limited to higher-frequency regions (Houtsma and Smurzynski, 1990; Krumbholz *et al.*, 2000). Accordingly, the influence of frequency region on the LLMP was investigated by bandpass filtering the harmonic complex tones of experiment I. Low-frequency regions were included since they are representative of the human voice, and many musical instruments include low-frequency energy.

B. Method

1. Stimuli

The bandpass filtering process was identical to Krumbholz *et al.* (2000). The filter had a nonattenuated section of constant width (600 Hz). The lower edge of this section is referred to as F_c, the filter cutoff. Five values of F_c were investigated: 200, 400, 800, 1600, and 3200 Hz. A frequency region with $F_c = 6400$ Hz was initially included but pilot data showed that the melody task was completely impossible with this high cutoff frequency. On both sides of the nonattenuated section, the filter had linear quarter-cosine skirts to minimize the effects of edge tones and to reduce the possibility of tracking individual harmonics. The lower skirt extended over 200 Hz and the upper skirt over 1 kHz. The repetition rate, R_{rep}, of the harmonic complex varied from 16 to 512 Hz in semitone steps. The stimuli were generated off-line by additive synthesis in the time domain. The filter was applied to the components during the additive synthesis. The overall presentation level of the harmonic complex was 55 dB SPL.

Continuous low-pass filtered pink noise was added to the stimuli before playback to mask any distortion products that might otherwise have been audible below F_c (Goldstein, 1967). The unfiltered pink noise had an overall level of 30 dB SL. It was filtered by an FIR filter with 158 taps, implemented by a TDT PF1. It was a brickwall-filter designed in the frequency domain with a cutoff at $F_c - 200$ Hz. The continuous pink noise was played from a DAT tape through a TDT PF1 filter and a PA4 attenuator. The noise and harmonic tones were then mixed by a SM3 summer.

2. Procedure and listeners

Thresholds were determined in separate adaptive runs for each filter condition. Within each filter condition, the procedure was the same as that of experiment I. The apparatus was also the same. The order of filter conditions was varied across listeners, and all runs for a given condition were done in the same session. Two of the listeners from experiment I (L1 and L2) took part in experiment II. A new listener, L4, with no musical training whatsoever also participated and was paid for her participation. She was 23 years old and had normal-hearing thresholds at standard audiometric frequencies.

C. Results and discussion

The results are shown in Fig. 2 (solid lines). The patterns are consistent across listeners and so the discussion is limited to the average data. There is a strong effect of frequency region on the LLMP. Listeners are extremely good at the melody task in the two lower-frequency regions; the adaptive procedure converges to a threshold around 35 Hz for both the 200- and 400-Hz filter conditions. In fact, listeners are just as good with the current band-limited stimuli as they were with the broadband stimuli in the previous experiment; for comparison the star symbol on the left-hand side of the graph presents the average threshold from that experiment. The LLMP then increases as F_c increases to the point where a repetition rate greater than 270 Hz is necessary to hear melodies in the highest filter condition, where F_c is 3200 Hz. The influence of frequency region on the LLMP region is similar to that reported by Ritsma (1962). The current results demonstrate that listeners can actually use the pitch cue to perform a melody task right down to the lower boundary of the existence region as defined by Ritsma (1962) and Moore (1973).

It is to be noted that listeners still reported hearing a pitchlike sensation at threshold in the higher filter conditions, which was not the case for the low filter conditions. There are pitch experiments performed with bandpass or highpass filtered click trains where the combination of repetition rate and frequency region would mean the pitch was below the LLMP as measured in the current study. Carlyon (1997) reported some informal melody recognition experiments with a 3900–5300 Hz passband and a base rate of 200 Hz. Melody recognition was possible, even though the pitch was below the current LLMP. Recognizing a familiar melody, however, is a task facilitated by higher cognitive processes, where a few cues might be enough to extract sufficient information

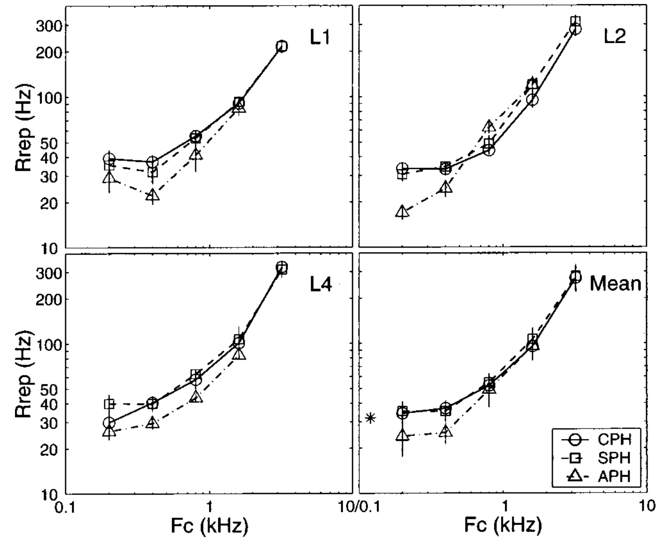

FIG. 2. Results for experiments II and III. The lowest repetition rate, R_{rep}, of the base note for which the melody task could be performed is plotted as a function of filter cutoff, F_c. The parameter is phase condition (CPH for experiment II, SPH and APH for experiment III). The star symbol on the bottom right panel is the mean result for broadband clicks (experiment I).

about the contour to do the task. The task in the current experiments was designed to minimize such cues. Houtsma and Smurzynski (1990) and Kaernbach and Bering (2000) reported that musical interval recognition was still possible for severely high-passed click trains, although with reduced accuracy. The current task involves a more stringent criterion as it requires semitone accuracy in interval recognition, as is the case for the Western musical scale. These observations suggest that the 30-Hz limit obtained in the broadband condition of experiment I is an absolute lower limit for melodic pitch. This value limits performance in the lower filter conditions ($F_c = 200$ or 400 Hz). For higher filter conditions, a pitchlike percept may be perceived by listeners between this absolute limit and the LLMP, but it is not sufficiently precise to support threshold performance on the melody task.

IV. EXPERIMENT III: THE EFFECT OF PHASE

A. Rationale

Performance in pitch-discrimination tasks is usually better when the stimuli include spectrally resolved components (Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994; Carlyon and Shackleton, 1994). The question then arises as to whether the LLMP might reflect the transition from resolved to unresolved components in the internal representation of the sound. To test this hypothesis, performance was compared for stimuli having the same amplitude spectra but different phase spectra. When the components are resolved, differences in phase have little or no influence on performance in pitch tasks; whereas when the components are not resolved phase differences can affect performance (Patterson, 1987; Houtsma and Smurzynski, 1990; Carlyon and Shackleton, 1994). The manipulation of phase also enables evaluation of the effect of wave shape on the LLMP.

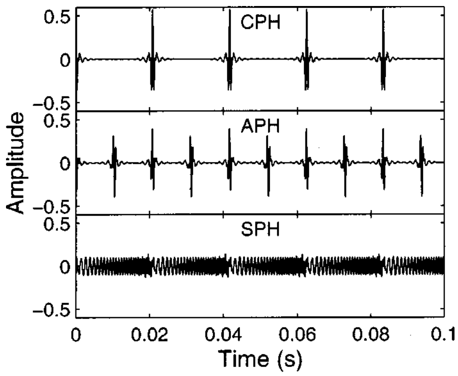

FIG. 3. Waveforms of the stimuli for experiments II (CPH) and III (APH and SPH).

B. Method

1. Stimuli

Three phase relationships were used. In the first condition, all components were in cosine phase (CPH) as in experiment II. This configuration produces the largest possible peak factor for a harmonic complex. The repetition rate is clearly visible in the stimulus waveform (Fig. 3, top panel). In the second condition, every other component was shifted $\pi/2$ radians. This alternating phase (APH; Patterson, 1987) produces a stimulus with a periodic Hilbert envelope at double the repetition rate, $2R_{\rm rep}$, although the repetition rate of the fine-structure is still $R_{\rm rep}$ (Fig. 3, middle panel). The last phase configuration was derived using the formula proposed by Schroeder (1970). This reduces the amplitude peak factor of the waveform markedly (Fig. 3, bottom panel). The sign of the phase in the Schroeder formula does not change the waveform peak factor. It does change the direction of the chirp in the waveform fine structure. An upward-chirping tone was chosen because evidence exists to indicate that this condition reduces the peak factor of the internal representation of the signal after auditory filtering (Smith *et al.*, 1986). This last condition will be referred to as Schroeder phase (SPH).

2. Procedure and listeners

The thresholds for each combination of filter condition and phase were obtained in separate adaptive runs. The procedure, apparatus and listeners were those of experiment II. The CPH conditions were not repeated; the results were taken directly from experiment II. All thresholds for the other two phase conditions were measured in an identical manner as for the CPH thresholds; the order was varied across subjects.

C. Results

The results are shown in Fig. 2. The CPH and SPH conditions produce very similar results. A Wilcoxon rank-sum test was applied to the raw data for all filter conditions and there were no significant differences ($p>0.2$). This appears to be at variance with the results of Houtsma and Smurzynski (1990) who reported differences in rate discrimination threshold for CPH and SPH harmonic complexes. However, the combinations of F_c and $R_{\rm rep}$ associated with the LLMP in the current experiment are different from the combinations where Houtsma and Smurzynski found reliable differences between the RDTs of CPH and SPH tones—200 Hz at a cutoff frequency of 3.2 kHz or higher. These parameters would produce a stimulus that falls below the pitch region revealed by the current study (270 Hz when F_c is 3.2 kHz). Another difference is that Houtsma and Smurzynski used wideband masking noise that might have interferred with the perception of the stimulus. Lowpass noise was used in the current experiment to avoid interference. Finally, this small discrepancy could reflect a difference between the RDT and LLMP tasks.

In the two lower filter conditions, performance is better for APH stimuli than for CPH or SPH stimuli, and the difference is significant ($p<0.01$, Wilcoxon rank-sum test on the raw data). As F_c increases, the LLMP increases faster for the APH stimuli than for the CPH and SPH stimuli, and no threshold could be measured with the adaptive procedure for the highest filter condition. Perceptually, for a given repetition rate, the pitch increases one octave for unresolved APH tones (Patterson, 1987; Carlyon and Shackleton, 1994). No octave shift has been reported for SPH tones.

V. SPECTRAL RESOLVABILITY

Two analyses were performed to determine whether the increase in the LLMP with frequency region reflected the transition from resolved to unresolved components. The hypothesis is that a clear pitch is required to do the melody task, and it can only be produced by resolved harmonics. One definition of resolvability is that the transition occurs at a constant harmonic number (Plomp, 1964). The precise value of this number varies from 6 to 12 between authors and according to the experimental task. Nevertheless, it should be a fixed value. The number of the lowest harmonic in the stimulus associated with each LLMP value was computed for all combinations of F_c and phase. The results are presented in Fig. 4 (upper panel), which shows that the LLMP does not correspond to a fixed harmonic number. Moreover, the lowest harmonic is the 14th or 15th in some cases.

Shackleton and Carlyon (1994) have proposed modifying the ''constant harmonic number'' rule to introduce a ''constant number of components per auditory filter'' rule. They suggest that the transition region between resolved and unresolved complexes occurs when there are 2 to 3.25 components within the 10-dB bandwidth of the auditory filter, as defined by Glasberg and Moore (1990). The number of components in the auditory filter centered on F_c was computed for stimuli at the LLMP, for each experimental condition. The results are presented in Fig. 4 (lower panel) which shows that number of components per filter at the LLMP varies with F_c, and that most of the observed values occur in the ''unresolved'' region as defined by Shackleton and Carlyon (1994).

In summary, these two analyses, which involve relatively large estimates of the upper limit of spectral resolution, nevertheless indicate (1) that the LLMP does not correspond to the loss of spectral resolution for either criterion, and (2) the LLMP is typically associated with stimuli having no resolved components. Further support for these conclu-

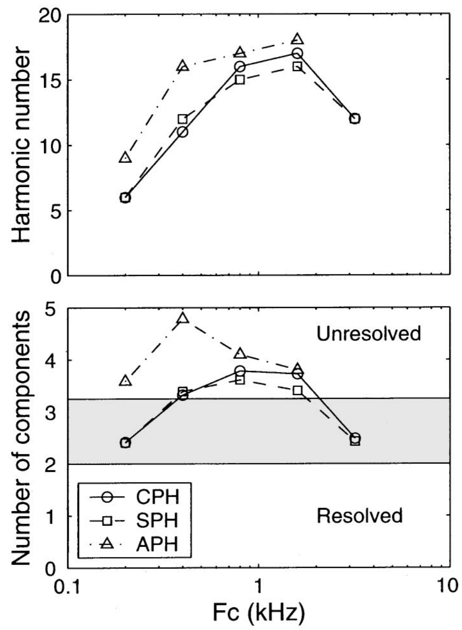

FIG. 4. Harmonic number (upper panel) and number of components between the 10-dB-down points of the auditory filter centered on F_c (lower panel), at threshold, for experiments II and III.

sions is provided by the phase effect in experiment III; the phase effect should only occur if the complexes are unresolved[2] (Patterson, 1987; Carlyon and Shackleton, 1994).

VI. SPECTRAL AND TEMPORAL MODELS OF THE LLMP

From the above discussion, it is clear that simple spectral models of pitch perception that only predict a pitch for perceptually resolved harmonics would fail to predict the LLMP. There exist, however, more sophisticated spectral models where harmonics can be represented above the psychophysical limit of resolution. In the Central Spectrum model (Goldstein, 1973; Srulovicz and Goldstein, 1983), the spectral representation is derived from time intervals in simulated auditory nerve fibres. The resolution of the model is improved by the use of a matched filter on the interval distribution for each fibre. Up to 15 components can be present in the central spectrum (Goldstein *et al.*, 1978; Srulovicz and Goldstein, 1983). The SPINET model (Cohen *et al.*, 1995) uses a on-center, off-surround mechanism that enhances spectral contrast. Both of these models could maintain some sprectral representation at rates corresponding to the LLMP. They would also predict a decrease of performance in high-frequency regions because of the reduction in the precision of the spectral components (Goldstein, 1973; Cohen *et al.*, 1995).

Temporal models where pitch is associated with the dominant periodicities in a range of frequency channels (Schouten *et al.*, 1962) do not immediately explain the LLMP. The temporal precision of the envelope at the output of auditory filters improves in high-frequency channels, but the LLMP nevertheless increases for high filter conditions. Moore (1973, 1997) proposed a theoretical model in which the time intervals in a given channel with center frequency, CF, were restricted to those between 0.5/CF and 15/CF ms. This CF-dependent limit would cause the LLMP to increase in high-frequency regions; however, 15/CF leads to very long intervals for low CFs (75 ms for a 200-Hz channel). The predicted LLMP value would be above the observed one. In the next section, we develop a modified autocorrelation model (Licklider, 1951; Meddis and Hewitt, 1991a) with a CF-independent limit on time intervals, that is able to reproduce the experimental LLMP results. Thus the present data do not preclude either modeling approach. Rather, they indicate that a residue mechanism, be it temporal or spectral, can convey melodies for notes as low as those that can be played on the lowest octave of the piano keyboard.

VII. AN AUTOCORRELATION MODEL OF THE LLMP

A. The autocorrelation model of pitch perception

Licklider (1951) produced the first computational model of pitch perception based on time intervals within auditory frequency channels. The incoming signal is bandpass filtered to simulate cochlear frequency selectivity; then, in each channel, a running autocorrelation function (ACF) is calculated to reveal any periodicity. The output is an array of ACFs and so the dimensions are autocorrelation lag versus filter center frequency, at a given moment in time. This representation is typically referred to as an autocorrelogram (ACG) and it has been used to explain pitch perception. There are distinct limitations to the autocorrelation approach, as noted by Kaernbach and Demany (1998). In this subsection, the Meddis and Hewitt (1991a) implementation of the autocorrelation model is described as a simple means to quantify the time-interval information present in auditory channels.

Meddis and Hewitt (1991a) introduced two additional stages to the autocorrelation model to enable quantitative predictions of pitch perception. First, they averaged the ACFs of all frequency channels to form a summary autocorrelogram (SACG). This emphasizes the time intervals common to a range of frequency channels. The dominant peak in the SACG specifies the period of the predicted pitch. Second, in order to model pitch discrimination performance, they defined a decision statistic equal to the Euclidean distance between the SACGs of pairs of sounds, referred to as d^2. The SACG and the d^2 statistic have been successful in accounting for a number of pitch phenomena (Meddis and Hewitt, 1991a, b; Meddis and O'Mard, 1997).

It is useful for the purposes of the following discussion to point out some of the properties of the autocorrelation calculation at the heart of the model. Licklider proposed performing a running autocorrelation of the signal, *s*, with an exponential time window [Eq. (1)]:

$$\mathrm{ACF}(t,\tau)=\int_0^{+\infty} s(t-T)s(t-\tau-T)e^{(-T/\Omega)}\,dT. \tag{1}$$

The time constant, Ω, determines the decay rate of the exponential window and so the time over which the ACF is averaged. Licklider (1951) suggested a value of 2.5 ms for Ω. The ACF fluctuates over time when this parameter is

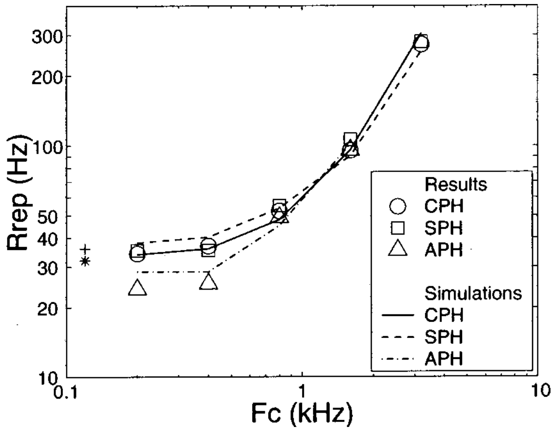

FIG. 5. Results and simulations for experiments I, II, and III. The mean experimental data are represented as unconnected symbols, the model simulation as curves without symbols. The star and cross symbols on the left-hand side of the figure are, respectively, the results and the simulation for the broadband condition.

short (Wiegrebe *et al.*, 1998). To reduce this variability when comparing the SACG of two stimuli, Meddis and Hewitt (1991a) suggested calculating the SACG at times corresponding to a multiple of the stimulus period. With this constraint and an Ω value of 2.5 ms, Meddis and Hewitt were able to explain a wide range of pitch phenomena with the SACG model. Alternatively, Ω can be set to a longer value to produce a better estimate of the long-term unbiased ACF. When explaining the perception of vowels, Meddis and Hewitt (1992) used a value of 25 ms. The time constant Ω is relevant to the integration of pitch over time. For stationary sounds, a relatively long value is usually satisfactory to stabilize the SACG. This parameter is not, however, directly involved in the determination of the LLMP.

B. Representation of the LLMP in the autocorrelation model

With regard to the LLMP, the important parameter is the maximum lag, henceforth denoted τ_m, for which the individual ACFs are computed. The model cannot explain the pitch of stimuli whose period is greater than this value since there would be no peak in the SACG to associate with the pitch. The role of τ_m has not been studied explicitly. Licklider (1951) pointed out the need to specify the maximum lag, and in the absence of experimental data, proposed a value of 33 ms (30 Hz). Similarly, Yost *et al.* (1996) noted that this parameter is related to the lowest audible pitch, and set the value of 35 ms. In general, however, τ_m seems to be set just long enough to ensure that the peaks produced by the stimulus of interest appear in the SACG (Meddis and Hewitt, 1991a, 1992; Meddis and O'Mard, 1997; Yost *et al.*, 1996). This has led to values of τ_m ranging from 10 to 35 ms, without any direct estimation of the parameter.

The problem with the current model is that there is a discontinuity at the boundary. If the maximum lag is set to $\tau_m = 33$ ms, for instance, a 35-Hz tone (29-ms period) produces some activity in the SACG whereas a 25-Hz tone (40-ms period) does not. As a consequence a large d^2 would be obtained predicting excellent discrimination for this semitone change, which is clearly at variance with experimental data (Ritsma and Hoekstra, 1974; Houtsma and Smurzynski, 1990; Krumbholz *et al.*, 2000).

A straightforward way to avoid this problem is to apply a weighting function to the SACG. The weighting chosen here decreases linearly from one at a lag of 0 ms to zero at a lag of 33 ms. This eliminates the discontinuity by making long time intervals disappear gradually from the SACG. Intuitively, the weighting function reduces pitch strength for long lags. It should be noted that the same result is obtained if, instead of a running ACF, a biased, long-term ACF is calculated with 33-ms, unwindowed, portions of the signal. This alternative implementation would depart significantly from the traditional structure of the autocorrelation model, so for the purposes of the current paper we focus on the established models of Licklider (1951) and Meddis and Hewitt (1991a).

C. Simulation of the experimental results

The first stages of the model were identical to Meddis and O'Mard (1997). There were 60 frequency channels regularly distributed on an ERB scale between 100 and 8000 Hz, each with a gammatone filter and a hair-cell simulator. The individual ACFs of the ACG were computed on the hair-cell outputs. The ACF was computed at the end of a complete period of the stimulus as in Meddis and O'Mard (1997). The time constant of the ACF, Ω, was increased from 2.5 to 15 ms. The linear weighting function was applied to the summary ACF terminating at $\tau_m = 33$ ms. The weighted SACG is also referred to as the SACG, for convenience.

LLMP values were produced with this model as follows: SACGs were calculated for all stimuli from the experiments with R_{rep} values between 16 and 340 Hz. Then, for each combination of filter and phase condition, d^2 was computed between the SACGs of stimuli with repetition rates separated by 6%. A threshold value, d^2_{thres} was fixed and the lowest note of the pair of notes whose d^2 just exceeded d^2_{thres} was taken as the estimate of threshold for that combination of stimulus filter and phase. It was verified that the d^2 increased monotonically for notes above threshold. A complete set of LLMP values was produced for the fixed d^2_{thres} and then, the value of d^2_{thres} was varied to find the set of LLMP values that matched the observed values. d^2_{thres} was the only parameter varied in the fitting process. The low-pass noise was not included in the simulations because the model does not produce distortion products. As it is a deterministic model, the addition of the random noise would have needlessly complicated the computation. The results are presented in Fig. 5.

The model reproduces most of the important features of the experimental data (Fig. 5). For the broadband condition, the simulated LLMP is 35 Hz, similar to that derived from experiment I. There is little difference between this condition and the CPH condition for $F_c = 200$ Hz, and the LLMP increases rapidly with increasing frequency region for CPH stimuli (experiment II). The model also exhibits the effect observed in experiment III, where the APH stimuli produce lower LLMP values than the CPH stimuli in the lowest filter conditions ($F_c = 200$ or 400 Hz). In the highest filter condi-

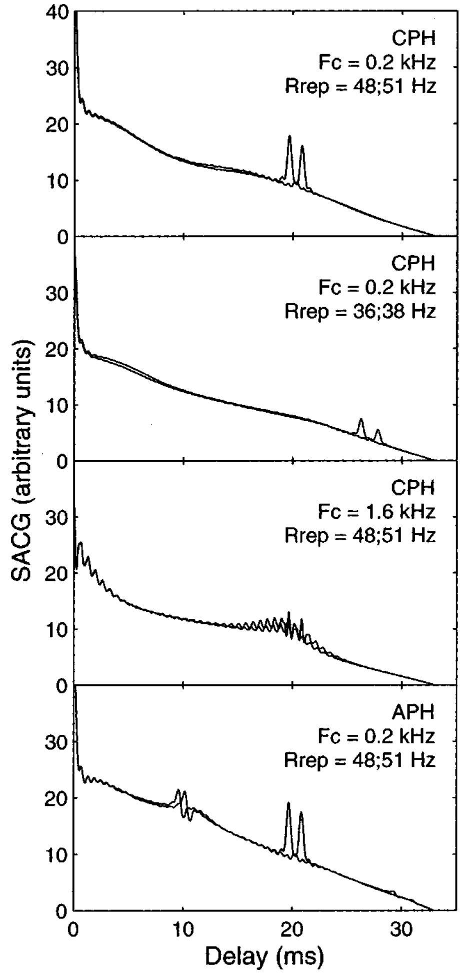

FIG. 6. Pairs of SACGs for stimuli differing by a semitone. See inserts for stimulus parameters.

tion ($F_c = 3200$ Hz), d^2 never reached the threshold criterion and so there is no predicted LLMP value. This is also the condition where the listeners could not perform the task. There is a slight discrepancy between the model values and the experimental data in the low filter conditions ($F_c = 200$ or 400 Hz); the LLMP for the CPH and APH conditions are greater than the experimental ones.

D. Discussion of the simulation

Having imposed a weighting function on the SACG that reduces it to zero at 33 ms, it is not surprising to find that the lowest LLMP values produced by the model are just greater than 30 Hz for CPH. The intriguing finding is that the LLMP values increase with frequency region at the same rate as observed in the data, and that the effect of APH is also reproduced. To understand how this arises, consider the SACGs presented in Fig. 6. Each panel presents two superimposed SACGs for notes separated by 1 semitone—the difference that distinguishes the pair of melodies in a given trial of the experiments and which was used to calculate the d^2 values.

In the upper panel, the stimulus filter has a cutoff of 200 Hz and the notes are well above the LLMP (48 and 51 Hz). The SACGs exhibit clear peaks at the delays of the two notes. In this case, d^2 is large and the model predicts good performance as is observed. As the pitch is lowered toward the LLMP at this filter cutoff (next panel, 36 and 38 Hz), the peaks in the SACG shrink, producing ever smaller d^2 values until eventually d^2 falls below threshold.[3] This is the direct result of the introduction of the weighting function.

Now consider what happens when the filter cutoff is increased to 1.6 kHz. The third panel presents the SACGs for the same notes as in the upper panel (48 and 51 Hz) but with the higher filter cutoff. The weighting function is the same but the peaks are shorter and more spread out. As a result, d^2 is reduced to the point where it falls below threshold and this condition is correctly predicted to be below the LLMP in this filter condition. The broadening of the peaks when energy is restricted to a band of high-frequency channels was noted by Meddis and O'Mard (1997). They argued that the effect arises from the loss of harmonic resolution at high harmonic numbers. However, as Carlyon (1998) noted, it is more an effect of frequency region than of resolvability. We suspect that there are several factors at work here. The first is the loss of phase locking that occurs at high frequencies and restricts the encoding of temporal fine structure. The second factor has to do with the averaging of ACFs across channels. This enhances activity at the stimulus period because there is activity at this lag in all channels. Activity associated with the center frequency of the channel varies with channel and cancels out in the cross-channel averaging. In experiments II and III, the width of the stimulus filter was fixed. As a result, in the low-frequency conditions, there were more active channels which leads to more summation of the common stimulus period and more cancellation of auditory filter activity than in the high-frequency conditions where the stimuli excite fewer filters. The spreading of the peaks is due to this lack of cancellation of auditory filter ringing combined with the loss of phase locking.

Finally, consider the SACGs of the APH stimuli when the cutoff of the stimulus filter is 200 Hz (bottom panel). The phase shift produces smaller peaks mid-way through the period of the stimulus which results in secondary peaks in the SACG mid-way between the main peaks. As the pitch is lowered and the period passes 33 ms, the listeners can switch from the main to the secondary peaks and so perform the melody task for periods that are nominally below the LLMP. This interpretation is compatible with previous explanations of the perception of APH sounds by autocorrelation models (Meddis and Hewitt, 1991b). When the filter cutoff is increased, the secondary peaks fade into the floor activity and so this cue cannot be used in high-frequency regions.

E. Limitations of the d^2 statistic

The model as presented in this paper is not invariant with respect to the bandwidth of the stimuli. The upper and middle panels of Fig. 7 show the output of the model in response to stimuli with a high filter cutoff ($F_c = 3.2$ kHz). In the upper panel, bandwidth is that used in experiments II and III. In the middle panel, the bandwidth has been enlarged to be proportional to auditory filter width (Glasberg and Moore, 1990). The proportional bandwidth equates the num-

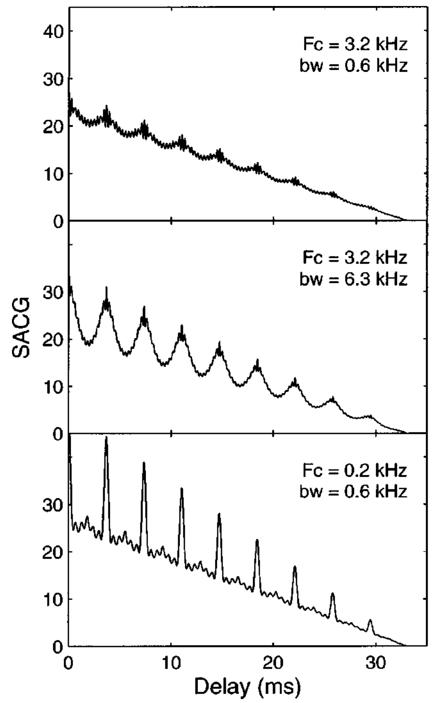

FIG. 7. SACGs for constant or proportional bandwidth stimuli. Phase condition for all stimuli is CPH, repetition rate is $R_{\mathrm{rep}}=271$ Hz. The top, middle, and bottom panels illustrate the cases of constant bandwidth, proportional bandwidth, and constant/proportional bandwidth, respectively.

ber of auditory channels activated in the low and high filter conditions ($F_c=200$ Hz and 3.2 kHz). The increased bandwidth leads to more activation in the SACG and larger peaks, which in turn produces larger values of d^2. As a consequence, lower values of the LLMP are predicted compared to the LLMP obtained with constant bandwidth stimuli.

In contrast to this prediction, Krumbholz *et al.* (2000) have shown that the RDTs obtained with a fixed passband were similar to those obtained by other experimenters using proportional bandwidth or high-pass stimuli. This suggests that the LLMP is not much affected by bandwidth. Figure 7 thus exhibits a limitation of the autocorrelation model in its current implementation, mainly because of the nonauditory aspect of the d^2 statistic. The peaks produced by the proportional bandwidth stimulus in the high filter condition are much broader than those produced in the $F_c=200$-Hz condition (bottom panel); it is possible that a more sophisticated statistic based on peak-picking, and taking into account the width of the peaks, could still account for the increase of LLMP in high-frequency regions. Such a modification is beyond the scope of the current paper. The autocorrelation model has served the purpose of demonstrating that the information present in time intervals within auditory channels can in principle explain the LLMP results, and this is sufficient to indicate that models based more closely on the physiology of temporal processing in the auditory system are worth investigating.

VIII. THE LLMP AND RATE DISCRIMINATION THRESHOLD

In a companion paper (Krumbholz *et al.*, 2000), the RDT for many of the stimuli used in the current study were measured—for CPH and APH stimuli in all frequency regions. It is the case that RDT rises rapidly around the LLMP. Krumbholz *et al.* (2000) found that the LLMP actually corresponds to a RDT of about 2.5%. This is substantially less than the 6% rate change that was available in the melody task. In fact, for most listeners, RDT did not exceed 5% for rates as low as 16 Hz in all frequency regions. The listeners would have been able to perform a semitone discrimination task with rates much lower than the LLMP. Thus the LLMP is not the point where the target note and its semitone neighbors become indiscriminable.

There is an obvious difference between the melody task and the task used to measure RDTs. In the RDT case, listeners compared a single pair of sounds. In contrast, the melody task involves eight different sounds. A simple model of the listeners' strategy would assume that they stored all the sounds, or some attributes of the sounds, in memory and then performed pairwise comparisons to make their judgments. In this case, the LLMP task is similar to four, parallel, RDT tasks, and signal detection theory predicts a reduction in discriminability by a factor of $\sqrt{4}$, as measured by d' (Macmillan and Creelman, 1991). Plack and Carlyon (1995) have shown that d' is proportional to the rate difference for RDT tasks with harmonic complex tones. Thus the fact that the threshold for the melody task with 6% between notes corresponds to a RDT of about 3% is consistent with the decision model based on the storage and retrieval of the sequence of sounds.

This leads to the argument that the cues used to perform the LLMP task were, indeed, pitch cues. The eight rates have to be stored over 4 s. Each note is perceived as a distinct sound event because it is 400 ms long, but the task clearly includes a memory component on a relatively long time scale. McFarland and Cacace (1992) have studied the perception of binary tone patterns similar in duration to the melodies used in the current experiments. The tone patterns were constructed by randomly alternating between two values of a given parameter, which was frequency, amplitude or duration. The maximum number of stimuli in the pattern that could be reliably memorized was determined by means of an adaptive procedure. To make comparison across parameters realistic, the difference between the two parameter values in the sequences was always a constant number of *jnd*'s for that parameter. McFarland and Cacace (1992) found that patterns based on alternation in frequency were retained longer than patterns based on alternation of amplitude or duration (4.4 s for frequency, as opposed to 1.7 s for amplitude or duration).

The superior performance with pitch sequences can be explained by the results of Semal and Demany (1991, 1993) who have shown that there is a pitch-specific memory that cannot be used for loudness. Clément *et al.* (1999) found that the accuracy of the pitch trace is maintained longer than that for loudness. Moore and Rosen (1979) had previously observed that no melody could be recognized when pitch intervals were replaced by loudness intervals, even when listeners were selecting from a closed set of familiar melodies. Interestingly, Semal and Demany (1993) also demonstrated that the timbre cues associated with repetition rates below 30 Hz, and hence below the LLMP, could not be stored in pitch memory.

These findings point to the following interpretation of the LLMP and its link to RDTs. For the higher repetition rates, pitch cues mediate both the melody task and RDT ($<1\%$). Listeners can store the pitch cues in memory and use them in the melody task until criterion performance is reached. As repetition rate decreases and the pitch cues become less reliable, listeners switch to other cues such as roughness or pulse rate to perform rate discrimination, and this explains the sudden rise in RDT to around 5%. The melody task requires a memory that operates for longer than a standard 2AFC rate discrimination trial, and thus pitch is the sole cue that works in the melody task.

IX. MELODIC AND MUSICAL PITCH

It is debatable whether the pitch cue used in the LLMP task can truly be called ''musical.'' This is why the more restricted term of melodic pitch was used. To establish whether pitch retains a musical quality down to the LLMP, further measurements involving interval recognition would be needed (Houtsma and Goldstein, 1972). It is likely that such measurements would be difficult to obtain, however, because of the need to control for (1) a restricted range and thus a limited number of intervals, (2) the ability of listeners to learn how to label arbitrarily a small set of rate differences with musical names. Randomly transposing the melodies seems a more appropriate approach to resolve this issue (de Boer and Houtsma, personnal communication). Also, the 69% correct value chosen as threshold might prove insufficient to convey melodies reliably in a musical context (Goldstein, 2000). Measurement of the rates that support performance closer to 100% correct is not practical with an adaptive procedure; they would have to be inferred from psychometric functions. In spite of these restrictions concerning the musical relevance of the LLMP and its relation to more established definitions of musical pitch, it is to be noted that the 30 Hz value corresponds to the lowest note available on the piano keyboard (A0, 27.5 Hz).

X. CONCLUSIONS

A melody-change task was used to measure the lower limit of melodic pitch which was found to be around 30 Hz, provided there was energy in the stimulus below 800 Hz. The value of 30 Hz corresponds roughly to the lowest note on the piano keyboard (27.5 Hz), but it is an octave above the 16-Hz pipe on large organs.

When the stimuli were bandpass filtered, the LLMP was found to increase rapidly with frequency in the region above 800 Hz. This effect is consistent with the results of Ritsma (1962) and Moore (1973) despite the differences in stimuli and experimental task. Above 30 Hz and below the LLMP, listeners still experience a weak pitch sensation but it is not sufficiently well defined to convey melodies with semitone accuracy.

The LLMP does not correspond to the loss of spectral resolution for individual harmonics, as usually defined; unresolved harmonics can support melodic pitch.

The data can be simulated by a modified autocorrelation model where a limit of about 33 ms is imposed on the time intervals that the pitch mechanism can accommodate.

ACKNOWLEDGMENTS

The authors would like to thank C. Kaernbach, C. Plack, and J. L. Goldstein for helpful comments on earlier versions of the manuscript. We also wish to thank A. de Cheveigné, L. Demany, and R. P. Carlyon for their much appreciated input. Supported by the U.K. Medical Research Council (G9703469). The first author is supported by the French Center National de la Recherche Scientifique.

[1]Throughout this paper, stimuli are described in terms of their waveform repetition rates, in Hertz. This is a convenient way to describe harmonic sounds that may or may not contain energy at their fundamental frequency.

[2]The APH effect occurs in lower-frequency regions, which seems to be the opposite of what is normally observed. Note, however, that the important factor for an octave shift is harmonic resolvability, not frequency region. The LLMP occurs at very low rates in low-frequency regions where the components are unresolved.

[3]In the modified model we assume that d^2 represents only the pitch cues that can be used for discrimination between the two notes. For long periods (>33 ms), the SACG would be empty and the model would not predict any discriminability. In this case, listeners would still be able to discriminate between sounds, but by using cues that are not pitch cues, and therefore, which would not appear in the current model.

de Boer, E. (**1976**). ''On the ''residue'' and auditory pitch perception,'' in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer Verlag, Berlin).

Carlyon, R. P. (**1997**). ''The effect of two temporal cues on pitch judgment,'' J. Acoust. Soc. Am. **102**, 1097–1105.

Carlyon, R. P. (**1998**). ''Comments on 'A unitary model of pitch perception','' 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.

Clément, S., Demany, L., and Semal, C. (**1999**). ''Memory for pitch versus memory for loudness,'' J. Acoust. Soc. Am. **106**, 2805–2811.

Cohen, M. A., Grossberg, S., and Wyse, L. L. (**1995**). ''A spectral network model of pitch perception,'' J. Acoust. Soc. Am. **98**, 862–879.

Crowder, R. G., and Morton, J. (**1969**). ''Precategorical acoustic storage,'' Percept. Psychophys. **5**, 365–373.

Dowling, W. J., and Fujitani, D. S. (**1971**). ''Contour, interval, and pitch recognition in memory for melodies,'' J. Acoust. Soc. Am. **49**, 524–531.

Dowling, W. J., and Harwood, D. L. (**1986**). *Music Cognition* (Academic, London).

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. (**1967**). ''Auditory nonlinearity,'' J. Acoust. Soc. Am. **41**, 676–689.

Goldstein, J. L. (**1973**). ''An optimum processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. **54**, 1496–1515.

Goldstein, J. L. (**2000**). ''Pitch perception,'' in *Encyclopedia of Psychology*, edited by A. E. Kazdin, Vol. 6 (American Psychological Association & Oxford University Press), pp. 201–210.

Goldstein, J. L., Gerson, A., Srulovicz, P., and Furst, M. (**1978**). ''Verification of the optimal probabilistic basis of aural processing in pitch of complex tones,'' J. Acoust. Soc. Am. **63**, 486–497.

Griffiths, T. D., Büchel, C., Frackowiak, R. S. J., and Patterson, R. D. (**1998**). ''Analysis of temporal structure in sound by the human brain,'' Nat. Neurosci. **1**, 422–427.

Guttman, N., and Julesz, B. (**1963**). ''Lower limits of auditory periodicity analysis,'' J. Acoust. Soc. Am. **35**, 610.

Guttman, N., and Pruzansky, S. (**1962**). ''Lower limits of pitch and musical pitch,'' J. Speech Hear. Res. **5**, 207–214.

Houtsma, A. J. M. (**1995**). ''Pitch perception,'' in *Hearing, Handbook of Perception and Cognition*, 2nd ed., edited by B. C. J. Moore (Academic, San Diego), pp. 267–295.

Houtsma, A. J. M., and Goldstein, J. L. (**1972**). ''The central origin of the pitch of complex tones: Evidence from musical interval recognition,'' J. Acoust. Soc. Am. **51**, 520–529.

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

Kaernbach, C., and Bering, C. (**2000**). ''Ist temporale tonhöhe musikalisch?,'' in Proceedings of DAGA 2000, Oldenburg (in press).

Kaernbach, C., and Demany, L. (**1998**). ''Psychophysical evidence against the autocorrelation theory of auditory temporal processing,'' J. Acoust. Soc. Am. **104**, 2298–2306.

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.

Macmillan, N. A., and Creelman, C. D. (**1991**). *Detection Theory: A User's Guide* (Cambridge University Press, Cambridge, UK).

McFarland, D. J., and Cacace, A. T. (**1992**). ''Aspects of short-term acoustic recognition memory: Modality and serial position effects,'' Audiology **31**, 342–352.

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

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

Meddis, R., and Hewitt, M. J. (**1992**). ''Modeling the identification of concurrent vowels with different fundamental frequencies,'' J. Acoust. Soc. Am. **91**, 233–245.

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**). ''Some experiments relating to the perception of complex tones,'' Quart. J. Exp. Psychol. **25**, 451–475.

Moore, B. C. J. (**1997**). *An Introduction to the Psychology of Hearing*, 4th ed. (Academic, London).

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

Parncutt, R., and Cohen, A. J. (**1995**). ''Identification of microtonal melodies: Effects of scale-step size, serial order, and training,'' Percept. Psychophys. **57**, 835–846.

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

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

Patterson, R. D., Peters, R. W., and Milroy, R. (**1983**). ''Threshold duration for melodic pitch,'' *Hearing—Physiological Bases and Psychophysics*, edited by Klinke and Hartmann (Springer, Berlin), pp. 321–326.

Plack, C. J., and Carlyon, R. P. (**1995**). ''Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics,'' J. Acoust. Soc. Am. **98**, 1355–1364.

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. (**1976**). *Aspects of Tone Sensation* (Academic, London).

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

Ritsma, R. J. (**1971**). ''Psychological correlates of a frequency shift,'' in Proc. 7th ICA (p. 19H16), Budapest.

Ritsma, R. J., and Hoekstra, A. (**1974**). ''Frequency selectivity and the tonal residue,'' in *Psychophysical Models and Physiological Facts in Hearing*, edited E. Zwicker and E. Terhardt (Springer, Berlin).

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

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

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

Semal, C., and Demany, L. (**1991**). ''Dissociation of pitch from timbre in auditory short-term memory,'' J. Acoust. Soc. Am. **89**, 2404–2410.

Semal, C., and Demany, L. (**1993**). ''Further evidence for an autonomous processing of pitch in auditory short-term memory,'' J. Acoust. Soc. Am. **94**, 1315–1322.

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.

Smith, B. K., Sieben, U. K., Kohlrausch, A., and Schroeder, M. R. (**1986**). ''Phase effects in masking related to dispersion in the inner ear,'' J. Acoust. Soc. Am. **80**, 1631–1637.

Smoorenburg, G. F. (**1970**). ''Pitch perception of two frequency stimuli,'' J. Acoust. Soc. Am. **48**, 924–942.

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

Terhardt, E. (**1970**). ''Frequency analysis and periodicity detection in the sensation of roughness and periodicity pitch,'' in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. Smoorenburg (A. W. Sijthoff, Leiden), pp. 278–287.

Warren, R. M., and Bashford, J. A. J. (**1981**). ''Perception of acoustic iterance: Pitch and infrapitch,'' Percept. Psychophys. **29**, 395–402.

Wiegrebe, L., Patterson, R. D., Demany, L., and Carlyon, R. P. (**1998**). ''Temporal dynamics of pitch strength in regular interval noises,'' J. Acoust. Soc. Am. **104**, 2307–2313.

Yost, W. A., Patterson, R. D., and Sheft, S. (**1996**). ''A time domain description for the pitch strength of iterated rippled noise,'' J. Acoust. Soc. Am. **99**, 1066–1078.