Dichotic pitches as illusions of binaural unmasking. III. The existence region of the Fourcin pitch

John F. Culling[a)]
University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, United Kingdom

(Received 19 December 1997; revised 10 August 1999; accepted 7 December 1999)

Two experiments explored the existence region of the Fourcin pitch. In each experiment, detectability was assessed by measuring listeners' ability to discriminate pitch changes. In the first experiment, the detectability of the pitch was measured as a function of the number of noises used to generate it. In the second experiment, the pitch was generated using two noises with equal and opposite interaural delays and detectability was measured as a function of the difference between these two delays, and thus of the perceived pitch height. In each case, the experimental results were compared with the predictions produced by a model of binaural unmasking, based on equalization cancellation, that had been designed to recover broadband sounds, such as speech, from interfering noise [Culling and Summerfield, J. Acoust. Soc. Am. **98**, 785–797 (1995)]. The model accurately predicted the results from experiment 1, but failed to show an adequate decline in performance for small differences in interaural delay (corresponding to higher perceived pitches) in experiment 2. A revised model, based on similar principles, but using data on listeners' sensitivity to interaural decorrelation, rather than an equalization-cancellation mechanism, was able to predict the results of both experiments successfully. © *2000 Acoustical Society of America.* [S0001-4966(00)04403-4]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Hg, 43.66.Pn [DWG]

INTRODUCTION

Dichotic pitches are heard when white noise is presented to the two ears under various binaural configurations. Listening to each earphone alone, the listener just hears noise, but when both earphones are used simultaneously a tone of some sort is heard standing out from the noise. Due to the tonal nature of each of these phenomena, they have been termed dichotic ''pitches,'' and have hitherto been investigated via pitch-matching experiments. However, one might more broadly describe them as dichotically evoked sounds.

Culling and co-workers (1998a, c) argued that the three most salient dichotic pitches, known as Huggins' pitch (Cramer and Huggins, 1958), the binaural edge pitch (Klein and Hartmann, 1986) and the Fourcin pitch (Fourcin, 1958, 1970) are all illusions produced by the mechanism of binaural unmasking. Durlach (1962) and Klein and Hartmann (1986) had previously invoked binaural unmasking as a mechanism for producing these pitches, but in the case of the Fourcin pitch, the suggestion was novel. As evidence for this claim, they showed that many features of each kind of pitch, both from the literature and from new experiments, could be predicted by a single model of binaural unmasking which had been designed to deal with the unmasking of complex sounds, without reference to dichotic pitches (Culling and Summerfield, 1995). The model was essentially a multichannel version of Durlach's equalization cancellation EC model (Durlach, 1960, 1962), although with the important caveate that the model should select equalization delays in each frequency channel independently. In many cases, Culling *et al.* (1998a, c) contrasted the performance of this model with the performance of competing models, based on selective direction of attention rather than on binaural unmasking (e.g., Bilsen, 1977; Raatgever, 1980; Raatgever and Bilsen, 1986) or different implementations of the EC model, which do not use different equalization delays in different frequency channels (e.g., Bilsen and Goldstein, 1974; Klein and Hartmann, 1986). In particular, Culling *et al.* (1998c) showed that the spectra which the model recovered from Fourcin-pitch stimuli corresponded to measurements of the perceived pitches which had been reported in the literature, while other models made qualitatively different predictions. A mathematical analysis showed that the model should produce the correct pitch for any configuration of two noises. Although this analysis showed that the model produces the correct pitches, it did not demonstrate that the model makes those predictions for all pitches which can be heard and for only those pitches. In other words, it did not predict the existence region of the Fourcin pitch.

The purpose of the current investigation was to extend the case developed in the earlier papers by exploring the existence region of the Fourcin pitch experimentally and comparing it with that predicted by Culling and Summerfield's modified EC (mEC) model. Notwithstanding a recent addition to the range of pitches that has been reported in Fourcin-pitch stimuli (Raatgever *et al.*, 1998), it is assumed throughout this article that the nature of the pitch which is evoked has been firmly established by others and that it is the detectability/salience of this dichotically evoked sound under different interaural configurations that most merits further investigation.

A. The Fourcin pitch

The Fourcin pitch can be demonstrated by presenting listeners with more than one (independent) broadband noise simultaneously and binaurally, over headphones. Each noise

[a)]Electronic mail: cullingj@cardiff.ac.uk

has a different interaural delay, and the differences in interaural delay between the different noises must be of the order of milliseconds. The perceived pitch is related to the difference in interaural delay between the different noises (Fourcin, 1958, 1970; Bilsen and Wesdorp, 1974; Bilsen, 1977) and decreases with increasing difference in delay. The pitch is ambiguous unless one of the two noises also has an interaural phase shift of 180°, whereupon the period of the pitch will be equal to the difference in delays. Since two or more noises are used and since they can each have different interaural delays, the Fourcin pitch has many parameters which may be varied. It was therefore necessary to constrain the current investigation to the most interesting manipulations. Those selected were (1) the number of noises employed and (2) the difference in interaural delay. These parameters were explored in experiments 1 and 2, respectively. Those aspects of the stimulus configuration which were not currently under test were always designed to make the pitch maximally clear (in the absence of empirical data on the effects of these parameters on the clarity of the pitch, assumptions were made based on the mechanism of the mEC model) and unambiguous; the delays were spaced evenly in interaural delay, with a symmetrical overall pattern (e.g., −2, 0 and +2 ms for 3 noises or −3, −1, +1, and +3 ms for four noises) and with alternate noises interaurally inverted. Since Fourcin (1970, p. 322) remarked that the phenomenon is most clearly heard when the pitch changes, the stimuli were also of an extended duration with continual or repeated movements in pitch, giving listeners time to pick the movements up.

Fourcin (1958, 1970) provides the only published reports of the use of more than two noises to generate the Fourcin pitch. Fourcin used up to five noises, which he spaced equally in interaural delay (e.g., −4, −2, 0, +2, +4 ms) with alternate noises inverted at one ear. Under these conditions, Fourcin observed that the clarity of the pitch did not improve with the number of noises. Experiment 1 provides the first formally presented data on this dimension of the existence region, using up to eight noises.

The extent of the existence region of the Fourcin pitch, in terms of the binaural configurations for which a pitch can or *cannot* be heard, has not been reported previously. However, various studies have shown that the pitch *can* be matched against other forms of pitch-evoking stimuli using differences in delays in the range 1–5 ms (Fourcin, 1958), 2–11 ms (Bilsen and Goldstein, 1974), and 2–9 ms (Bilsen and Wesdorp, 1974; Bilsen, 1977). Clearly the pitch exists in these regions, but the breakdown of the phenomenon outside them has not been documented. Experiment 2 seeks to explore the limits of the existence region.

B. The mEC model

Culling and Summerfield's (1995) mEC model is a modified version of Durlach's EC model. Briefly, the left- and right-channel wave forms are filtered by twin gammatone filterbanks (Patterson *et al.*, 1987, 1988) and processed by the Meddis (1986, 1988) hair-cell model. Then, corresponding frequency channels from the two sides are equalized first in level and then (so far as possible) in delay, before they are subtracted one from the other. Equalization delays of up to ±5 ms are permitted and the best delays are selected independently in each frequency channel. The residual energy in each frequency channel is a measure of the binaural activity at that center frequency and a plot of rms residual energy as a function of center frequency forms the ''recovered spectrum''. See Culling *et al.* (1998a) for a more detailed description. The model gives a measure of the deviation in the interaural correlation from 1.0 at each frequency. Such deviations in interaural correlation are widely thought to be the percentual cues underlying binaural masking release (Gabriel and Colbum, 1981; Durlach *et al.*, 1986; Koehnke *et al.*, 1986; Jain *et al.*, 1991; Culling and Summerfield, 1995; Bernstein and Trahiotis, 1992, 1996a, b).

I. EXPERIMENT 1

Experiment 1 measured the detectability of the Fourcin pitch as a function of the number of noises used in generating the pitch, termed the ''order'' of the Fourcin pitch. Orders of 2–8 were used. Listeners were presented in each trial with an 11-pitch sequence, which traversed a wide range of frequencies in approximately half-octave steps, and were instructed to discriminate the direction of pitch movement.

A. Stimuli

To make a single Fourcin-pitch sequence, a series of Fourcin pitches were generated and then concatenated together. Each pitch was generated in the following way. Between two and eight 409.6-ms broadband noises (0–10 kHz) were generated digitally at a 20-kHz sampling rate. A copy of each of the noises was delayed, using frequency-domain filtering. The original and copy were combined into a stereo file. The left channel of every second stereo file was inverted and the files created for each noise were summed. The interaural delays were evenly spaced at intervals of the period of the desired pitch period and were symmetrically distributed about zero delay. These files could then be concatenated in both ascending and descending order of pitch, to create ascending and descending sequences with approximately half-octave steps between successive notes. After concatenation, the overall stimulus was gated with a 10-ms raised-cosine rise/decay function.

Since separately generated stimuli were directly concatenated, the transition between one pitch and the next was accompanied by a brief period (up to 5 ms) during which the noise in each channel was uncorrelated. This short period of interaural decorrelation was not noticeable in the finished stimuli and disrupted perception of the pitches less than gating the sound off and then back on between each pitch. Akeroyd and Summerfield (1999) have measured the threshold duration for the detection of burst a of decorrelation in otherwise correlated noise and found that only one of their six listeners could detect bursts of decorrelation shorter than 5 ms.

Figure 1 shows the broadband cross-correlation functions for Fourcin pitches of order 2–8, which demonstrates this arrangement. The maintenance of symmetry meant that for an odd order, one noise was at zero delay, whereas for an even order, two noises lay equally spaced on either side. In order to maintain maximal perceptual salience for an unam-

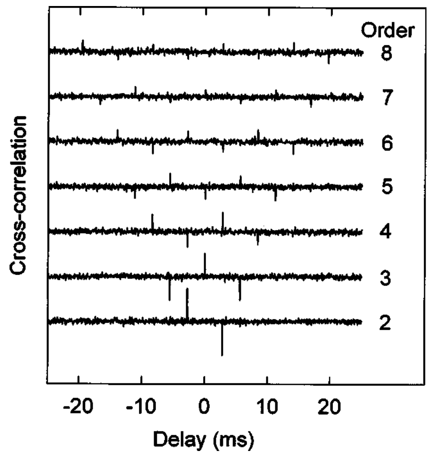

FIG. 1. Broadband cross-correlation functions for stimuli with between two and eight noises (order 2–8) in experiment 1. The interaural delays of the constituent noises are distributed at 5.6-ms intervals, corresponding to a perceived pitch of 179 Hz The cross-correlation used an exponentially tapering window with 50-ms time constant.

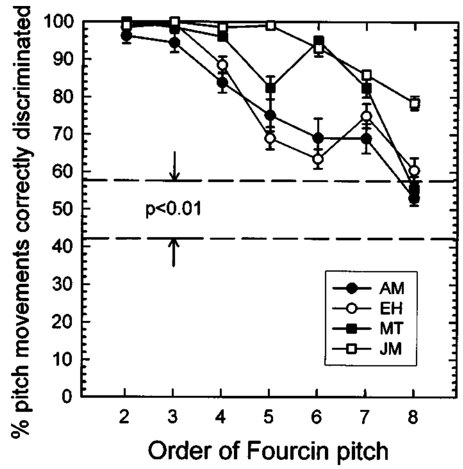

FIG. 2. Percentage of upward/downward pitch movements correctly discriminated as a function of the number of noises used to make the stimuli (the order). The data from four listeners is plotted separately with the different symbols. The error bars are standard errors of the mean for ten runs.

biguous pitch, every other noise was interaurally phase shifted by 180° (inverted). The levels of the constituent noises were adjusted so that each noise in a given stimulus was the same level and their combined power was the same for each condition. Five examples of each sequence were generated for each of the 7 conditions (orders 2–8) and, for each example, the 11 pitches were concatenated in both ascending and descending sequences. So, there were $5\times7\times2=70$ stimuli in all.

B. Procedure

Four listeners with no known hearing problems participated in experiment 1. They were trained without trial-by-trial feedback on Fourcin-pitch stimuli of the kind used in the experiment until they could discriminate ascending from descending sequences with 90% accuracy. Some listeners picked up the pitch quickly, while others were trained for many hours. Listeners were not selected for aptitude in the task. During the early stages of training, listeners were given sets of stimuli in which Fourcin pitches were interspersed with ''filler'' stimuli which were designed to sound similar, but be more perceptually salient than the Fourcin-pitch stimuli. Using these filler stimuli to assist listeners in training was found to be essential for two of the four listeners. Various filler stimuli were used, but the most effective were bases on the MPS pitch (Bilsen, 1977).

The listeners attended five 1-h sessions, during each of which they completed two experimental runs. All the stimuli were presented twice in a randomized sequence during each run, so that each run yielded a score out of 20 for each condition.

C. Results

Figure 2 shows the percentage of stimuli for which each of the four listeners correctly discriminated ascending from descending sequences as functions of the order of the Fourcin pitch. The figure also shows thresholds for statistical significance ($p<0.01$) for a single listener's data in a single condition [derived from binomial probability: 200 trials, $p(\text{correct})=0.5$]. All four listeners showed a progressive decline in discrimination accuracy with increasing order. By order 8, only two listeners performed significantly above chance ($p<0.01$).

D. Modeling

Figure 3 shows the spectra recovered by the mEC model from the stimuli used in experiment 2. The model was run on portions of the stimulus where the perceived pitch should be 179 Hz. The model correctly predicts that listeners will perceive a pitch of that frequency, but like the listeners, the model detects less evidence of a pitch as the order of the Fourcin pitch is increased. For order 2, the output of the model is well modulated, but, as the order of the pitch increases, the modulation decreases and the recovered spectrum becomes more and more ragged. For order 8 the output spectrum is virtually flat.

The most likely reason for the decline in salience is that unlike autocorrelation, the principle of superposition does

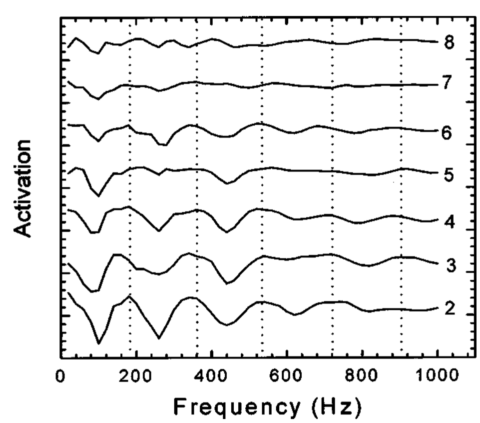

FIG. 3. Spectra recovered by the mEC model from the segment of noise whose Fourcin pitch has a perceived frequency of 179 Hz in experiment 1. Dotted vertical lines indicated the frequencies of harmonics of 179 Hz. Each spectrum is for a stimulus generated using a different number of noises (the order), indicated by the numeric labels 2–8.

not apply to the cross-correlation of finite-duration signals. That is to say that when two stimuli with different cross-correlation functions are added together, the cross-correlation of the resulting stimulus is *not* the sum of the two separate cross-correlation functions. The broadband cross-correlation functions shown in Fig. 1 show clearly that the consequence of adding extra noises with different interaural delays is that the cross-correlation (measured over a fixed interval of time) becomes weaker at the delays of the existing noises. By the time eight noises have been added, the spikes in the cross-correlation marking the delay of each individual noise are barely discernible from random fluctuations in the function. This situation contrasts with the monaural effect of echo pitch (also known as ''rippled noise'' or ''repetition pitch'') for which the addition of extra noises at regularly spaced delays increases pitch strength (e.g., Yost *et al.*, 1996).

II. EXPERIMENT 2

Experiment 2 investigated the classical existence region of the second-order Fourcin pitch, i.e., the range of pitches which can be heard. Investigating this aspect of the phenomenon was more difficult than the effect of order, since the stimulus could no longer be swept in an extensive sequence of pitches through several octaves. These pitch sequences were very helpful to listeners in enabling them to detect the pitch.

In pilot experiments, shorter sequences were employed that covered a smaller frequency range. However, even the most sensitive listeners had great difficulty detecting the pitch from such stimuli. As a result, the final design of experiment 3 included three features designed to help the listeners tune-in to the correct pitch while performing the task. First, the stimuli at each pitch frequency were presented in separate blocks, and the start of each block was preceded by a monaural repetition pitch stimulus with a pitch equal to the pitch frequency under test. Second, the Fourcin pitch stimuli in each block were interspersed with an equal number of modified multiple-phase-shift (MPS pitch) stimuli (Bilsen, 1976). These ''filler'' stimuli were designed to sound similar to, but be slightly more salient than, the Fourcin pitches. Third, the first two stimuli in a given block were always such MPS fillers.

A. Stimull

Fourcin pitches were generated in a similar manner to the second-order Fourcin-pitch stimuli from experiment 1. Each stimulus was constructed from eight 409.6-ms segments which had expected pitches 5% above and 5% below the pitch frequency under test. These segments were concatenated into sequences which either alternated through four cycles high–low–high–low... or low–high–low–high... . The stimuli were then gated with 10-ms raised-cosine onset/offset ramps. The same 11 pitches were tested as were used in Experiment 1, i.e., 31, 45, 63, 89, 125, 179, 250, 357, 500, 714 and 1000 Hz. Five examples of each stimulus were made. With 11 frequencies×5 examples×2 alternations, there were 110 Fourcin-pitch stimuli.

The filler stimuli were based on the MPS pitch described by Bilsen (1976). The MPS pitch is made by introducing a series of 360° interaural phase transitions at harmonic frequencies into otherwise diotic noise. In other words, it contains a harmonic series of Huggins' pitches (Cramer and Huggins, 1958). The pitch is highly salient if made with transition bandwidths which are 6% of the transition frequencies. In order to make the MPS pitches less salient they were created with 1% transition bandwidths. The narrower bandwidths reduced the strength of the pitch somewhat, but the pitch was still strong and the stimulus still differed from the Fourcin pitch perceptually; for the MPS pitch the noise is centered in the head while the pitch is either lateralized or diffuse, whereas for the Fourcin pitch, neither component of the percept is well localized. In order to diffuse the intracranial position of the noise, and also to reduce the pitch salience further, the noise was partially interaurally decorrelated: the phases of each component of the noise were offset at one ear form their original values by rectangularly distributed offsets in the range ±30°. The resulting stimuli were still easy to discriminate from Fourcin-pitch stimuli, but were sufficiently similar for the purposes of the experiment. In common with the Fourcin pitches, the resulting sounds were assembled into alternating-pitch stimuli and five examples of each stimulus were created. The cue tone which preceded each block was a single 409.6-ms monaural repetition pitch (Basset and Eastmond, 1964; Bilsen, 1966). This sound was generated by creating a 409.6-ms Gaussian noise, delaying a copy of this noise by the period of the pitch under test, and adding the delayed noise to the original. The resulting stimulus has a clear pitch with a noisy timbre.

B. Procedure

The same four listeners attended 11 1-h sessions, during each of which they completed two experimental runs. Each run was composed of 11 blocks of 20 stimuli. Each block was preceded by a single monaural-repetition-pitch cue tone. The noisy timbre of such a cue tone was thought more suitable than a pure tone as a cue for the stimuli which were to follow. The pitch used in successive blocks either ascended or descended throughout a run, except when the end of the scale had been reached whereupon the pitch jumped to the other end of the scale. The starting point varied progressively from one run to the next, so that each block would occupy each position in the sequence in different runs. For eleven runs the blocks ascended in pitch and for eleven it descended; two subjects did blocks of ascending pitch for the first eleven runs while the other two did blocks of descending pitch.

The 20 stimuli in a block were each of the 10 Fourcin-pitch stimuli (5 examples×2 alternations) and each of the corresponding fillers. The listeners' task was to listen to the alternation of high and low pitch and determine whether the sequence was high–low–high–low..., or the reverse. The four cycles of alternation were important, because listeners rarely heard the entire sequence, and found the optimal strategy was to wait until they picked up the alternation and then decide whether the final sound was high or low.

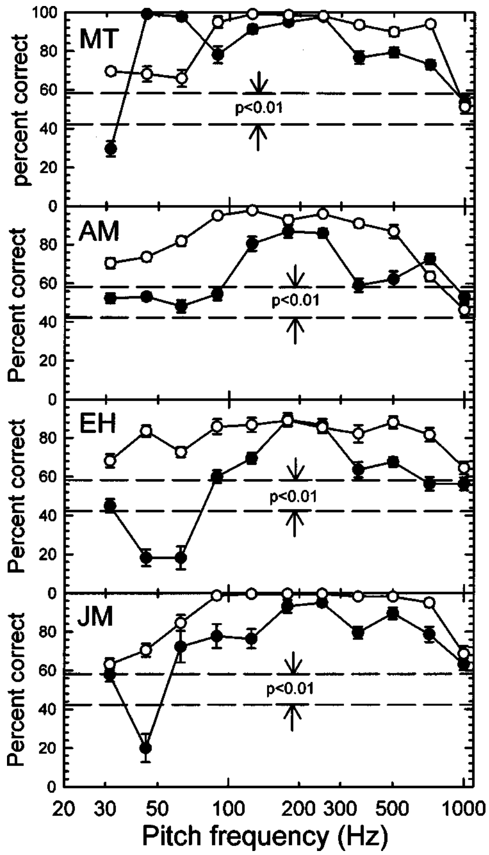

FIG. 4. Discrimination of high–low vs low–high alternation of Fourcin pitches (closed symbols) as a function of Fourcin pitch frequency for the four listeners in experiment 2. The dashed horizontal lines represent the thresholds of statistical significance ($p<0.01$) for individual data points.

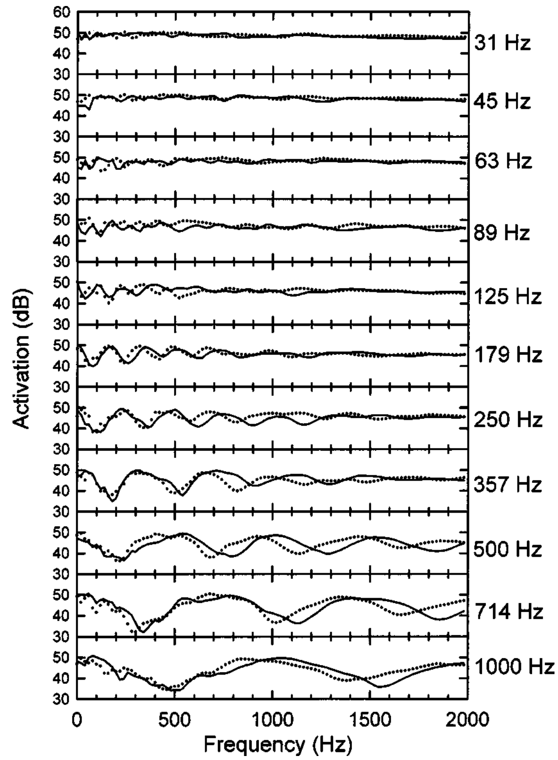

FIG. 5. Spectra recovered by the mEC model for Fourcin pitches at each pitch frequency used in experiment 2. Each panel shows the recovered spectra for the two Fourcin-pitch stimuli that listeners compared in experiment 2 for the indicated nominal pitch frequency. These stimuli had expected pitches 5% above (solid lines) and 5% below (dotted lines) the nominal pitch frequency.

C. Results

Figure 4 shows the effect of pitch frequency on listeners' ability to discriminate between high–low and low–high alternation of both the Fourcin pitch (closed symbols) and the modified MPS pitch (i.e., the fillers, open symbols). Each panel shows the results for one listener. The dotted lines show thresholds for significant deviations from chance ($p<0.01$) for each data point [from binomial probability: 220 trials, $p(\text{correct})=0.5$]. The MPS-pitch data are shown only to illustrate the fact that they were more easily discriminated than the Fourcin pitch stimuli.

Taking first the features of the Fourcin-pitch data which the listeners show in common, the pitch appears to be most salient around 125–250 Hz and is very difficult to hear for all listeners at the two extremes of the stimulus set (31 and 1000 Hz). All the listeners show a more or less monotonic decline in discrimination performance between 250 and 1000 Hz. The listeners performance at frequencies between 31 and 125 Hz is more variable. In particular, listeners MT, EH, and JM all show performance which is significantly *below* chance for one or more pitch frequencies.

D. Modeling

Figure 5 show the spectra recovered by the model for examples of the two Fourcin-pitch stimuli which were used in each condition of experiment 2. The two pitches should differ by 10%, but listeners had difficulty detecting the direction of movement in experiment 2 when the pitch was lower than 89 Hz or higher than 357 Hz (the limits of this range varying across listeners). In order for the model to predict that the direction of a given pitch change should be discriminable, the corresponding panel of Fig. 5 should show peaks in the two curves which are displaced from each other in frequency, indicating harmonic series with different fundamental frequencies. The model recovers spectral peaks at appropriate harmonic frequencies for Fourcin pitches above about 100 Hz F_0. Unlike the listeners, no deficit in its performance is evident for pitch frequencies above 250 Hz.

III. DISCUSSION

A. The empirical existence regions

Experiment 1 shows that the Fourcin pitch becomes progressively less detectable as the number of noises used to generate it is increased (Fig. 3). Experiment 2 shows that pitches in the 125–250-Hz region (generated using interaural delays of 4–8 ms) are most easily detected, but that deviations from chance performance are displayed by the majority of listeners at all frequencies from 45 to 714 Hz. In the cases where listeners scored below chance, the most likely explanation is that the listeners were unable to hear all the harmonics of the pitch and that they picked up different harmonics during the high and low-pitch phases of the stimuli; if, for instance, decisions were based on single harmonics of different number, it is not surprising that the wrong pitch movement was perceived. This explanation is supported by the fact that listeners reported a mismatch between the cue tones used in the conditions with low pitch frequencies and the pitches which they heard in the test stimuli. The test stimuli had much higher pitches, which were consistent with the detection of single high-numbered harmonics. Since listeners did detect evidence of the Fourcin pitch which influenced their decisions in a consistent manner, these deviations from chance may be regarded as detection of the pitch.

B. The predicted existence regions

With the exception of the decline in salience for high pitch frequencies in experiment 2, the spectra recovered by the model predicts the pattern of results displayed by the listeners in both experiments. The modulation of the model's output spectrum is affected by the order of the Fourcin pitch. The spectra become increasingly featureless as the number of noises is increased, mirroring the decline in the listeners' ability to discriminate different pitch movements in these conditions. The spectra produced by the model in response to very low pitch frequencies, where listeners have difficulty hearing the pitch, are quite flat (Fig. 5); they become better modulated at higher frequencies where listeners performance is at its best (125–250 Hz), but unlike the listeners, the model seems to work well (produce pairs of spectra with different harmonic structures) up to the highest pitch frequency used (1000 Hz). In contrast, the listeners show a gradual decline in their ability to discriminate different pitch movements at high pitch frequencies.

C. A revised model

The mismatch between model and data for high Fourcin pitches is probably attributable to the mEC model's lack of internal noise. The internal noise in Durlach's original formulation was principally intended to model the reduction in size of the binaural masking level difference with increasing frequency. The mEC model was designed for the purpose of making qualitative rather than quantitative predictions, and so does not feature internal noise as used in Durlach's original formulation of equalization cancellation. Consequently it performs too well at high frequency. Bernstein and Trahiotis (1992, 1996a, b) have recently shown that the decline in binaural masking release above 1500 Hz can be modeled by including peripheral nonlinearities which encode only the envelope of the stimulus wave form at higher frequencies. The model might be revised by adding internal noise or by changing its peripheral nonlinearities. [The existing peripheral nonlinearities, provided by the Meddis (1986, 1988) hair cell model, provide a degree of desynchronization to the carrier frequency at high frequencies, but this loss of synchrony is rather less than would be necessary for accurate predictions of binaural phenomena.] However, since contemporary models of binaural unmasking interpret binaural detection of masked sounds as resulting from the detection of interaural decorrelation of the stimulus, one can, equivalently, use empirical measurements of listeners' sensitivity to interaural decorrelation to predict their ability to detect sounds in noise and, in this case, to directly detect the interaural decorrelation which is present in dichotic-pitch stimuli.

Culling *et al.* (1998b, 2000) have collected data on listeners' sensitivity to interaural decorrelation. They measured listeners sensitivity to changes in correlation of one subband embedded within a broadband correlated noise. This sensitivity was expressed in terms of cumulative d' and a family of functions was derived which relate correlation to cumulative d' at each frequency (see the Appendix). These functions can be used to transform interaural correlations onto a perceptual salience scale. By measuring the interaural correlation of each frequency channel and calculating the cumulative d' for the difference between a correlation of 1 and each interaural correlation, ρ, $d'_{(1,\rho)}$ can be calculated. $d'_{(1,\rho)}$ represents the perceptual salience of the interaural decorrelation at that frequency, so a spectrum of values derived from different frequency channels constitutes a perceptually scaled binaurally recovered spectrum.

The revised model is similar to the mEC model in that it permits the application of delays of up to 5 ms, which are independently selected for each frequency channel. Like the mEC model, it assumes similar frequency selectivity to the monaural system (see Kohlrausch, 1988; Kollmeier and Holube, 1992). So the stimuli are still passed through a pair of gamma-tone filterbanks (Patterson *et al.*, 1987, 1988). As before, the wave forms are optimally delayed, but rather than canceling the corresponding left- and right-ear frequency channels these wave forms are correlated within an exponentially decaying window. [For the Fourcin-pitch stimuli used in experiment 2, a delay of $10^6/4f$ μs (where f is the channel center-frequency) must be applied to either the left- or right-hand channel in order to achieve maximal correlation. The side to be delayed alternates with increasing channel frequency, switching whenever f is a multiple of the pitch frequency.] The window was exponentially decaying with a 100-ms time constant. The equivalent rectangular duration of the window (also 100 ms) was thus brought into line with recent measurements of the binaural temporal window (Culling and Summerfield, 1998; Akeroyd and Summerfield, 1999). The resulting product–moment correlations can then be transformed according to the measured sensitivity of listeners to deviations in correlation from one ($d'_{(1,\rho)}$).

Figure 6 shows $d'_{(1,\rho)}$ as a function of frequency for examples of the two Fourcin-pitch stimuli which were used in each condition of experiment 2. The two pitches should differ by 10%, but listeners had difficulty detecting the direction of movement in experiment 2 when the pitch was lower than 89 Hz or higher than 357 Hz (the limits of this range varying across listeners). The d'-based model appears to make this prediction quite accurately. In order for the model to predict that the direction of a given pitch change should be discriminable, the corresponding panel of Fig. 6 should show peaks in the two curves which are displaced from each other in frequency, indicating harmonic series with different fundamental frequencies. None of these pairs of curves are identical, indicating that there may always be some audible difference between the two stimuli. However, systematic shifts in the peaks, indicating the correct differences in pitch, are only apparent for the middle range of pitch frequencies,

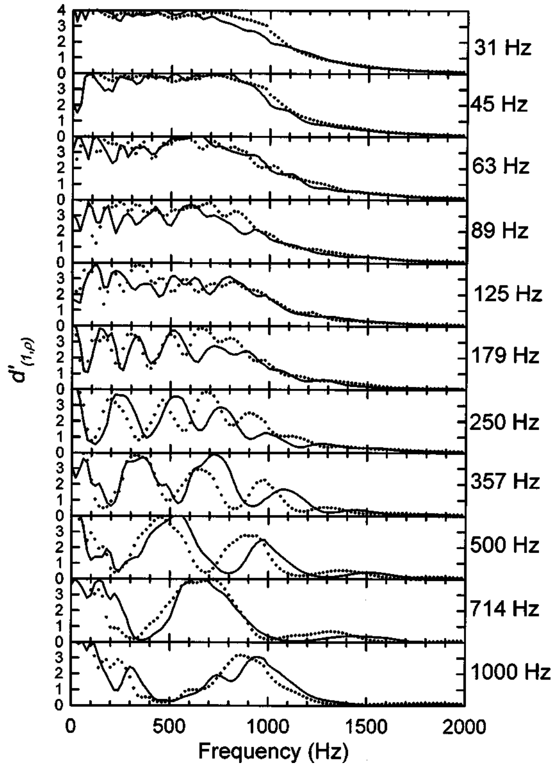

FIG. 6. As in Fig. 5, but showing perceptually scaled spectra for each Fourcin pitch in experiment 3. $d'_{(1,\rho)}$ is the expected sensitivity of listeners to the decorrelation of the stimulus within each frequency channel. The transform between ρ and $d'_{(1,\rho)}$ was taken from Culling *et al.* (2000). ρ was calculated on the corresponding frequency channels emerging from twin gammatone filterbanks (Patterson *et al.*, 1987, 1988) fed with the left- and right-hand channels of the stimuli.

where all the listeners were able to make the discrimination.

The revised model was also run on the stimuli from experiment 1 in order to check that it can still correctly predicted a decline in salience with increasing order. The results of this test are shown in Fig. 7 in identical format with Fig. 3 for easy comparison. The results are very similar in this case.

IV. CONCLUSIONS

The results of the two experiments reported here are in broad agreement with the predictions of the mEC model and are therefore consistent with the view that the Fourcin pitch is an illusion of binaural unmasking. Where disagreement between the mEC model and the data exists, a similar modeling method which incorporates measurements of the discriminability of different levels of correlation gives more accurate predictions.

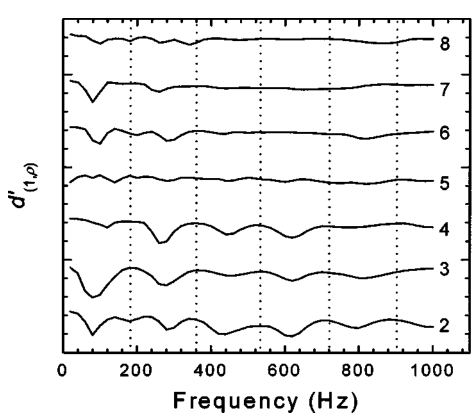

FIG. 7. As in Fig. 3, but showing perceptually scaled decorrelation spectra similar to those of Fig. 6.

ACKNOWLEDGMENTS

The author would like to thank Quentin Summerfield for comments on a previous draft of this manuscript and two anonymous reviews for their thorough and thoughtful reviews. Work supported by M.R.C.

APPENDIX

Sensitivity to interaural decorrelation has been summarized by Culling *et al.* (1998b, 2000) as follows. The growth in perceptual salience, measured using cumulative d', as a function of deviation in correlation from one, $d'_{(1,\rho)}$, can be described by

$$d'_{(1,\rho)} = e^{(k+n)} - e^{(k\rho+n)}. \tag{A1}$$

The parameters of this function varied with frequency according to the following logistic functions. The parameters of these logistic functions have been updated in line with additional data collected since Culling *et al.* (1998b).

$$k = \frac{4.68}{1+e^{0.0027(f-666)}} + 0.0027, \tag{A2}$$

$$n = \frac{3.17}{1+e^{-0.0047(f-560)}} - 2.75. \tag{A3}$$

Akeroyd, M. A., and Summerfield, Q. (**1999**). ''A binaural analog of gap detection: Gap thresholds,'' J. Acoust. Soc. Am. **105**, 2807–2820.

Basset, I. G., and Eastmond, E. J. (**1964**). ''Echolocation: Measurements of pitch versus distance for sounds reflected from a flat surface,'' J. Acoust. Soc. Am. **36**, 911–916.

Bernstein, L. R., and Trahiotis, C. (**1992**). ''Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies,'' J. Acoust. Soc. Am. **91**, 306–316.

Bernstein, L. R., and Trahiotis, C. (**1996a**). ''On the use of the normalized correlation as an index of interaural envelope correlation,'' J. Acoust. Soc. Am. **100**, 1754–1763.

Bernstein, L. R., and Trahiotis, C. (**1996b**). ''The normalized correlation: Accounting for binaural detection across center frequency,'' J. Acoust. Soc. Am. **100**, 306–316.

Bilsen, F. A. (**1966**). ''Repetition pitch: monaural interaction of a sound with the repetition of the same, but phase shifted sound,'' Acustica **17**, 295–300.

Bilsen, F. A. (**1976**). ''A pronounced binaural pitch phenomenon,'' J. Acoust. Soc. Am. **59**, 467–468.

Bilsen, F. A. (**1977**). ''Pitch of noise signals: Evidence for a 'central spectrum,''' J. Acoust. Soc. Am. **61**, 150–161.

Bilsen, F. A., and Goldstein, J. L. (**1974**). ''Pitch of dichotically delayed noise and its possible spectral basis,'' J. Acoust. Soc. Am. **55**, 292–296.

Bilsen, F. A., and Wesdorp, C. (**1974**). ''Dichotic pitch phenomena and their common basis,'' in *Eighth International Conference on Acoustics* (Goldcrest, London).

Cramer, E. M., and Huggins, W. H. (**1958**). ''Creation of pitch through binaural interaction,'' J. Acoust. Soc. Am. **30**, 858–866.

Culling, J. F., Colburn, H. S., and Spurchise, M. (**2000**). ''Interaural correlation sensitivity,'' unpublished.

Culling, J. F., Marshall, Q., and Summerfield, D. H. (**1998a**). ''Dichotic pitches as illusions of binaural unmasking II: The Fourcin pitch and the dichotic repetition pitch,'' J. Acoust. Soc. Am. **103**, 3527–3539.

Culling, J. F., Spurchise, M., and Colburn, H. S. (**1998b**). ''Interaural correlation sensitivity,'' J. Acoust. Soc. Am. **103**, 3081–3082.

Culling, J. F., and Summerfield, Q. (**1995**). ''Perceptual separation of competing speech sounds: Absence of across-frequency grouping by common interaural delay,'' J. Acoust. Soc. Am. **98**, 785–797.

Culling, J. F., and Summerfield, Q. (**1998**). ''Measurements of the binaural temporal window using a detection task,'' J. Acoust. Soc. Am. **103**, 3540–3553.

Culling, J. F., Summerfield, Q., and Marshall, D. H. (**1998c**). ''Dichotic pitches as illusions of binaural unmasking II: The Fourcin pitch and the dichotic repetition pitch,'' J. Acoust. Soc. Am. **103**, 3509–3526.

Durlach, N. I. (**1960**). ''Note on the equalization-cancellation theory of binaural masking level differences,'' J. Acoust. Soc. Am. **32**, 1075–1076.

Durlach, N. I. (**1962**). ''Binaural signal detection: Equalization and cancellation theory,'' J. V. Tobias (Ed.), *Foundations of Modern Auditory Theory*, Vol. II (Academic, New York).

Durlach, N. I., Gabriel, K. J., Colburn, H. S., and Trahiotis, C. (**1986**). ''Interaural correlation discrimination II: Relation to binaural unmasking,'' J. Acoust. Soc. Am. **79**, 1548–1557.

Fourcin, A. J. (**1958**). ''Speech perception and bandwidth compression,'' Signals Research and Development Establishment Report No. 1126.

Fourcin, A. J. (**1970**). ''Central pitch and auditory lateralization,'' R. Plomp and G. F. Smoorenburg (Eds.), *Frequency Analysis and Periodicity Detection in Hearing* (Sijthoff, The Netherlands).

Gabriel, K. J., and Colburn, H. S. (**1981**). ''Interaural correlation discrimination I: Bandwidth and level dependence,'' J. Acoust. Soc. Am. **69**, 1394–1401.

Jain, M., Gallagher, D. T., Koehnke, J., and Colburn, H. S. (**1991**). ''Fringed correlation discrimination and binaural detection,'' J. Acoust. Soc. Am. **90**, 1918–1926.

Klein, M. A., and Hartmann, W. M. (**1986**). ''Binaural edge pitch,'' J. Acoust. Soc. Am. **70**, 51–61.

Koehnke, J., Colburn, H. S., and Durlach, N. I. (**1986**). ''Performance in several binaural-interaction experiments,'' J. Acoust. Soc. Am. **79**, 1558–1562.

Kohlrausch, A. (**1988**). ''Auditory filter shape derived from binaural masking experiments,'' J. Acoust. Soc. Am. **84**, 573–583.

Kollmeier, B., and Holube, I. (**1992**). ''Auditory filter bandwidths in binaural and monaural listening conditions,'' J. Acoust. Soc. Am. **92**, 1889–1901.

Meddis, R. (**1986**). ''Simulation of mechanical to neural transduction in the auditory receptor,'' J. Acoust. Soc. Am. **79**, 702–711.

Meddis, R. (**1988**). ''Simulation of auditory-neural transduction: Further studies,'' J. Acoust. Soc. Am. **83**, 1056–1063.

Patterson, R. D., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1987**). ''An efficient auditory filterbank based on the gammatone function,'' paper presented to the IOC speech group on auditory modeling at the Royal Signal Research Establishment, 14–15 December.

Patterson, R. D., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1988**). ''Spiral vos final report, Part A: The auditory filter bank,'' Cambridge Electronic Design, Contract Report No. (APU 2341).

Raatgever, J. (**1980**). ''On the binaural processing of stimuli with different interaural phase relations,'' doctoral dissertation, Delft University.

Raatgever, J., and Bilsen, F. A. (**1986**). ''A central theory of binaural processing: Evidence from dichotic pitch,'' J. Acoust. Soc. Am. **80**, 429–441.

Raatgever, J., Bilsen, F. A., and Mungra, R. (**1998**). ''New experiments beyond the traditional Fourcin pitch range,'' J. Acoust. Soc. Am. **103**, 2767 (A).

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