On the pitches of the components of a complex tone

William Morris Hartmann and Sandra L. Doty[a)]
Department of Physics, Michigan State University, East Lansing, Michigan 48824

(Received 11 April 1995; revised 25 July 1995; accepted 2 August 1995)

The pitches of the harmonics (numbers 1, 2, 3, 4, 5, 7, 9, and 11) of a complex tone were measured in a matching experiment. The harmonics to be matched were mistuned (8% or less) either positively, or negatively, or not at all. For all mistuned harmonics and all listeners the matching pitches were found to be exaggerations of the mistunings, i.e., the data exhibited pitch shifts with the same sign as the mistunings. This result is shown to be contrary to place models of pitch perception, such as the spectral pitch algorithm of Terhardt, in which pitch shifts are caused by the interaction of excitation patterns for the individual harmonics. An alternative model, in which pitch is determined by neural timing, also fails to account for the data. However, a hybrid model, combining effects of excitation pattern interaction with neural timing, does agree with most of the data. © *1996 Acoustical Society of America.*

PACS numbers: 43.66.Hg, 43.66.Ba

INTRODUCTION

A complex periodic tone, such as an idealized steady vowel or musical instrument tone, is normally heard as a single entity. The harmonics of the tone are not heard individually; rather they are integrated by the auditory system to determine the pitch, the loudness, and the tone color of the entity. Under special circumstances, however, the individual harmonics can be heard (Helmholtz, 1885). Updates of these circumstances were given a century later (McAdams, 1984; Hartmann, 1988). The question then arises, how does the percept of a harmonic segregated from its complex tone context compare with the percept of that harmonic when it is presented alone, as a single sine tone?

In 1971 Terhardt reported experimental results showing that the pitch of a harmonic in a complex tone is different from the pitch of an equal-level sine tone having the frequency of that harmonic. His data showed that the pitch of the fundamental component is shifted downward compared to the pitch of the corresponding sine. The pitches of all other harmonics were shifted upward. (Data also appear in Terhardt, 1972, 1979.)

The discovery of these shift effects influenced subsequent research in pitch perception, and did so for several reasons. First, it is possible that the auditory system derives the virtual (low) pitch of a complex tone from the (shifted) pitches of the harmonics (Terhardt, 1974). If this is true then shift effects are potentially involved in every aspect of complex tone pitch. Second, pitch shifts of all kinds provide powerful experimental methods to learn more about the process of pitch perception itself (e.g., Houtsma, 1981).

In 1979 Terhardt presented a semiempirical algorithm whereby the pitch shifts can be calculated. (See also Terhardt *et al.*, 1982b.) The algorithm owes at least as much to actual measurements of pitch shifts as it does to first principles. Such principles as are involved have the strong flavor of a place theory, beginning with a neural excitation pattern as a function of a tonotopic coordinate (Terhardt, 1972). The excitation might be understood as driven neural firing rate or synchronous firing rate. The tonotopic coordinate might be the critical-band scale or a place along the basilar membrane.

In this theory, the pitch of each harmonic in a complex tone is determined by a corresponding hump in the excitation pattern. The shape of each hump is modified by excitation due to neighboring harmonics because of mutual masking. Such modifications are the origins of the pitch shifts. For the fundamental, the masking can only be due to higher harmonics. The excitation for the fundamental is masked from above and therefore the pitch of the fundamental must be lower in the complex tone than in isolation. Thus the theory predicts a negative pitch shift for the fundamental, in agreement with experiment. For a higher harmonic the excitation is affected by harmonics both above and below. The interaction is easiest to understand for spectrally resolved harmonics (approximately two through six) where the prediction is that there should be a positive pitch shift, again in agreement with Terhardt's experiment.

In 1983 Peters *et al.* made a search for pitch shifts in the harmonics of a complex tone, using experiments similar to those reported by Terhardt in 1971. Although they tried several different methods, both adjustment and forced choice, they were unable to measure statistically significant pitch shifts. This result cast some doubt on the generality of both the experiments and the algorithm described by Terhardt and his colleagues.

In 1989 Hartmann *et al.* obtained data pertinent to the pitches of the harmonics in a complex tone. The experiment, called the "mistuned-harmonic matching" experiment (MHM), studied complex tones in which one harmonic had been mistuned from its correct value. The goal of the experiment was to determine the amount of mistuning required for a listener to hear out the mistuned harmonic as a segregated tone. The method required listeners to tune a sine tone to match the pitch of the mistuned harmonic.

A by-product of the MHM experiment was a set of data on the pitches of the components of a complex tone. The

---

[a)]Present address: Department of Science and Math, GMI Engineering and Management Institute, 1700 W. Third Ave., Flint, MI 48504.

 0001-4966/96/99(1)/567/12/$6.00

main difference between that experiment and the experiments by Terhardt or by Peters *et al.* was that the harmonics were not exact but were mistuned. The data showed significant pitch shifts, as might have been expected from Terhardt's experiments. However, the sign of the pitch shift was curiously correlated with the sign of the mistuning of the mistuned harmonics. For example, if the fourth harmonic of a 200-Hz fundamental was mistuned by +4% so that its frequency became 832 Hz, then a typical match was 849 Hz, a pitch shift of +2%. If the fourth harmonic was mistuned by −4% so that its frequency became 768 Hz, then a typical match was 760 Hz, a pitch shift of −1%. (By definition, pitch shifts are always computed with respect to the actual frequency of the mistuned harmonic.)

The pitch shifts observed in the MHM experiment could be fairly compared with shifts predicted by Terhardt's algorithm. Because of its place-theory origins, the algorithm is not restricted to periodic complex tones; it can be used to compute pitches for inharmonic components as naturally as for harmonic components. For example, Terhardt *et al.* (1982a) used the algorithm to predict the pitch of church-bell tones. The algorithm predicts that for positive mistuning the pitch shift should be positive. As it turns out, both the sign and the magnitude of the predicted shift are in reasonable agreement with the pitch shifts observed in the MHM experiment for positive mistuning.

For negative mistuning the algorithm predicts that the pitch shift is again positive, even more positive than for an equivalent positive mistuning. This prediction (for harmonics 2–7) follows from the place theory. The fourth harmonic will serve as an example: The pitch of the fourth harmonic is shifted upward because of partial masking from below, mainly by the third harmonic. If now the fourth harmonic is mistuned in a negative direction it moves closer to the third harmonic. Therefore, the amount of masking is greater, and this leads to a greater positive pitch shift. However, the prediction for negative frequency shifts is completely contrary to the results observed in the MHM experiment. The experiment found that for negative mistunings the preponderance of the matches showed *negative* pitch shift.

The discrepancy between the predicted and observed pitch shifts is striking. It is more serious than a simple numerical matter. Rather, the discrepancy in sign for negative mistunings challenges any model of pitch perception based upon place processes incorporating mutual masking because shifts due to masking do not change sign when the mistuning changes sign.

To check the results obtained as a by-product in the MHM experiment, we decided to perform experiments that were designed specifically to measure pitch shifts. These experiments are reported below. Like the original MHM experiment these experiments used a pitch matching procedure. However, the present experiments included symmetries in method and analysis that minimize biases. Specifically we wanted to make a careful test of the MHM observation that pitch shifts are positive for positive mistunings and negative for negative mistunings.

I. EXPERIMENT 1

Experiment 1 was a pitch matching experiment in which listeners adjusted the frequency of a sine tone to match the pitch of a mistuned harmonic in a complex tone.

A. Procedure

The listener was seated in a sound-treated enclosure, holding a response box that controlled the events of an experimental trial. When the listener pressed a yellow button there was a pause of 300 ms, and then a complex tone, with one of its harmonics mistuned. When he pressed an orange button there was a pause of 300 ms, and then a sine tone, with a frequency that could be adjusted by means of a ten-turn potentiometer on the box. The pause prevented interaction between the complex tone and the matching sine tone (Rakowski and Hirsh, 1980); the potentiometer allowed the listener to make the pitch match. The listener could call up the complex tone or the matching tone, in any sequence, as often has he liked. When the listener was satisfied with his match he pressed a green button to finish the trial. The stimulus and matching frequencies were then recorded, and then the next trial, with a different amount of mistuning, began. After 11 trials a run was complete, and the listener could come out of the enclosure to rest.

During the course of an experimental run the number of the particular harmonic that was mistuned was fixed for all trials. In different runs, harmonics 2, 3, 4, 5, 7, and 9 were mistuned. There were eleven different mistunings, −8%, −4%, −2%, −1%, −0.5%, 0%, 0.5%, 1%, 2%, 4%, and 8%, each used once in a run for a total of eleven matches. On the first trial of a run the mistuning was always 8%. A harmonic mistuned by that amount was easy to hear so that the listener was immediately cued to the correct frequency region for the run. On subsequent trials the amount of mistuning was selected randomly from the other ten percentages.

B. Stimuli

The fundamental frequencies of the complex tones were in the vicinity of 200 Hz. For a given trial, a complex tone with a mistuned harmonic was loaded into a digital buffer 16 382 samples long. The buffer was recycled endlessly, and samples were converted by a 16-bit DAC at a nominal sample rate of 16 382 samples/second. To prevent the listener from using his memory for pitch to do the task, the sample rate was actually different on every trial. It was randomized over a range of +10% to −10%, with a rectangular distribution. Therefore, the fundamental frequency ranged from 180 to 220 Hz.

The analog signal was low-pass filtered at 7 kHz, −115 dB/oct. The complex tone, as presented to the listener, was shaped by a computer-controlled amplifier to give it an envelope with a 10-ms raised-cosine onset and offset and a full-on duration of 400 ms. The electrical signal had 16 components, all of the same amplitude. It was presented diotically via Yamaha YH1000 headphones at a level of 40 dB SPL, nominally 28 dB per component. In this experiment no runs were done with a mistuned fundamental because the low level made it difficult to hear the fundamental component.

The matching sine tone was generated by a voltage-controlled function generator, Wavetek VCG116, controlled by the potentiometer on the response box. Its frequency was read to an accuracy of 0.01 Hz by a triggered clock on the computer bus. The level of the matching tone was the same for all experimental conditions, fixed at a value 18 dB less than the level of the complex tone, i.e., 6 dB less than the level of each component of the complex tone. As a result, the matching tone was approximately equally as loud as an individual component. It was expected that equal loudness would make the pitch matching task easier.

C. Listeners

The listeners, E, J, K, and S, were four male undergraduates, between the ages of 19 and 24. They were selected from a larger pool of listeners based upon their accurate performance in a pitch-matching test for sine tones in the range 150 to 1000 Hz. All the listeners had negative otological histories and some training as performers of musical instruments.

D. Protocol

In the data-collection phase we obtained at least ten matches (ten runs) from each listener for each percentage of mistuning for each mistuned harmonic. Runs were blocked by mistuned harmonic number, a listener completing one block before starting on the next. The order of blocks for each listener was somewhat haphazard, but as a rule, easy harmonics (mistuned 2 and 3) were done first and harder harmonics (mistuned 1 and 9) were done later. Listeners differed greatly on the amount of time spent on a run. Listener J could complete a run in as little as 2 min. At the other extreme listener E sometimes needed as long as 20 min for a run. Listener J could easily complete a block of ten runs within a single 2-h session; listener E required several 2-h sessions for a block.

E. Results

The goal of the experiment was to find accurate estimates of the pitch of mistuned harmonics by averaging the matches made by a listener. Averages, however, can misrepresent the majority of the data if there are a few outlying points at extreme values. Our experiment was prone to such outlying data points because listeners occasionally matched the wrong harmonic. Therefore, we subjected the data to a self-consistency test. All the data from the runs for a particular listener, a particular mistuned harmonic, and a particular percentage mistuning were averaged to find a mean pitch shift, measured in percent. The standard deviation (s.d.) (N−1 weight) was computed as well. If the s.d. was less than 2.5% then the mean was accepted. If the s.d. was greater than 2.5%, then the data point that differed most from the mean was discarded and a new mean and s.d. were computed. This process was iterated until the s.d. became less than 2.5%. Unlike the data windowing used in the original MHM experiment there was no absolute requirement on the data. The

TABLE I. For four listeners, E, J, K, and S, and for different mistuned harmonics the denominator shows the total number of matches, summed over the different amounts of mistuning, and the numerator shows the number of matches accepted by the analysis. The difference between denominator and numerator is the number of matches that were rejected to make the standard deviation for each amount of mistuning less than 2.5%. Experiment 1, 28 dB per component.

Mistuned harmonic	E	J	K	S
2	110/110	121/121	111/111	121/121
3	100/100	100/100	100/100	100/100
4	100/100	100/100	100/100	110/110
5	111/111	109/110	109/110	108/110
7	112/121	113/121	114/121	111/121
9	48/48	59/60	48/48	60/60

data selection was based entirely on self-consistency. Table I shows the number of matches and the number of data points included in the average, summed over all values of the amount of mistuning.

For each listener and each mistuned harmonic we made plots showing the pitch shift, in percent, as a function of the frequency shift of the mistuned harmonic, also in percent. This led to a total of 28 figures, not shown here. Instead Figs. 1–5 show collective plots which are averages over listeners, obtained by pooling all the matches for a given mistuned harmonic and a given mistuning. Comments below indicate the individual differences. Solid lines on the figures show predictions from the algorithm of Terhardt *et al.* (1982b).[1]

Mistuned 2nd harmonic (Fig. 1): Listeners E and S showed small shifts, but shifts were as large as 2% for listener K. The tendency for pitch shifts to saturate for large positive and negative mistunings was observed for all four listeners.

Mistuned 3rd harmonic (Fig. 2): Listener S showed small shifts, the others were like the average. Saturation for large mistuning occurred for all listeners. For all listeners negative mistunings produced entirely negative shifts.

Mistuned 4th harmonic (Fig. 3): As for the mistuned 3rd

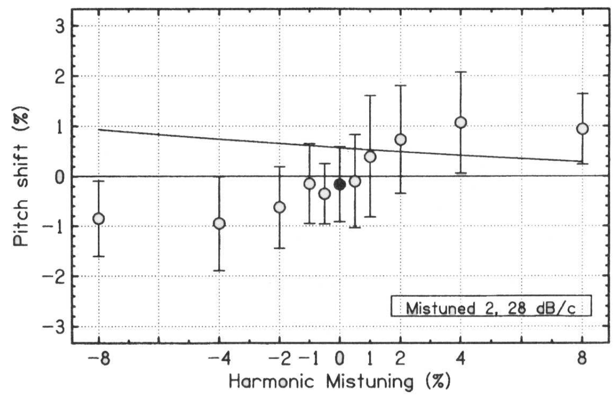

FIG. 1. Collective graph showing average pitch shifts for four listeners for the mistuned second harmonic, presented at a level of 28 dB per component. The filled symbol in the middle is for zero mistuning. Error bars are two standard deviations in overall length. The solid line shows the prediction of the algorithm of Terhardt *et al.* (1982b).

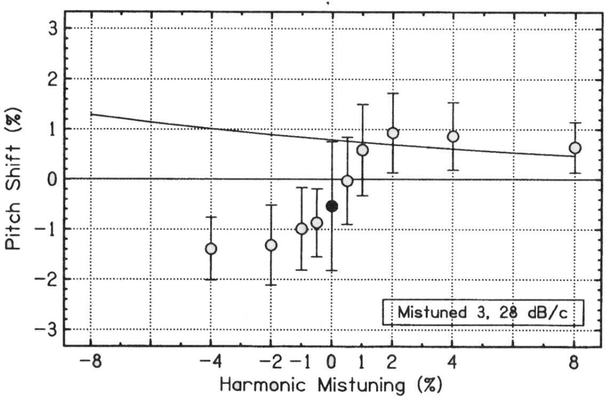

FIG. 2. As in Fig. 1, for the mistuned 3rd harmonic.

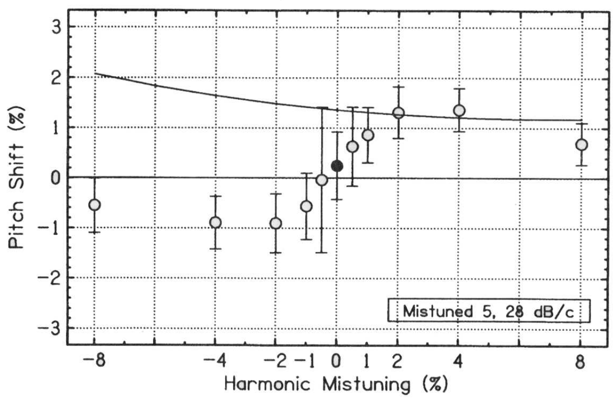

FIG. 4. As in Fig. 1, for the mistuned 5th harmonic.

harmonic, no data were collected for a mistuning of −8%. Shifts tended to be small for listeners E, J, and S. Except for E, negative mistunings led to negative pitch shifts and positive mistunings led to positive pitch shifts. Saturation occurred for all listeners.

Mistuned 5th harmonic (Fig. 4): Data from individual listeners are very similar. For all of them, all positive mistunings led to positive pitch shifts and all negative mistunings led to negative pitch shifts. The tendency for the shift to be smaller for ±8% mistuning than for ±4% was common to all listeners.

Mistuned 7th harmonic (Fig. 5): Data for listeners K and J resemble those for the mistuned 5th. Listeners E and S contributed most of the variability. For J, K, and S negative mistunings led to negative pitch shifts. Saturation effects were strong.

Mistuned 9th (Fig. 6) and 11th harmonics: Only −8%, −4%, 4%, and 8% mistunings were used, each three times on each of four runs. Data for individual listeners are given in Fig. 6 for the mistuned 9th. Data for both mistuned 9th and 11th show positive pitch shifts for positive mistunings and negative pitch shifts for negative mistunings. However, for these high harmonics we are not sure that the listeners matched the mistuned harmonic. Possibly they matched the harmonic immediately above or below the mistuned harmonic. On a ratio scale these higher harmonics are close together. In the extreme case, an 11th harmonic shifted by 8% lies only 1% below the 12th harmonic. The frequencies of the nearest-neighboring harmonics are given by asterisks in Fig. 6. It is likely that listeners did correctly match the mistuned 9th harmonic, at least for ±4%. Data are not shown for the mistuned 11th.

To summarize the data of experiment 1 (28 dB per component) negative mistunings tended to lead to negative pitch shifts, as found in the MHM experiment and contrary to the place theory algorithm. Quite generally the pitch shifts saturated for large mistunings. This was true for all mistuned harmonics and for all the listeners, with only one exception in 28 plots.

II. EXPERIMENT 2—INCREASED LEVEL

A. Method

Experiment 2 was identical to experiment 1 except that the levels of the tones were increased by 30 dB. Each harmonic of the complex tone had a level of 58 dB SPL.[2] These conditions are similar to those used by Moore *et al.* (1984) in a difference limen experiment: 12 harmonics at 60 dB per component and envelope durations identical to ours. Experiment 2 also included the mistuned fundamental. The listeners were the same as in experiment 1. Most of the runs of ex-

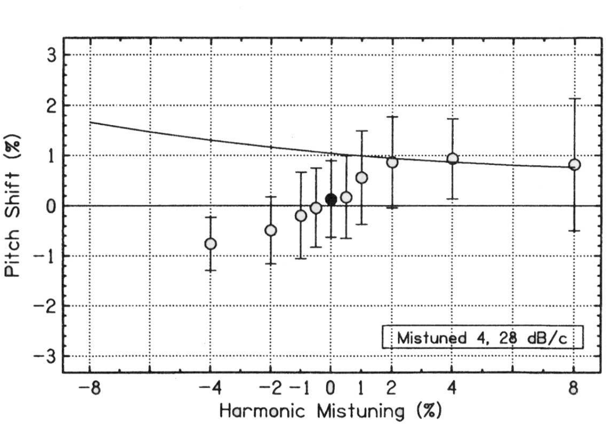

FIG. 3. As in Fig. 1, for the mistuned 4th harmonic.

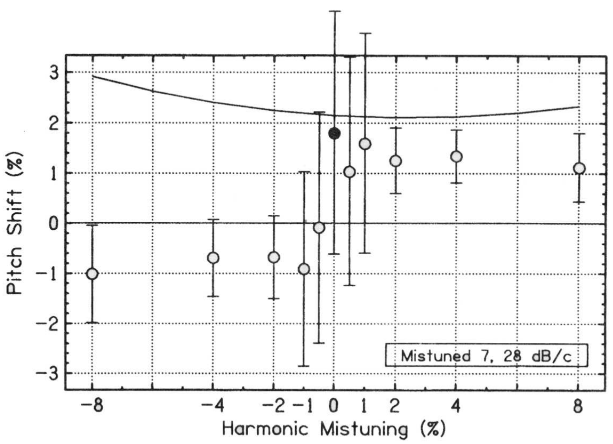

FIG. 5. As in Fig. 1, for the mistuned 7th harmonic.

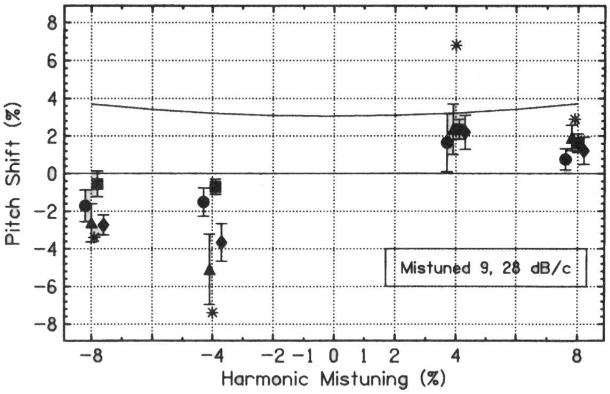

FIG. 6. Pitch shift for the mistuned 9th harmonic versus the amount of mistuning, with signals presented at 28 dB per component. Listeners are E (circles), J (triangles), K (squares), and S (diamonds). Error bars are two standard deviations in overall length. Asterisks show the locations of the frequencies of neighboring harmonics, the 8th (on the left) and the 10th (on the right).

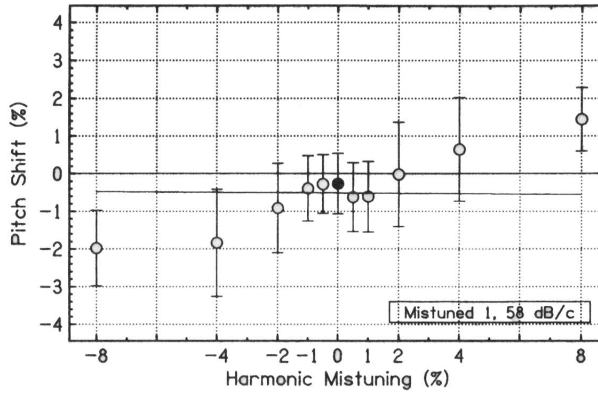

FIG. 7. Collective graph showing average pitch shifts for four listeners for the mistuned fundamental in a signal presented at a level of 58 dB per component. The filled symbol in the middle is for zero mistuning. Error bars are two standard deviations in overall length. The solid line shows the prediction of the algorithm of Terhardt *et al.* (1982b).

periment 2 were done before the runs of experiment 1; otherwise runs from the two experiments were interleaved in a haphazard way.

B. Results

The total number of matches and the number of matches that passed the 2.5% criterion for the s.d. are shown in Table II. This is directly comparable to Table I. A comparison of these tables shows that more matches for high mistuned harmonic numbers had to be rejected in the experiment done at the higher level. The pitch shifts themselves are here compared with those found in experiment 1.

Mistuned fundamental (Fig. 7): Data for individual listeners fell within the error bars of the collective graph. The largest s.d. on the collective graph is 1.4%, considerably smaller than the 2.5% criterion. Negative pitch shifts predominated; on individual listener plots they outnumbered positive pitch shifts by a 3 to 1 ratio. Negative pitch shifts are predicted by Terhardt's algorithm.

Mistuned 2nd and 3rd harmonics: The data are essentially the same as the low-level data shown in Figs. 1 and 2. Comparing individual graphs for 58 and 28 dB per component shows that the corresponding error bars overlap with one exception out of 84 data points. To conserve space, plots are not shown for harmonics 2 and above. Interested readers may request a set of figures for experiment 2 from the first author.

TABLE II. Same as Table I but for experiment 2, 58 dB per component.

Mistuned harmonic	E	J	K	S
1	110/110	110/110	132/132	121/121
2	109/110	121/121	132/132	121/121
3	100/100	100/100	100/100	100/100
4	100/100	97/100	94/100	109/110
5	100/100	104/110	98/100	99/100
7	73/110	93/100	89/100	79/100

Mistuned 4th harmonic: Pitch shifts are about 50% larger at the higher level compared to experiment 1; they are as large as 4% for listener K. Large negative mistunings always led to negative pitch shifts for all listeners. A tendency for shifts to saturate was common to all listeners.

Mistuned 5th harmonic: As for mistuned harmonics 1–4, no matches were rejected from the collective graph that were not rejected from individual graphs. However, the largest s.d. was 2.4%, close to the 2.5% limit; it occurred for 0% mistuning. A tendency for the shift to have smaller magnitude for 8% mistuning than for 4% (defined as "supersaturation") was common to all listeners. In the central cluster positive shifts were more common than negative, and shifts for listener K were positive for all mistunings. For the other listeners large negative mistunings led to negative pitch shifts. Compared to experiment 1, pitch shifts were larger at the higher level, particularly for positive mistunings. For individual listeners they were often a factor of 2 larger. Error bars were also larger at the higher level and the shift data were less orderly.

Mistuned 7th harmonic: Error bars were dramatically larger at the higher level and the shift data were less orderly. This result came as no surprise to the listeners. Informally, they reported greater confidence in their matches at the lower level. Tables I and II show the same effect; only 7% of the matches were rejected by the s.d. criterion at the lower level, compared to 19% at the higher level. Two listeners were unable to make reasonable matches for zero mistuning. Individual data plots show that positive shifts considerably outnumbered negative shifts, though the average pitch shift was negative for negative mistunings.

Mistuned 9th and 11th harmonics: Only −8%, −4%, 4%, and 8% mistunings were used. The data showed positive pitch shifts for positive mistunings and negative pitch shifts for negative mistunings, in overall agreement with data for other mistuned harmonics. However, the scatter in the data make us less sure that the listeners matched the mistuned harmonic.

To summarize experiment 2 (58 dB per component) level effects were found to be negligible for low-numbered

TABLE III. Quadrant count. For each mistuned harmonic and for two levels, the entries indicate significant pitch shifts, summed over listeners. With mistuning on the horizontal axis and pitch shift on the vertical, the entry in the first quadrant gives the number of significant positive pitch shifts given a positive mistuning, expressed as a percentage of the total number of matches to tones with positive mistunings. The entry in the second quadrant gives the percentage of pitch shifts that are significantly positive given a negative mistuning. The third quadrant gives the percentage of pitch shifts that are significantly negative given a negative mistuning, and the fourth quadrant gives the percentage of pitch shifts that are significantly negative given a positive mistuning. Numbers in a column do not sum to 100% because not all matches resulted in significant shifts.

Mistuned harmonic	58 dB/component		28 dB/component	
1	0	35		
	55	20		
2	15	40	0	45
	35	0	45	10
3	0	45	0	55
	44	0	58	0
4	0	45	0	47
	25	0	44	5
5	0	50	0	85
	38	0	60	0
7	25	94	10	75
	0	0	42	5
9			0	100
			100	0

harmonics 2 and 3. For higher harmonics pitch shifts increased in magnitude with increasing level. Variability also increased with increasing level especially for mistuned harmonics 5 and above. In many cases pitch shifts were somewhat more positive at the higher level. Pitch shift saturation for large mistuning was present at both levels but was less marked for the higher level.

III. DISCUSSION

A. Quadrant count

The quadrant count is a way to display the data that effectively checks the most elementary hypothesis under test, namely, that a positive mistuning produces a positive pitch shift and a negative mistuning produces a negative pitch shift. The quadrant count is based upon graphs for individual listeners, with mistuning on the horizontal axis and pitch shift on the vertical. We define a pitch shift to be "significant" if the mean shift is positive or negative and the error bar ($N-1$ weight) does not cross the horizontal axis. (This criterion is moderately stringent. If the matches are normally distributed then 84% of them must lie on one side or the other of the horizontal axis in order to be counted as significant.) The quadrant count is simply the number of significant pitch shifts in each quadrant. For this purpose data obtained with zero percent mistuning are ignored. If the hypothesis is absolutely true then counts for the second and fourth quadrants, where mistunings and pitch shifts have opposite signs, will be exactly zero. The counts, summed over listeners and expressed as a percentage of the total, are shown in Table III, one column for each level.

Table III shows that the hypothesis is effectively verified. Entries in the first and third quadrants greatly exceed entries in the second and fourth. There is one exception, the mistuned 7th at the 58-dB level, where a strong tendency toward positive shifts put all the significant shifts in the second quadrant for negative mistunings.

B. Variability

The variability observed in our matching data, both for a single listener and across listeners, is considerably larger than the variability expected for matching one sine tone to another. As shown by Moore *et al.* (1984) one effect of embedding a sine tone in a context of harmonic components is an increase in the frequency difference limen. Error bars on individual plots for our best listeners are smaller than the difference limens of Moore *et al.* for small harmonic numbers and comparable at larger harmonic numbers.

As indicated above, for a level of 58 dB per component there was a sudden degradation in the reproducibility of the data as the mistuned harmonic number increased from 5 to 7. This result is in agreement with the difference limen data from Moore *et al.* It is presumably due to the difficulty of hearing out harmonics that are closely spaced, as studied by Plomp (1964) and Plomp and Mimpen (1968). In fact, the transition takes place exactly where Plomp (1964) says it should; for a 200-Hz fundamental the 5th harmonic should be separately audible, the 7th not. This result is normally associated with resolving power along the tonotopic coordinate of the auditory system, as expressed by critical bandwidths. According to Zwicker (1961), near our 5th harmonic (1000 Hz) the critical band width is less than 200 Hz; near our 7th harmonic (1400 Hz) the critical band width is greater than 200 Hz. The correspondence between the theory of auditory resolution and the variability of our data at 58 dB per component is therefore good. If one accepts that correspondence, however, then the considerable improvement in the ability of listeners to resolve the 7th harmonic at 28 dB per component suggests that something about the critical band, its width or its shape, changes for a 30 dB reduction in level.

C. The strategic advantage of mistuned harmonics

The special nature of our experiment is that listeners matched harmonics that were mistuned. Plots of pitch shift versus mistuning were found to be reproducible, rather stable, and orderly. The data are of high enough quality that one is motivated to try to find a model to explain them.

It is possible, of course, to extract from our data only those pitch shifts observed for zero mistuning. The result should be comparable to previous results obtained by Terhardt (1971) or by Peters *et al.* (1983). Plots of these pitch shifts are shown in Figs. 8 and 9 as a function of harmonic number. The figures do not show an orderly pitch shift. Errors and individual differences are large enough that the average is consistent with the idea that there is no pitch

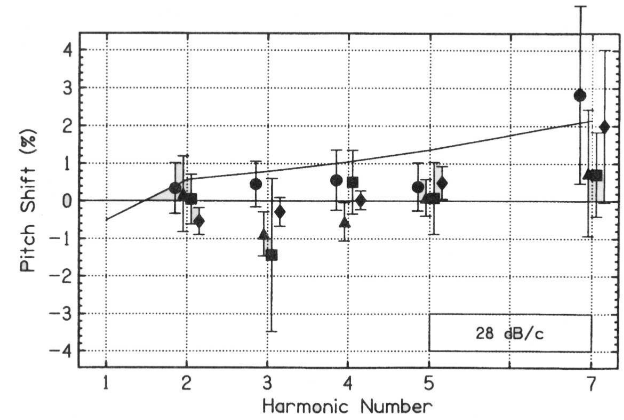

FIG. 8. Experimental pitch shift for five different harmonics that are not mistuned. The signal level is 28 dB/component. Symbols for four listeners are as in Fig. 6. Error bars are two standard deviations in overall length. The solid line shows the prediction of the algorithm of Terhardt.

shift at all, with the possible exception of the mistuned 7th harmonic at 28 dB per component. If we only had the data for perfectly harmonic components we would agree with Peters *et al.* that the spectral pitches of the harmonics of a complex tone are are not shifted. The complete set of data, however, makes it evident that reproducible pitch shifts do exist and that they can be large.

IV. PREDICTIONS OF A NEURAL TIMING MODEL

The data shown above show that a masked-excitation place model is inadequate to represent the pitch of the components of a complex tone. Although the model, as exemplified by Terhardt's algorithm, agrees with the data for positive mistunings, everything goes wrong when mistunings are made negative.

A. Introduction to the timing model

As alternatives to place models for pitch, there are models based upon neural timing or counting. Among a wide variety of possible candidates is a model that begins with the interspike interval (ISI) histogram and supposes that pitch

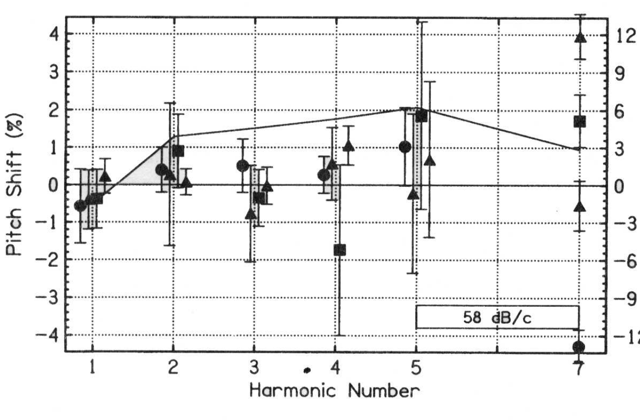

FIG. 9. As in Fig. 8, for 58 dB/component. For the seventh harmonic (only) the vertical scale is multiplied by a factor of 3, as shown by the labels on the right-hand axis.

perception is derived from a function rather like that. This approach has the advantage that the ISI histogram is flexible (Ohgushi, 1983). Unlike other timing processes where the spike rate is rigidly tied to the periodicity of the stimulus, the ISI histogram does not necessarily have its peaks at exact multiples of the period. In fact, we are aware of only two restrictions on the ISI histogram; both of them are sum rules: First, the integral of the ISI histogram must equal the total number of spikes in the record. Second, the first moment of the ISI histogram must equal the total duration of the record (Hartmann, 1993). These constraints permit considerable variation. For example, the first peak in the histogram may occur at intervals longer than the period, if there are compensating changes elsewhere in the function. As a result, pitch models based upon the ISI histogram can accommodate pitch shift effects, diplacusis for example.

The ISI histogram has been successfully employed in several pitch studies. Ohgushi (1983) argued that the pitch shift phenomenon known as octave stretch can be explained on the basis of delays seen in experimental ISI histograms as the frequency is doubled. Jones *et al.* (1983) and Tubis *et al.* (1986) used model ISI histograms incorporating refractory effects to derive Stevens' rule for the average pitch shifts as signal intensity is changed. Apart from these pitch shift studies, we note that Goldstein and Srulovicz (1977) obtained highly satisfactory agreement between frequency difference limen data and the predictions of a model in which neural encoding is represented by the ISI histogram.

The model explored in this section extends the application of the ISI histogram to the case of complex tones. The way that the model works is not hard to describe: To calculate the pitch of a component of a complex tone we consider an auditory filter with its characteristic frequency f_c in the neighborhood of the component of interest. The filter passes mainly, but not exclusively, the component itself. Components of higher and lower frequencies will be passed, more or less depending upon the slopes of the filter passband. From a model of neural encoding we calculate the ISI histogram for the spike train in that auditory filter. The intervals corresponding to the peaks of the ISI histogram give estimates of the pitch for the component. To predict the results of a pitch matching experiment we next do an equivalent calculation for the pitch of the matching sine tone. The difference between the component pitch and the sine tone pitch gives an estimate of the pitch shift.

It is also not hard to see why we expect this model to work in a way that can account for our data. When the frequency of a component of a complex tone is mistuned positively the locus of maximum excitation moves on the tonotopic axis. Therefore the characteristic frequency of the appropriate filter is similarly shifted. This filter passes a larger fraction of those components having frequencies higher than the component of interest. The peaks of the resulting ISI histogram are therefore shifted to smaller intervals, corresponding to higher pitches, and this leads to a pitch shift in a positive direction. An analogous argument should predict negative pitch shifts for negative mistunings in agreement with experiment.

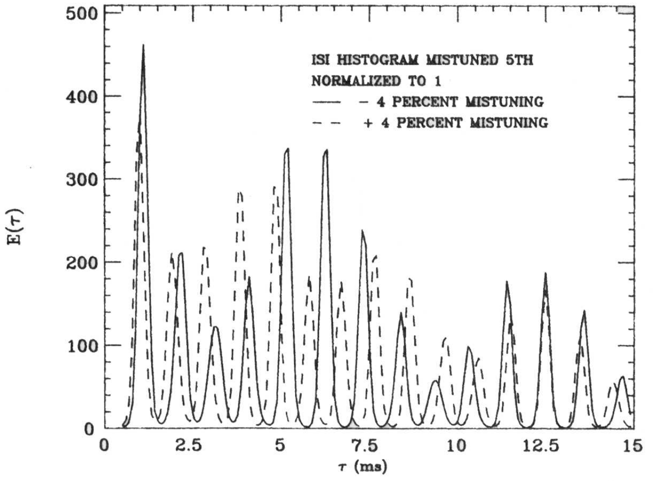

FIG. 10. Interspike interval histograms computed from the model of the Appendix for a mistuned 5th harmonic. The solid line shows -4% mistuning, the dashed line shows $+4\%$ mistuning. Histograms are shown as line plots to make them easier to follow by eye. The bin width was 0.1 ms.

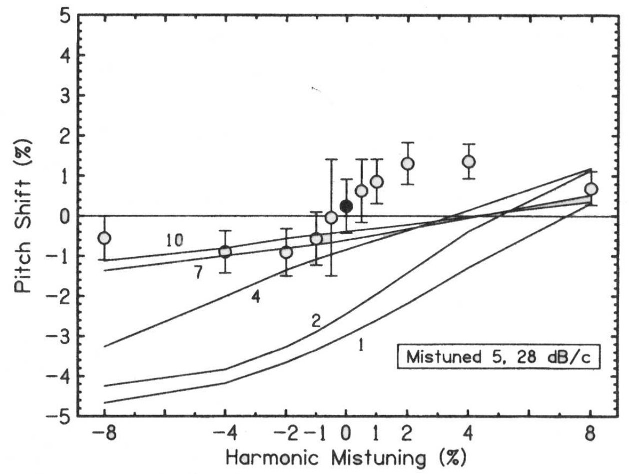

FIG. 11. Experimental pitch shift data from Fig. 4 for the mistuned 5th harmonic compared with calculations from the timing model wherein the filter center frequency tracks the mistuned harmonic. Model predictions are shown as lines, the lowest for pitch based upon the first peak in the ISI histogram, others based upon information accumulated over 2,4,...,10 peaks, showing convergence.

B. Timing model calculations

The calculation of the ISI histogram for complex tones is intricate and is described in the Appendix. An important assumption in the calculation is that the ensemble of neurons that contribute ISI histograms to the pitch perception process can be represented by a single representative neuron with a characteristic frequency equal to the frequency of the mistuned harmonic. This assumption seems reasonable because it is expected that the process of mistuned harmonics segregation focuses attention on the mistuned harmonic.

Two model ISI histograms are shown in Fig. 10. Both are for mistuned 5th harmonics, one with $+4\%$ mistuning, the other with -4% mistuning. The calculated histograms assume that all harmonics have equal amplitudes and that these are large enough that neural firing rate is saturated. The average firing rate is taken to be $\bar{R}=100$ spikes/s, but our calculations are quite insensitive to the value of $\bar{R}$.

The reciprocal of the ISI for the first peak provides an estimate of the pitch of the mistuned harmonic. Twice the reciprocal of the ISI for the second peak provides a second estimate, etc. The final estimate of the pitch is a weighted average of the estimates derived from individual peaks, where the weighting functions are given by the square roots of the heights of the peaks. Figure 11 shows the results of this calculation in successive stages for the mistuned 5th harmonic. The lines on the figure show calculated values of the pitch shift based upon a total of 1, 2, 4, 7, or 10 peaks. As more peaks are included in the average the predicted shifts converge.

Figure 11 also shows the measured pitch shifts for the four listeners in the low-level experiment. It is evident that the pitch shifts calculated by the model for seven peaks or more have about the right overall span. It is also evident that the calculated shifts are too negative and have the wrong shape. This discrepancy for the mistuned 5th harmonic is part of a trend. A comparison between measured and calculated shifts for mistuned 7th and 9th harmonics reveals a discrepancy that is even larger. Comparison for 4th and lower harmonics shows a decreasing discrepancy, such that agreement between measured and calculated shifts is good for the mistuned 3rd.

C. Combined place and timing model

Our hybrid place and timing model is essentially a timing model in that it determines pitch from the reciprocals of the times of peaks in an ISI histogram, just as in the timing model above. Place principles enter the model in the following way: There are many different neurons that synchronize with the mistuned harmonic. Some of them have characteristic frequencies somewhat above the mistuned harmonic frequency, and their ISI histograms are affected by higher harmonics. Neurons tuned on the low-frequency side of the mistuned harmonic are affected oppositely. A proper pitch calculation should take account of all of these neurons, weighting their inputs to the central pitch processor by their firing rates or their synchrony indices. The weights are determined by the spectrum of the tone and by partial masking, especially partial masking from below. Thus the interaction of excitation patterns along the tonotopic coordinate can shift pitches because it changes the weights attached to different ISI histograms.

Our actual calculations, however, did not combine the outputs of many different neurons. Because of computational constraints we continued to use the output of a single representative neuron. What was changed was that the characteristic frequency of the representative neuron was no longer taken to be equal to the mistuned harmonic frequency. Instead it was higher because of the upward spread of masking.

Combining timing and place principles in this way was a sensible approach in view of our experimental data because the partial masking of one harmonic by another is largest for high harmonics, where the timing model by itself does not agree with experiment. The partial masking becomes small

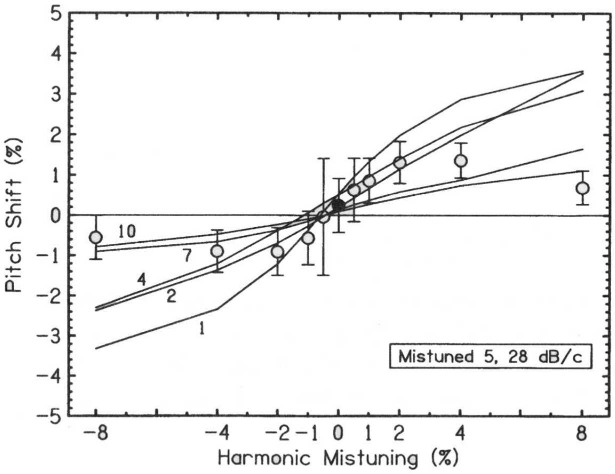

FIG. 12. Experimental pitch shift data for the mistuned 5th harmonic compared with a timing model wherein the characteristic frequency of the auditory filter is shifted by partial masking of the excitation of the mistuned harmonic.

for low harmonics where no correction is needed. The sign of the masking contribution to the shift is also right. Because of the upward spread of excitation, there is a positive displacement of the characteristic frequency. This leads to an increased contribution of higher harmonics to the ISI histogram which increases the predicted pitch, as needed. The dependence on signal level is also in the right direction. Increased upward spread of masking with increasing level results in larger positive pitch shifts, in agreement with the data for the higher harmonics 4, 5, and 7.

For a specific implementation, the displacement in characteristic frequency was taken to be proportional to the reciprocal of the percentage by which the frequency of the mistuned harmonic exceeded the frequency of the harmonic immediately below. This choice led to larger displacements for larger harmonic numbers where harmonics are more closely spaced on a ratio scale. For the mistuned 5th harmonic, the displacement in characteristic frequency varied monotonically from 10%, to 4% as the mistuning varied from −8% to +8%. Figure 12 shows the resulting predictions for the pitch shift. As the number of peaks of the ISI histogram increases, the calculation agrees better with the experiment. The tendency for predicted pitches to be too low has been cured. However, the agreement with experiment is not perfect; the tendency for the shifts to saturate with increasing mistuning is not modeled accurately in the calculation. This matter is treated further in the concluding section.

V. CONCLUSION

The experiments described above measured the pitches of the harmonics of a complex tone having 16 equal-amplitude harmonics. Harmonic levels were either 28 or 58 dB SPL. Harmonics numbered 1, 2, 3, 4, 5, 7, 9, and 11 were measured. When the harmonics were not mistuned from their correct integer values, their pitches were found to be equal to their frequencies, consistent with the zero shift results of Peters *et al.* (1983).

When harmonics were mistuned positively, positive pitch shifts appeared such that the pitch of a mistuned harmonic was higher than its frequency. This result is consistent with the excitation pattern algorithm of Terhardt. But when the harmonics were mistuned negatively, the pitch shifts became negative. By contrast, the excitation pattern algorithm predicts that the pitch shift should become even more positive. This disagreement does not seem to be reparable by any reasonable modification of the excitation pattern model.

An alternative model based upon neural timing, represented by the ISI histogram, is also in disagreement with the data, but in a way that is more easily corrected. A hybrid model combining timing and excitation pattern models is suggested in which pitch is determined by the ISI histogram, but the ensemble of neurons that contribute ISI histograms to the pitch perception process consists of the most strongly excited neurons. The relative excitation strengths, in turn, are determined by masking interaction among excitation patterns.

The hybrid model can predict pitch shifts that agree approximately with measured data for mistuned harmonics 2 and greater. Some details, however, are missing. The saturation of the pitch shift for large mistunings is a highly reproducible result that should be explained by a model, but the model calculations shown in Fig. 12 show little saturation. In fact, we have done numerous calculations in which saturation and supersaturation does occur, for both positive and negative extreme mistunings. These calculations tend to be those in which the peripheral filters are wider than those recommended by Colburn (1973). The problem with filters that are wide enough to predict saturation is that they introduce a phase dependence into the predicted pitch shifts, a result for which we have no experimental support.

The most serious flaw of the hybrid model is that it fails to predict the correct pitch shift for a mistuned fundamental (harmonic number 1). The model prediction, unique for the fundamental, is that the pitch shift should be very small and opposite in sign to the mistuning. By contrast, the experimental pitch shift is the same as for any other harmonic, namely substantial and of the same sign as the mistuning. The problem is endemic to the model. Essentially, it occurs because every peak in the ISI histogram of a neuron synchronized to the fundamental coincides with some peak in the ISI histogram for any other harmonic. Based upon many calculations we think it likely that, with the right choice of parameters, the model could account for the data for all other mistuned harmonics, but not for the mistuned fundamental.

ACKNOWLEDGMENTS

We are grateful to Dr. Arnold Tubis and to Dr. Julius Goldstein for helpful conversations on model ISI histograms. Dr. Ernst Terhardt kindly reviewed our computer program for the pitch shifts predicted by his algorithm. Jian-Yu Lin made a particular study of the model predictions for the mistuned fundamental. This work was supported by a grant from the National Institutes of Health, NIDCD.

APPENDIX: MODEL ISI HISTOGRAM

Most of the work on model ISI histograms has been specific to sine-tone stimuli. The so-called "MIT model" based on auditory nerve data (Goldstein, 1973; Colburn, 1973) is the most visible example. To construct a model ISI histogram for a complex tone, as needed in our calculations, is rather more complicated. Details of auditory filtering become important, and nonlinear effects, including suppression, must be accommodated in some way by the model. The present development of an ISI histogram for a complex tone represents a minimal modification of the model for sine tones. Essentially it is the model proposed by Srulovicz and Goldstein (1983).

The development begins with the model response of an auditory-nerve fiber to a sine tone with frequency f_i. The signal is given by

$$s(t)=A_i\cos(2\pi f_i t+\phi_i). \tag{A1}$$

The response of an auditory-nerve fiber to the tone is described by a nonhomogeneous Poisson process with rate function λ. The model therefore neglects the refractory nature of neurons. The rate function for the Poisson process is given by the output of an exponential rectifier, which appears to be a reasonable approximation for signals with line spectra (Evans, 1968; Siebert, 1968). Therefore, the rate function is

$$\lambda(t)=\frac{\bar{R}}{B}\exp[Z_i\cos(2\pi f_i t+\phi_i)], \tag{A2}$$

where $\bar{R}$ is the average firing rate, and Z_i is the synchrony parameter. Constant B normalizes the exponential function. Here, as below, we pay little respect to normalization because normalization is not relevant to our calculation.

To estimate the value of Z_i, we expand the exponential function in a Fourier series,

$$\exp[Z_i\cos(x)]=I_0(Z_i)+2\sum_{k=1}^{\infty}I_k(Z_i)\cos(kx). \tag{A3}$$

The coefficients I_k are the modified Bessel functions of order k. A measure of synchrony is the ratio $S(Z)=I_1(Z)/I_0(Z)$. This ratio has been called the "synchrony index" or "vector strength" (see references in Johnson, 1980). It is simply one half of the maximum in a cross-correlation integral between a period histogram of the measured firing rate and a sine function with the period of the signal. Function $S(Z)$ is a monotonically increasing function of Z, with asymptotic value 1. The largest measured synchrony indices are at low frequencies; for high-level signals they are about 0.92, which suggests that the low-frequency limit of the argument of S should be about $Z=6.5$.

We now generalize the above to the case of a complex tone. For a complex tone with N partials the signal is

$$s(t)=\sum_{i=1}^{N}A_i\cos(2\pi f_i t+\phi_i). \tag{A4}$$

Then the rate function becomes

$$\lambda(t)=\frac{\bar{R}}{B}\exp\left(\sum_{i=1}^{N}Z_i\cos(2\pi f_i t+\phi_i)\right). \tag{A5}$$

This is the form used by Srulovicz and Goldstein (1983), who have also suggested the following approach to describe the synchrony. Parameters Z_i are given in terms of the amplitudes of the components, A_i, the transfer function of a hypothetical linear auditory filter H_i and an expression for the maximum synchrony G_i.

$$Z_i=\frac{A_iG_iH_i}{1+(\sum_{j=1}^{N}A_j^2H_j^2)^{1/2}}. \tag{A6}$$

We now consider a specific application, namely, the pitch of a harmonic of a complex tone. It is assumed that listeners can monitor selected places along the tonotopic axis, i.e., selected auditory filters. Because the task is to match the pitch of a harmonic it is to the listener's advantage to monitor that place where there is the most information about the harmonic in question. Therefore the model confines its attention to the neural signal in an auditory filter with a characteristic frequency f_c that is near the frequency of the harmonic, specifically the mistuned harmonic in our case.

The filter function is from Colburn (1973)

$$H_i=(f_i/f_c)^{\alpha},\quad f_i\leq f_c, \tag{A7a}$$

$$H_i=(f_i/f_c)^{-2\alpha},\quad f_i>f_c, \tag{A7b}$$

where the slope is

$$\alpha=4,\quad f_c\leq 800\ \text{Hz}, \tag{A8a}$$

$$\alpha=f_c/200,\quad f_c>800\ \text{Hz}. \tag{A8b}$$

Function G_i for the maximum possible synchrony is given by Johnson's (1974) expression:

$$G_i=6.5\{[1+(f_i/630)^2][1+(f_i/3000)^2]\}^{-1/2}. \tag{A9}$$

This completes the description of the model rate function for neural firing, assumed to be Poisson. The transformation of the model derived for sine tone response into a model for complex tone response has been simple indeed. Some concession to the expected behavior of real neurons appears in Eq. (A6) for Z_i, where suppression of the response at f_c by neighboring components of the tone is included by means of the sum in the denominator. Here, neighboring components contribute to suppression according to their strength in the auditory filter at f_c. Absent from our description is any account of phase shift on the basilar membrane. It is implicitly assumed that the phase angles in the neural response are those of the signal. This assumption is dubious at low frequencies (Allen, 1983).

The basis for our model of pitch perception is the ISI histogram, E, with bin width $\Delta\tau$. This is given by

$$E(\tau)=\Delta\tau\int_0^T dt\ \lambda(t)\lambda(t+\tau)\bar{P}(t+\tau|t), \tag{A10}$$

where T is the duration of the tone, and where $\bar{P}(t+\tau|t)$ is the probability that if a neuron fires at time t it does not fire in the interval between t and $t+\tau$. For a Poisson process, $\bar{P}$ is given by

$$\bar{P}(t+\tau|t)=\exp\left(-\int_t^{t+\tau}dt'\ \lambda(t')\right). \tag{A11}$$

To make the integral in E mathematically tractable we approximate $\bar{P}$ by a time-independent average,

$$\bar{P}(t+\tau|t)=\exp(-\bar{R}\tau). \tag{A12}$$

To neglect the correlation between $\bar{P}$ and the rate functions in the integral for E seems *a priori* to be a crude approximation. A special feature of our application, however, makes the approximation better than in the general case. Below we shall be concerned with the peaks of function $E(\tau)$. These occur at values of τ that span a period (or an approximate period) of the rate function. In that case the value of $\bar{P}(t+\tau|t)$ is, in fact, independent of t, or nearly so.

A standard trigonometry identity then leads to a simplified expression for the ISI histogram:

$$E(\tau)=\left(\frac{\bar{R}}{B}\right)^2 \Delta\tau\ e^{-\bar{R}\tau}\int_0^T dt\prod_{i=1}^N \exp\{[2Z_i\cos(\pi f_i\tau)] \times\cos(2\pi f_i t+\pi f_i\tau+\phi_i)\}. \tag{A13}$$

The integrand can be rewritten in series form using the expansion of Eq. (A3):

$$\prod_{i=1}^N \left(I_0[2Z_i\cos(\pi f_i\tau)]+2\sum_{k=1}^\infty I_k[2Z_i\cos(\pi f_i\tau)] \times\cos[k(2\pi f_i t+\pi f_i\tau+\phi_i)]\right). \tag{A14}$$

The integrand can be regarded as a product of binomials with first term I_0 and second term given by the series in k. Expanding the product of binomials leads to an expansion for the ISI histogram:

$$E=E_0+E_1+E_2+\cdots . \tag{A15}$$

The zeroth-order term is

$$E_0=T\left(\frac{\bar{R}}{B}\right)^2 e^{-\bar{R}\tau}\Delta\tau\prod_{i=1}^N I_0[2Z_i\cos(\pi f_i\tau)]. \tag{A16}$$

The first-order term E_1 involves $(N-1)$ factors of I_0 and a single factor which is a sum of cosine terms. This term vanishes so long as the duration is longer than several periods because the average of a trigonometric function is zero.

The second-order term E_2 involves $(N-2)$ factors of I_0 and a product of two infinite sums on k. By the orthogonality of the cosine functions only a limited number of these terms contribute to E, but those that do contribute introduce a phase dependence into the ISI histogram. Thus the lowest order nonvanishing correction to E results in an ISI histogram that depends upon the phases of the components.

Neglecting the second-order and higher-order terms is a better approximation in some cases than in others. In the limit of infinitely sharp filters the approximation is, of course, exact. One suspects, based upon the form of the second-order correction, that the approximation is worst if all phase angles are zero. Then all terms in the k sum contribute the maximum amount and all are positive. If all phase angles are 90 deg, as in our experiments, there is cancellation among the terms and E_0 should be a better approximation to E. Further, some advantage is gained in our calculation because the harmonic of interest is actually mistuned. The largest contributions to the ISI histogram come from terms involving the mistuned harmonic because of the peripheral filtering. But the integrals involving the mistuned harmonic frequency tend to vanish because of the orthogonality of the trigonometric functions. Therefore, E_0 is a better approximation to E for the case of mistuned harmonics than for a strictly periodic tone.

We have checked the conjectures in the paragraph above by comparing analytical values of E_0 with numerical calculations of E as given in Eq. (A13) using a large computer. The numerical comparison showed that with filters as narrow as those of Eqs. (A7) and (A8) and with values of mistuning per our experiments the analytical values E_0 are rather good approximations to E. That means that the combination of filter width and mistuning makes E quite insensitive to the phases of the components. This is a satisfying result; the E_0 approximation is computationally fast and the form is simple enough that one can readily understand the effects observed in the calculations. Our numerical work also showed that when filters are broader, for example, when they are twice as broad, then E_0 is a rather poor approximation to E.

[1]The calculations were done using Eq. (10) from Terhardt *et al.* (1982b). The level of the sine matching tone was taken to be 6 dB less than the level of each component of the complex tone, as in the experiments of this paper.

[2]The experiments for mistuned 3rd and 5th harmonics were repeated with signals that had been equalized to produce harmonics of equal amplitude as measured at the headphones with a flat-plate coupler, exactly 58 dB per component. For all listeners the matches for this experiment fell within the error bars for the unequalized stimuli. We conclude that the pitch matches of this study are stable against small changes in the levels of the harmonics.

Allen, J. B. (**1983**). "Magnitude and phase-frequency response to single tones in the auditory nerve," J. Acoust. Soc. Am. **73**, 2071–2092.

Colburn, H. S. (**1973**). "Theory of binaural interaction based on auditory-nerve data I. General strategy and preliminary results on interaural discrimination," J. Acoust. Soc. Am. **54**, 1458–1470.

Evans, J. E. (**1968**). "Characterization of synchronization to low frequency tones by fibres in the cat's auditory nerve" (unpublished).

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

Goldstein, J. L., and Srulovicz, P. (**1977**). "Auditory-nerve spike intervals as an adequate basis for aural spectrum analysis," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, New York), pp. 337–345.

Hartmann, W. M. (**1988**). "Pitch perception and the segregation and integration of auditory entities," in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 623–645.

Hartmann, W. M. (**1993**). "On the origin of the enlarged melodic octave," J. Acoust. Soc. Am. **93**, 3400–3409.

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

Helmholtz, H. von (**1885**). *On the Sensations of Tone*, translated by A. J. Ellis (reprinted in 1954 by Dover, New York).

Houtsma, A. J. M. (**1981**). "Noise-induced shifts in the pitch of pure and complex tones," J. Acoust. Soc. Am. **70**, 1661–1668.

Johnson, D. H. (**1974**). "The response of single auditory-nerve fibers in the cat to single tones; synchrony and average rate," thesis, MIT.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115–1122.

Jones, K., Tubis, A., and Burns, E. M. (**1983**). "Temporal neural response correlates of pitch-intensity effects and diplacusis," J. Acoust. Soc. Am. Suppl. 1 **74**, S8.

McAdams, S. (**1984**). "Spectral fusion, spectral parsing and the formation of auditory images," Ph.D. thesis, Stanford University.

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

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

Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (**1983**). "Pitch of components of a complex tone," J. Acoust. Soc. Am. **73**, 924–929.

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

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

Rakowski, A., and Hirsh, I. J. (**1980**). "Post-stimulatory pitch shifts for pure tones," J. Acoust. Soc. Am. **68**, 467–474.

Siebert, W. M. (**1968**). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, edited by P. Kolers and M. Eden (MIT, Cambridge, MA).

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. (**1971**). "Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon," Proc. 7th ICA **3**, 621–624.

Terhardt, E. (**1972**). "Frequency and time resolution of the ear in pitch perception of complex tones," *Hearing Theory 1972*, edited by B. L. Cardozo (Institute for Perception Research, Eindhoven, The Netherlands).

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

Terhardt, E. (**1979**). "Calculating virtual pitch," Hear. Res. **1**, 155–182.

Terhardt, E., Stoll, G., and Seewann, M. (**1982a**). "Pitch of complex signals according to virtual-pitch theory: Test, examples, and predictions," J. Acoust. Soc. Am. **71**, 671–678.

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

Tubis, A., Jones, K., and Burns, E. M. (**1986**). "Pitch-intensity effects and model calculations of the level dependence of VIIIth nerve response functions," J. Acoust. Soc. Am. Suppl. 1 **79**, S80.

Zwicker, E. (**1961**). "Subdivision of the audible frequency range into critical bands (Frequenz-gruppen)," J. Acoust. Soc. Am. **33**, 248.