The pitch of a mistuned harmonic: Evidence for a template model

Jian-Yu Lin and William M. Hartmann
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

(Received 18 July 1996; accepted for publication 15 January 1998)

A harmonic of a periodic complex tone can be heard out as a separate entity if the harmonic is slightly mistuned from its correct frequency. Pitch matching experiments show that the pitch of such a mistuned harmonic differs systematically from its frequency. The shift in pitch is found to be an exaggeration of the frequency mistuning. This article considers two classes of model for the pitch shift. In the first class are tonotopically local interaction models which attribute the pitch shift to interactions between the mistuned harmonic and neighboring harmonics, where the neighborhood is established by peripheral filtering. The second class of model attributes the pitch shift to a contrast between the mistuned harmonic and a broadband harmonic template. This article describes six pitch matching experiments using complex tones having spectral gaps, strategically chosen to compare local interaction and template models. The results show that when a competition is set up between local interactions and a template, the template proves to be dominant. A parallel between the pitch shifts of mistuned harmonics and periodicity pitch, also attributed to a harmonic template, is seen as the frequency range of the mistuned harmonic is changed. Tonotopically local influences are evident in several experiments, but they are of secondary importance. [S0001-4966(98)00305-1]

PACS numbers: 43.66.Hg, 43.66.Ba [JWH]

INTRODUCTION

A complex periodic tone with many harmonics is normally perceived as a single entity with a pitch approximately equal to the fundamental frequency of the tone. The individual harmonics are not heard separately. However, if a harmonic is mistuned from its correct frequency, it may be detected as a separate entity (Moore *et al.*, 1986; Hartmann, 1988). The mistuned harmonic sounds like a sine tone against the background of the low-frequency complex tone. The background, with a complex timbre, has a pitch close to the fundamental, though this pitch can be shifted as a result of the mistuning (Moore *et al.*, 1985; Darwin *et al.*, 1994). The mistuned harmonic appears as a pure tone, and its pitch is also shifted. Matching it with a (nonsimultaneous) sine tone gives a matching frequency that systematically disagrees with the mistuned harmonic frequency (Hartmann *et al.*, 1990).

The pitch shift of the mistuned harmonic can be described succinctly as an exaggeration of the frequency mistuning. For example, the third harmonic of a 200-Hz tone is normally at 600 Hz. If the harmonic is mistuned by 8%, its frequency becomes 648 Hz. Typically its pitch would increase to 655 Hz, corresponding to a pitch shift of 1% $[(655-648)/648=0.01]$. If the harmonic is mistuned by -8%, its frequency decreases to 552 Hz, but its pitch would typically decrease to 546 Hz, corresponding to a pitch shift of -1% $[(546-552)/552=-0.01]$. The reader will notice that pitch shifts are computed with respect to the mistuned harmonic frequency. The present article is a study of the pitch shifts of mistuned harmonics.

To find a model for the pitches of mistuned harmonics, it is natural to begin with previous work, which has been on the pitches of correctly tuned harmonics. In 1979, Terhardt presented a semi-empirical algorithm to calculate the pitches of the harmonic components of complex tones. These then became the *spectral pitches* that were combined to form a virtual pitch for the tone as a whole (see also Terhardt *et al.*, 1982b). According to the algorithm, the pitches of individual harmonics are shifted due to partial masking caused by the excitation patterns of neighboring harmonics.

The algorithm is based on experimental data showing that when a sine tone is partially masked by a sound having a lower frequency than the sine itself, the pitch of the sine is shifted upward. Similarly, masking by a sound having a higher frequency than the sine tone tends to shift the sine tone pitch downward.[1] Terhardt extended these sine tone masking ideas to the pitches of components in a complex tone. A harmonic, somewhere in the middle of the spectrum, tends to be shifted upward by neighboring harmonics having lower frequencies, and it tends to be shifted downward by neighboring harmonics having higher frequencies. Because of the asymmetry of auditory filters there is an upward spread of masking, and the usual net result is that the pitch of a harmonic is shifted upward. An exception occurs for the fundamental component, where there is no lower frequency excitation. Then the model predicts that the pitch shift is downward.

Because Terhardt's model is a place-based excitation pattern model, it can also be used to predict pitch shifts for inharmonic components. For example, Terhardt *et al.* (1982a) used the model to predict the pitches of church-bell tones. However, when the model was applied to mistuned harmonics it predicted positive pitch shifts no matter whether the mistuning was positive or negative. Although it correctly predicted the pitches found experimentally for positive frequency mistuning, it failed to predict the negative shifts for negative mistuning (Hartmann *et al.*, 1990; Hartmann and Doty, 1996).

To try to explain the observed pitch shifts for mistuned harmonics, Hartmann and Doty developed an alternative model based on neural timing. The model began with the idea that pitch should be determined by the peaks of the interspike interval (ISI) histogram, as suggested by Goldstein and Srulovicz (1977). According to the model, the pitch of a component is shifted when excitation from neighboring harmonics appears in the same auditory filter, leading to a complicated temporal pattern. In general, the effect of such interaction on the ISI histogram is to cause a harmonic to be attracted in pitch to its nearest neighbors. Therefore the higher neighbor tends to make a positive pitch shift, and the lower neighbor tends to make a negative pitch shift, an ''attraction'' effect which is opposite to Terhardt's ''repulsion'' effect. Decreasing the spacing in frequency between the harmonic and a neighbor by mistuning the harmonic results in an increased attraction.

The timing model was capable of producing the positive and negative pitch shifts that occur for positive and negative mistunings. With appropriate model parameters, calculated pitch shifts could be brought into agreement with experimental results for mistuned harmonics 2, 3, 4, 5, 7, 9, and 11 [Doty (Smith), 1989; Hartmann and Doty, 1996]. However, there was one glaring exception, namely the mistuned fundamental. The model said that the pitch shifts should be unusually small and opposite in sign to the mistuning. In contrast, experiment showed unambiguously that the mistuned fundamental behaves just like any other mistuned harmonic—pitch shifts are substantial and in the same direction as the frequency shift.

This difficulty led us to consider an alternative paradigm. Terhardt's excitation pattern model and the timing model of Hartmann and Doty are local interaction models is the sense that the pitch shifts depend mainly on neighboring harmonics. The two nearest neighbors are especially important. A limitation of such local interaction models is that they give no recognition to the role of the entire spectrum in integrating a correctly tuned harmonic into a complex tone, or to the role that a broadband harmonic template might play when the harmonic is segregated by mistuning. A harmonic template is an internalized pattern for the components of a periodic tone. Current models of complex tone pitch perception (Goldstein, 1973; Terhardt, 1974) say that the pitch is the fundamental frequency, or harmonic spacing, of the best fitting template. Further, the ''pitch meter'' studied by Duifhuis *et al.* (1982) and Scheffers (1983) uses a self-consistent harmonic template to determine whether each resolved spectral component of the physical tone should be integrated into a complex tone percept or should be segregated. The experiments of this article were designed to search for a role for a broadband template and to establish a contrast between template models and local interaction models in the pitch of mistuned harmonics.

I. EXPERIMENTS

A. Method

Experiments were performed to measure a quantity defined as the *pitch shift gradient*. The gradient concept originated in the observation that pitch shifts for mistuned harmonics are in the same direction as the frequency mistuning. Therefore, if one takes the difference between the pitch shifts observed for positive and negative mistunings, the result should be larger and more robust than either pitch shift alone. An example is shown in Fig. 1; part (a) plots pitch, and part (b) plots pitch shift. The pitch shift observed for a mistuning of $+8\%$ is $\Delta p_+ = 1.1\%$. The pitch shift observed for a mistuning of -8% is $\Delta p_- = -1.3\%$. The pitch shift gradient is $\Delta p_+ - \Delta p_- = 2.4\%$, and there is an average pitch shift of $(\Delta p_+ + \Delta p_-)/2 = -0.1\%$.

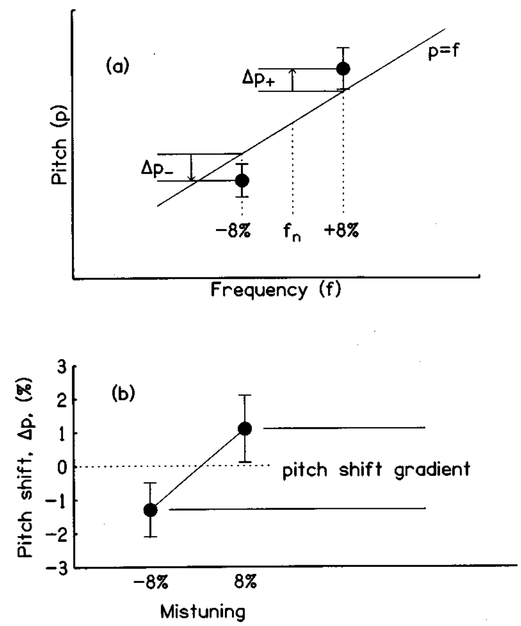

FIG. 1. Pitch shift gradient. A positive pitch shift gradient occurs when pitch changes more rapidly than frequency. Part (a) shows the mean and standard deviation of pitch matching data for a positively mistuned harmonic $(+8\%)$ and a negatively mistuned harmonic (-8%). In the absence of a gradient, the pitch shifts Δp_+ and Δp_- would be the same. In fact, however, these shifts even have different signs. Part (b) plots the values of Δp_+ and Δp_- to show the pitch shift gradient. This gradient would be called *large* because the error bars for matching positive and negative mistunings do not overlap. The error bars are two standard deviations in overall length. All the data presented in this article are represented by graphs like part (b). To measure a pitch shift gradient does not necessarily require positive and negative mistuning of a harmonic. A gradient can be measured using any two frequencies that are not greatly different.

A pitch shift gradient is a perturbation in the smooth functional relationship between pitch and frequency. It occurs when, over a small range in frequency, the pitch changes differently from the frequency itself. Therefore, a pitch shift gradient can be measured by measuring the percentage pitch shifts for any two frequencies that are rather close together and taking the difference. If the measured pitch shifts turn out to be a constant percentage of the frequencies, then the pitch shift gradient is zero. The experiments below all used mistunings of plus and minus 8% to measure the gradients. Doty's data indicate that pitch shifts grow with increasing mistuning up to mistunings of about $\pm 4\%$ and remain flat, or decrease slightly out to $\pm 8\%$. Therefore, $\pm 8\%$ is a reasonable standard. The results would likely have been the same had the mistunings been as small as $\pm 4\%$.[2]

1. General procedure

The listener was seated in a sound-treated enclosure, holding a response box that controlled the events of an experimental trial. The listener pressed a yellow button to hear a complex tone with one of its harmonics mistuned. The listener pressed an orange button to hear a matching sine tone with a frequency that the listener could adjust by means of a ten-turn potentiometer on the box. Each tone was preceded by a 300-ms silent interval to minimize interactions between intervals. The listener could call up the complex tone or the matching tone in any sequence, with no time limit. When satisfied with the match, the listener pressed the green button to finish the trial. The stimulus and matching frequencies were recorded, and then the next trial began, with a different complex tone. There was no feedback to the listener.

The experiments used a wide variety of tones with as few as three and as many as six different complex spectra in a run. Each experimental run included at least one presentation of each spectrum with a mistuned harmonic mistuned by $+8\%$ and -8%. Therefore, a run included from 6–12 trials. It took from 10–20 min for a listener to finish a run. After a run was completed, the listener could come out to rest. Runs continued until each listener had done at least 12 runs in each experiment. Every data point shown in the figures that follow was based on the final 12 matches.

2. Stimuli

The complex tones with mistuned harmonics contained from 8–16 partials of equal amplitude. The fundamental frequencies varied from 150–800 Hz. Complex tones were generated by sound files, which were specific as to the fundamental frequency, the harmonic content, the mistuned harmonic number, and the percentage of mistuning—either plus or minus 8%. For a given trial, the appropriate sound file was loaded into a digital buffer 16k (16 384) samples long and was converted by a 16-bit DAC at a nominal sample rate of 16k/s. To prevent the listener from using pitch memory in 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. Because of this randomization, all the numerical values of frequencies given in this article are nominal values.

The output of the DAC was low-pass filtered at 7 kHz, -115 dB/oct, and then 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 1 s. The signal was presented to the listener via Sennheiser HD480 headphones such that each component had a level of 58 dB SPL. Therefore, a 16-component complex tone had a level of 70 dB SPL.

The matching tone was generated by repeatedly cycling a buffer using fractional-addressing technology. One cycle of a sine wave was loaded into a 16k buffer and sampled at the rate of 32k/s. The fractional address increment was determined by the potentiometer on the control box, as read by a 12-bit ADC. An exponential frequency control law was applied in software, leading to a frequency resolution of 0.06%.

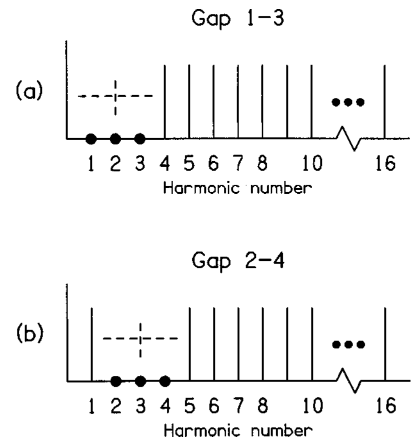

FIG. 2. Spectra for experiment 1. (a) Gap 1–3. A single mistuned harmonic (number 1, 2, or 3) added to harmonics 4–16. Vertical solid lines represent perfectly tuned harmonics. A vertical dashed line represents a mistuned harmonic. A dot is put on the horizontal axis where a harmonic is omitted. (b) Gap 2–4.

The matching tone was filtered and enveloped like the complex tone. The matching tone level was 55 dB SPL, 3 dB lower than a component of the complex tone, causing the matching tone to be approximately as loud as the mistuned harmonic. It was expected that equal loudness would make the pitch matching task easier.

3. Listeners

There were four listeners, three males, B, J, T, and one female, C. Their ages ranged from 19–55. All the listeners had negative otological histories and had some training as performers of musical instruments. They could perform accurate sine–sine pitch matching in the range 150–1000 Hz. Listeners J and B were the authors.

B. Experiment 1

1. Spectra of the complex tones

The spectra of the two complex tones in experiment 1 are shown in Fig. 2. The stimulus in Fig. 2(a) was made by starting with a periodic complex tone with a nominal fundamental frequency of 200 Hz and harmonics 4–16. Then one (and only one) of the omitted harmonics (1, 2, or 3) was reinserted as a mistuned harmonic, mistuned by either $+8\%$ or -8%. There were, therefore, six different stimuli in this experiment. This experiment is called ''gap 1–3.''

The stimulus in Fig. 2(b) was made in the same way except that harmonics 2, 3, and 4 were initially omitted. Then one of those harmonics was reinserted as a mistuned harmonic, mistuned by either $+8\%$ or -8%. This experiment also had six different stimuli, and it is called ''gap 2–4.''

2. Results—The zigzag effect

Figure 3 shows the results for gap 1–3 and gap 2–4 experiments, one panel for each subject. There are several observations to be made: First, there is a large pitch shift gradient for each plus–minus mistuned pair. Second, there are average shifts for some of the pairs, though these differ in size from subject to subject. Finally, although the magni-

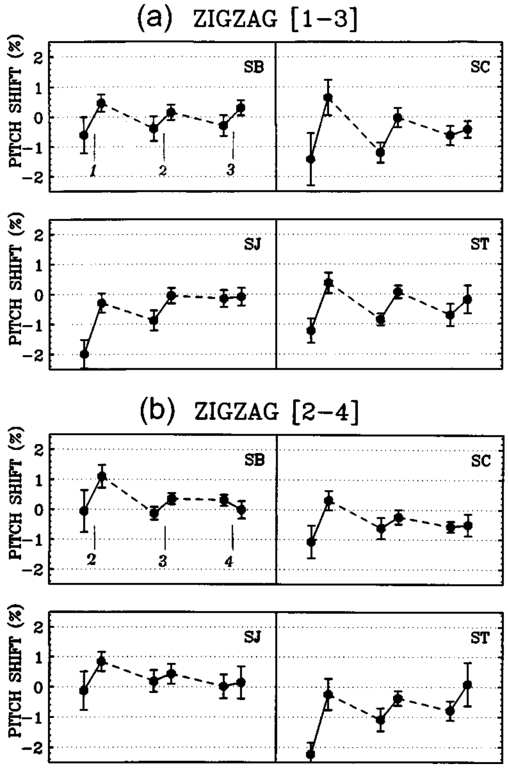

FIG. 3. Results of experiment 1. There are four panels (a) for the gap 1–3 experiment, and four panels (b) for the gap 2–4 experiment. In each panel there are three pairs of data points representing the ±8% mistuned pairs for three mistuned harmonics. Members of a pair are connected with a solid line so that one can compare the gradients of different pairs. A dashed line connects pairs to illustrate the ragged dependence of the pitch shift on the mistuned harmonic frequency, indicating the *zigzag effect*. The horizontal axis is monotonic (but not linear) with mistuned harmonic frequency.

tudes of the gradients are different for different listeners, there is a common tendency for low mistuned harmonics to have larger gradients. This low-frequency tendency will be seen in other experiments too.

Experiment 1 shows that the pitch shifts do not vary smoothly with mistuned harmonic frequency. The patterns shown in Fig. 3(a) and (b) are zigzags. These zigzag functions are evidence against an important class of models in which pitch shifts are mediated by local interaction. The gap 1–3 experiment is critical. The stimuli for gap 1–3 can be regarded as six target sine tones, increasing in frequency from harmonic 1 mistuned by −8% to harmonic 3 mistuned by +8%. As its frequency increases, the target tone gets closer to the rest of the spectrum (harmonics 4–16), which is responsible for the pitch shifts. Local interaction models based on the interaction of excitation patterns logically require that a shift varies smoothly with the separation between the target and the rest of the spectrum because the interaction decreases monotonically with increasing separation in local models. This prediction is contrary to the zigzag pattern observed experimentally. Local interaction models based on timing, like Hartmann–Doty, do not necessarily predict a smooth behavior for the gap 1–3 experiment, but the peripheral filtering in these models causes them to predict that gradients should increase in magnitude as the mistuned harmonic number increases from 1 to 3. In fact, the zigzag pattern does the reverse. It is largest for mistuned harmonic

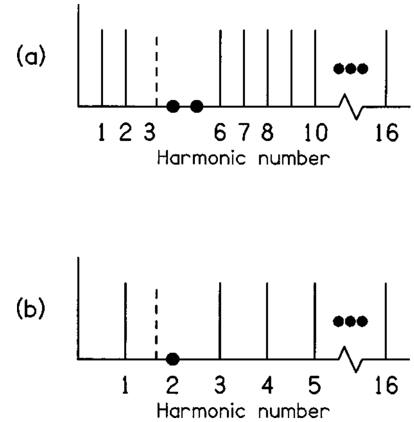

FIG. 4. Spectra for experiment 2. Part (a) shows mistuned harmonic number 3. Part (b) shows mistuned harmonic number 1.5. A dot appears on the horizontal axis where a harmonic is omitted.

number 1 and smallest for number 3. The zigzag pattern seen for gap 1–3 occurs also for the gap 2–4, except that the gradients are somewhat smaller in gap 2–4.[3]

The results of experiment 1 suggest that local interaction models are not appropriate models for the pitches of mistuned harmonics. Instead, it suggests that there is something special about the integer harmonic numbers, even when they are in a gap. A large pitch shift gradient occurs when the pair straddles a special frequency where a harmonic is supposed to be. This, in turn, is evidence that the complex tone forms a harmonic template within the auditory system, signalling the location of harmonics.

C. Experiment 2

1. Spectra of the complex tones

The purpose of experiment 2 was to make a direct comparison between two stimuli: (a) in which the target tone was a mistuned harmonic, and (b) in which it was not. If a harmonic template is the dominant influence on the pitch shifts of mistuned harmonics, as suggested by experiment 1, then shifts should be evident for the mistuned harmonic in (a) but not for the target in (b). The key to experiment 2 was that the local spectral structure around the target was similar for both stimuli. Therefore, local interaction models predict similar pitch shifts for both (a) and (b).

The spectra for the tones of experiment 2 are shown in Fig. 4(a) and (b). For stimulus (a), the fundamental frequency was 200 Hz. Harmonics 4 and 5 were omitted, and the third harmonic was mistuned, either by +8% or −8%, i.e., it was nominally either 648 or 552 Hz. For stimulus (b), the fundamental frequency was 400 Hz, and the second harmonic was omitted to make a gap. A target, having the same frequency as the target in stimulus (a), was inserted in the gap. Figure 4(b) shows that the target, with frequency of 648 or 552 Hz, looks like a mistuned component with harmonic number 1.5. Therefore, stimulus (b) makes an adequate test: the harmonics neighboring the target have the same frequencies as in stimulus (a) [the spectral differences between (a) and (b) are six (Cambridge) critical bandwidths away from the target region], but 1.5 is not a legitimate harmonic number.

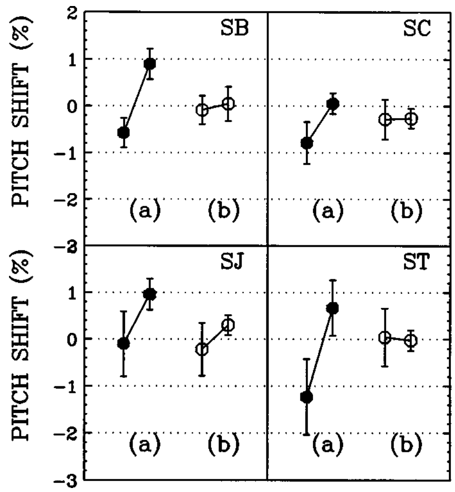

FIG. 5. Results of experiment 2. There are four subjects. Letters (a) and (b) refer to spectra in Fig. 4. The gradient is seen to be smaller in case (b), mistuned harmonic number 1.5, indicating the *integer effect*.

2. Results—The integer effect

The results of experiment 2 are shown in Fig. 5, one panel per subject. For all the subjects there are substantial shift gradients for stimulus (a), where the frequencies of the pair straddle harmonic 3, but there are no similar gradients for stimulus (b), where the frequencies of the pair are not close to an integer harmonic. The effect is significant $[F(1,3)=16, p<0.03]$. The result of experiment 2 is called the ''integer effect,'' meaning that large pitch shift gradients occur when the frequencies of the pair straddle integer harmonic frequencies. The integer effect suggests that a harmonic template plays an important role in the shifts.

For local interaction models, one expects pitch shifts to be mainly determined by the closest frequency components. Therefore, local interaction models tend to predict approximately the same pitch shifts for stimuli (a) and (b). Specific calculations with the tonotopic model of Terhardt *et al.* (1982a, 1982b) and the timing model of Hartmann and Doty both predict that (a) and (b) should have the same pitch shifts to five significant figures. These predictions are contrary to the large difference between (a) and (b) shown in Fig. 5.

D. Experiment 3

1. Spectra of the complex tones

The main purpose of experiment 3 was to see whether spectrally distant harmonics contribute to the pitch shift of a mistuned harmonic. The experiment used five complex tones, all with a period of 1/200 s. All had a mistuned fundamental. The spectra are shown in Fig. 6: Tone (a) was a 16-component complex tone. Tones (b)–(d) were the same as tone (a) but with increasing number of omitted harmonics. Tone (e) had only harmonics 1, 2, and 3.

2. Results—The proximity effect

The results of experiment 3 are shown in Fig. 7. Because listener SB did not complete this experiment, data are shown for only three listeners; the results are essentially the same for all three. As the neighboring harmonics 2, 3, and 4 were

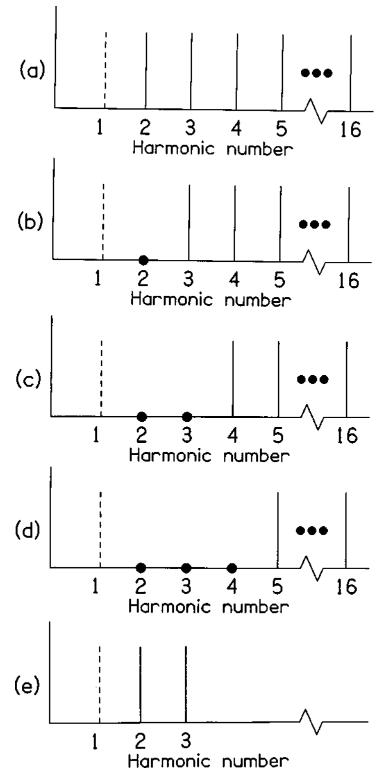

FIG. 6. Spectra for experiment 3. Tones (a)–(e) have a mistuned fundamental. Harmonic spectral blocks run from proximate (a) to remote (d). Tone (e) had only three harmonics. Dots on the horizontal axis mark the positions of omitted harmonics.

omitted one by one, the pitch shift gradient became smaller, but some gradient was always present. In only one case, tone (d) for SJ, was the gradient less than the error bars.

The results of the experiment show that distant harmonics are capable of generating a pitch shift gradient. Because the distant harmonics were well outside the critical band containing the target, the pitch shifts cannot be understood from a model that treats only a single peripheral channel. Specific calculations using Terhardt's model show negligible pitch shift ($<0.1\%$) for a mistuned fundamental when harmonic 2

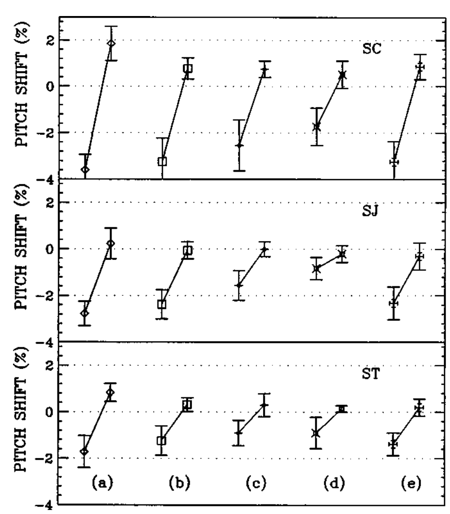

FIG. 7. Results of experiment 3. There are three subjects. Letters (a)–(e) refer to spectra in Fig. 6. Pitch shift gradients decrease in order (a) through (d), indicating the *proximity effect.*

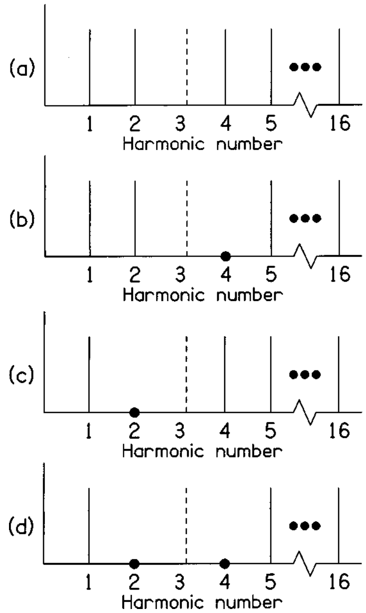

FIG. 8. Spectra for experiment 4. A mistuned third harmonic in symmetrical [(a) and (d)] and asymmetrical [(b) and (c)] harmonic environments. Dots on the horizontal line indicate omitted harmonics.

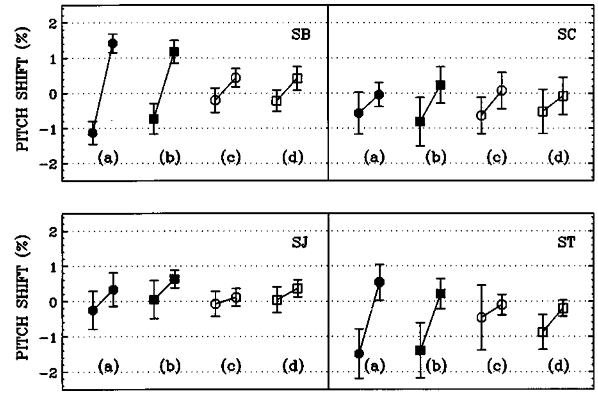

FIG. 9. Results of experiment 4. There are four subjects. Letters (a)–(d) refer to spectra in Fig. 8. Larger gradients occur for tones (a) and (b) where the second harmonic is present, indicating a *local asymmetry effect*.

is omitted and even smaller shifts for other omitted-harmonic stimuli. The Hartmann–Doty model underpredicts the pitch shift for the mistuned fundamental, as always.

The fact that appreciable pitch shift gradients were caused by distant harmonics suggests that pitch shifts originate at a higher level, where the outputs of tuned channels are combined. This result is consistent with a template model for the shifts.

A second result of experiment 3 is that the shift gradients were smaller when the target was farther from the harmonics of the complex tone. This result is a ''proximity effect,'' which might be regarded as evidence for the importance of a local interaction such as partial masking. However, the large size of the shifts themselves in Fig. 7 seems out of proportion to the small weight that local interaction plays in other experiments, such as experiment 2. It seems more likely that the proximity effect represents increasing influence for a harmonic template when the components that establish the template become closer to the target.

E. Experiment 4

1. Spectra of the complex tones

Experiment 4 created spectral asymmetries for the mistuned harmonic target by selectively eliminating nearest-neighboring harmonics. The motivation for this experiment was that both the Terhardt place model and the Hartmann–Doty timing model are sensitive to local spectral structure. The former is particularly sensitive to the removal of a lower harmonic; the latter is sensitive to both lower and upper neighbors of the mistuned harmonic. Experiment 4 looked for the expected large changes in average shifts.

The target was a mistuned third harmonic, mistuned by $+8\%$ or -8%. The fundamental was 200 Hz. The spectral contexts are shown in Fig. 8: In tone (a) all harmonics were present. In tone (b) the fourth harmonic was omitted. In tone (c) the second harmonic was omitted. In tone (d) both the second and the fourth harmonics were omitted. The Terhardt model predicts that eliminating the second harmonic changes the average shift from 1.9% to 0.5%. The Hartmann–Doty model predicts that all shifts, regardless of mistuning, are positive for an omitted second harmonic, and all shifts are negative for an omitted fourth harmonic.

2. Results—The local asymmetry effect

The data from experiment 4 are shown in Fig. 9, one panel for each subject. The first observation to be made is that the large changes in average shifts (average over positive and negative mistunings) predicted by local interaction models for asymmetrical spectra (b) and (c) did not materialize. The changes in average shifts are small, and the directions of the changes are different for different listeners. This result is evidence against local interaction models. The largest effect seen in this experiment was a reduction in gradient caused by removing the second harmonic (spectra c and d), whether or not the fourth was removed. For SB and ST the change was dramatic. Overall, the reduction in pitch shift upon removal of the second harmonic was significant [$F(1,7)=11.3$, $p<0.02$]. The special importance of the second harmonic in producing the gradient, in comparison with the fourth, is a *local asymmetry*. Generalized from this experience with a mistuned third harmonic, a statement of the local asymmetry effect says that the harmonics below the target are more important than the harmonics above the target in making a pitch shift gradient. Like the proximity effect, the local asymmetry effect has some of the character of a local interaction model. The upward spread of excitation might be expected to result in just this kind of asymmetry. However, local interaction models have the problem that an interaction large enough to generate a large change in gradients (e.g., SB and ST) also tends to generate a large shift in average pitch, which is not seen experimentally.

F. Experiment 5

1. Spectra of the complex tones

Experiment 5 was done to see the how pitch shift gradients for the mistuned fundamental depend on the period of the complex tone. There were three complex tones with eight

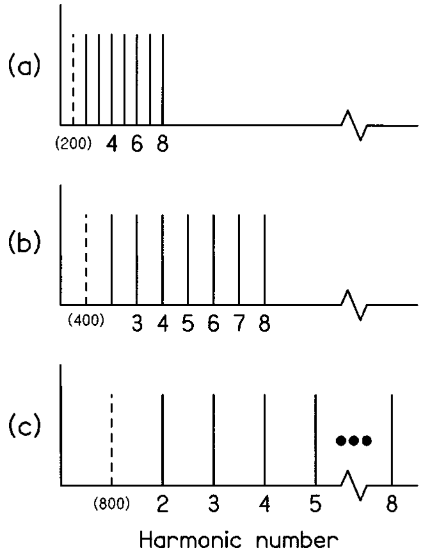

FIG. 10. Spectra for experiment 5. The three complex tones are scaled so that nominal fundamental frequencies are 200, 400, and 800 Hz. The fundamentals are mistuned.

components and mistuned fundamentals. The only difference among the tones was the nominal fundamental frequency: (a) 200 Hz, (b) 400 Hz, and (c) 800 Hz, as shown in Fig. 10.

2. Results—The frequency effect

The results of Experiment 5 are shown in Fig. 11. The results clearly show that the pitch shift gradient decreases with increasing fundamental frequency. Whereas the gradients at 200 Hz are large, greater than 2% for all listeners, the gradients at 800 Hz are almost zero. This is the frequency effect. The gradient is larger for 200 Hz than for 400 Hz [$F(1,3)=53$, $p<0.01$], and it is larger for 400 Hz than for 800 Hz [$F(1,3)=24$, $p<0.05$].

The Terhardt model predicts the frequency effect, largely because of the increased sensitivity to SPL excess as the frequency decreases. It also predicts the dominance of negative pitch shifts seen experimentally. However, the predicted pitch shift gradient has the wrong sign, and, therefore, the model fails. The Hartmann–Doty model fails for a mistuned fundamental, as usual, because it predicts that the gradient is too small to be observed, whatever the fundamental frequency.

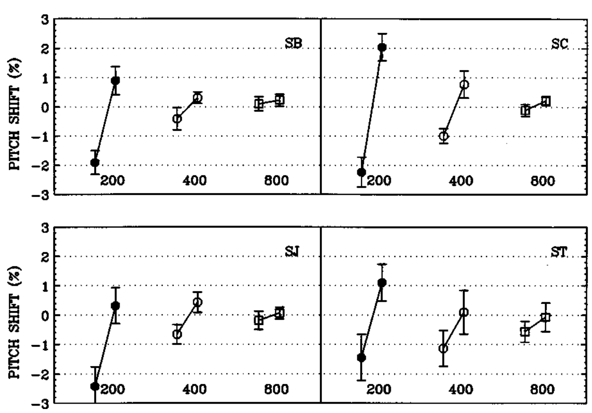

FIG. 11. Results of experiment 5. There are four subjects and three fundamental frequencies with spectra given in Fig. 10. The gradients are seen to be smaller for higher fundamental frequencies, indicating a *frequency effect.*

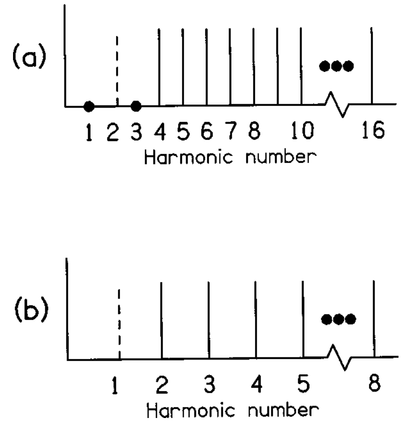

FIG. 12. Spectra for experiment 6. A mistuned ''400-Hz'' component serves as a mistuned second harmonic (a), or as a mistuned fundamental (b).

The small gradients seen at 800 Hz might have been predicted from a harmonic template model because periodicity pitch is weak at 800 Hz. Modern pitch perception models suggest that periodicity pitch, as seen in tones with missing fundamentals, is the result of a central harmonic template. The DWS pitch meter is an example (Scheffers, 1983). According to Ritsma's 1962 amplitude modulation experiments, 800 Hz is the upper limit for the existence of periodicity pitch. It seems natural to attribute this limit to a weakness in template influence for fundamental frequencies as high as 800 Hz. An ineffective template, in turn, would produce small pitch shift gradients, as observed in experiment 5. A possible further interpretation for the results of experiment 5 is that templates with low fundamental frequencies or long periods (e.g., 1/200 s are more effective than templates with short periods (e.g., 1/800 s). This interpretation was tested in experiment 6.

G. Experiment 6

1. Spectra of the complex tones

Experiment 6 had two goals. The first was to test the template period conjecture from experiment 5, the second was to explore the dependence of pitch shift gradients on harmonic number. The idea was to create a similar local spectral environment for mistuned harmonics, subject to two different overall periodicities. The spectra are shown in Fig. 12. In tone (a) the second harmonic of the complex tone was mistuned and its neighbors (the fundamental and the third harmonic) were omitted. The fundamental frequency was 200 Hz. In tone (b) the stimulus was a 400-Hz complex tone with its fundamental mistuned. The local spectral structures for the mistuned harmonic were similar for both (a) and (b), but the harmonic number in (a) was 1 (the fundamental), and the harmonic number in (b) was 2 (the second harmonic). Both stimuli had the same power. According to local interaction models both should have the same pitch shift (less than 0.2% in Terhardt's model).

2. Results

The results of experiment 6 are shown in Fig. 13. For all the subjects, the lower harmonic number (the fundamental)

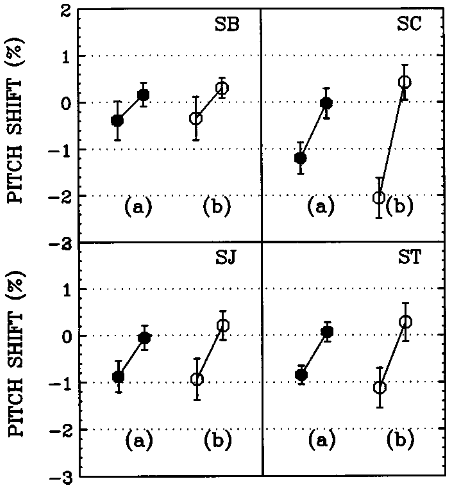

FIG. 13. Results of experiment 6. There are four subjects. Letters (a) and (b) refer to spectra in Fig. 12. The gradients for (b) tones are slightly larger, indicating a *harmonic number effect*.

suffered the larger gradient. This is a harmonic number effect. The effect is not significant according to our usual test $[F(1,3)=4.4]$ because the size of the difference varies among subjects leading to a large variance. It is important, however, that the reverse effect did not occur.

The results of experiment 6 affect our interpretation of experiment 5. If long-period templates are more effective than short, as conjectured in the discussion of experiment 5, then experiment 6 should find a larger pitch shift gradient in case (a) (200-Hz complex tone) than in case (b) (400-Hz complex tone). In fact, this did not occur. Instead, the larger gradient occurred for the 400-Hz complex tone. Therefore, our interpretation of the frequency effect in experiment 5 is that templates are less effective in shifting the pitches of higher-frequency fundamentals.

II. DISCUSSION AND CONCLUSION

A. Outcomes and interpretations

The pitch matching experiments of this article confirmed that the pitch of a mistuned harmonic differs from its frequency and that the shift in pitch exaggerates the frequency difference caused by mistuning. The experiments used computer-generated complex tones with specially tailored spectra. The principal goal was to distinguish between two classes of models for the pitch shifts, one based on tonotopically local interactions, the other based on a broadband template.

In a local interaction model, either place based or timing based, the pitch shifts are attributed to the effect of neighboring harmonics. The effects of distant harmonics are imagined to be attenuated by peripheral auditory filters. It is reasonable to try to apply local interaction models to mistuned harmonic pitch because these models have provided a qualitative account of many other pitch shift effects that occur for sine tones when noise or other tones are added to the signal (Webster and Muerdter, 1965; Terhardt, 1971; Rakowski and Hirsh, 1980; Burns and Osterle, 1980; Houtsma, 1981; Hartmann, 1996, 1997).

In a template model, by contrast, it is hypothesized that pitch shifts result from the influence of a broadband framework established by all the harmonics of the complex tone. Because the template model is more central than local interaction models, it is less clear how one should proceed to make it quantitative. The fact that pitch shifts are almost always found to be exaggerations of the frequency mistuning means that the template acts in such a way as to enhance contrasts. The enhancement favors the segregation of the mistuned harmonic as an independent entity.

Important evidence in favor of the template model came from the *zigzag* effect in experiment 1, where the presence of upper harmonics led to a ragged pitch shift pattern for mistuned lower harmonics. We found it helpful to think about this experiment in the following physical terms: In experiment 1 (gap 1–3), there is an empty spectral region where harmonics 1–3 have been omitted. Above this region there is spectral strength in harmonics 4–16. Experiment 1 consisted of placing a sine tone probe somewhere in the empty spectral region and observing ''forces'' on it caused by harmonics 4–16. The imagined forces are the influences that shift the pitch of the sine tone. The zigzag effect indicated an apparent force whereby pitches are ''repelled'' from sites where harmonics ought to be. The gap 2–4 variation of experiment 1 could be analyzed similarly.

The *integer* effect, found in experiment 2, gave additional evidence for the influence of a harmonic template because no significant pitch-shifting forces were found for sine tones half way between harmonic sites. By contrast, a large pitch shift gradient appeared, for the same local spectrum, when the sine tone was near a harmonic site.

Experiments 3 and 4 led to the *proximity* and *local asymmetry* effects, which indicated a role for local interaction. These effects are strong enough that any complete model for pitch shifts must take them into account. These effects essentially showed that spectral components near the mistuned harmonic target are more important in shifting the pitch of the target than are remote components. This result suggests a role for peripheral filtering of the kind that is always observed in masking experiments, but it is not clear that the effects actually resemble masking. Masking models, or other local interaction models consistent with the proximity effect, also tend to predict large average pitch shifts for asymmetrical spectra, and these were not observed. Instead, the proximity and local asymmetry effects may indicate a restricted ''range of interaction'' for a template.

Experiment 5 studied mistuned fundamentals for three tones with spectra that were identical except for octave scaling factors. This experiment found that a low-frequency mistuned fundamental was shifted considerably more than a high-frequency mistuned fundamental. We interpreted this observation as a *frequency* effect, which says that a template is more effective at a low frequency, such as 200 Hz, than at a higher frequency, such as 800 Hz. Such a frequency effect is well known in experiments that study the strength of periodicity pitch, which has often been attributed to template matching, and it was gratifying to find it in our experiment. An alternative interpretation for experiment 5 would be that templates with long periods are more effective than templates with short periods. However, evidence contrary to the latter interpretation was found in experiment 6, which systematically changed the period of the template. In general, experiments 3–5 exhibited larger pitch shift gradients for lower frequencies.

B. Caveats and conjectures

The experiments reported in this article measured pitch shifts, defined as the difference between the frequency of a mistuned harmonic and the frequency of a matching sine tone. A measured shift is, therefore, the change in the pitch of a sine tone when it is embedded in an otherwise periodic complex, with the period chosen in such a way that the sine tone becomes a mistuned harmonic. To avoid misinterpretation, it is important to note that our experiments did not measure the change in the pitch of a harmonic as the harmonic frequency changes from correctly tuned to mistuned. Therefore, our measured shifts can be interpreted as the pitch shifts caused by mistuning a harmonic only if it is assumed that the pitch of a correctly tuned harmonic is equal to the frequency of the harmonic. The preponderance of evidence suggests that this this assumption may be valid. In 1971 and 1979 Terhardt reported that the pitches of harmonics of a complex tone differed from their frequencies. However, extensive experiments by Peters *et al.* (1983) failed to confirm that result. Further, when the zero-frequency-shift data were extracted from the pitch measurements made by Doty (1989), it was found that harmonic pitches were statistically the same as harmonic frequencies, in agreement with Peters *et al.* Therefore, it is likely that our measurements actually do reflect the change in pitch as a spectral component becomes inharmonic, even though that is not what we measured. Our focus on pitch shift gradient made it unnecessary to measure the pitch shift caused by mistuning *per se*.

The focus on pitch shift gradient has another advantage, especially when comparing data to theory. There are pitch shifts with signal level, even for the sine-wave matching tone, and these effects are included in models such as Terhardt's. However, these level effects cancel in the measured gradient. Therefore, they do not confound the spectral context effects of interest here.

As noted in the Introduction, mistuning a harmonic in a complex tone leads to a shift in the low pitch (virtual pitch) of the tone overall as well as a shift in the pitch of the harmonic. Possibly these two shifts are related. Specifically, it might be imagined that pitch of the mistuned harmonic is shifted in order to agree with a harmonic of the shifted virtual pitch. This conjecture, however, does not square with the facts. According to the measurements of Moore *et al.* (1985), changing the frequency of a harmonic by a certain percentage leads to a change in virtual pitch that is always a considerably smaller percentage. By contrast, changing the frequency of a harmonic leads to a *larger* percentage change in the pitch of that harmonic.

It seems evident that the pitch shift observed for a component that has been perceptually segregated from a complex tone is likely to be closely associated with the process that segregates the component in the first place. It is the pitch, after all, that identifies a perceptually segregated component in a mistuned harmonic experiment. Recently, de Cheveigné (1997) has proposed a statistical decision theory model that successfully predicts the pitch shift of a mistuned harmonic as a function of the mistuning as measured by Doty. The model begins by assuming that a component of a complex tone can be represented by a distribution of pitch cues on an internal tonotopic coordinate. Pitch cues that are close to a harmonic place are perceptually integrated with the rest of the complex tone. Pitch cues that are more remote lead to a segregated tone. Precisely because they are more remote, these pitch cues lead to a pitch shift for the segregated tone. Although the model in its present version cannot account for both the pitch shifts and the high hit rates for mistuned harmonic identification found by Hartmann and Doty, this model does appear to be consistent with the pitch shift effects reported in the present article.

III. CONCLUSION

When a harmonic of a complex tone is mistuned, the harmonic can be heard out as a separate entity, distinct from the complex tone background. Pitch matching experiments show that the pitch of the mistuned harmonic does not agree with its frequency. The shift in pitch is, very reliably, found to be an exaggeration of the physical mistuning.

This article considered two possible kinds of model to account for the pitch shift. A tonotopically local model attributes the shift to the influence of neighboring harmonics. The influence may be the result of partial masking of firing rate excitation patterns (Terhardt) or it may stem from alterations in neural spike timing, as appear in the interspike interval histogram (Hartmann–Doty). In either case the effects are restricted by peripheral filtering. A second class of model attributes the pitch shift to a contrast between the mistuned harmonic and a spectral pattern in the mind of the listener that serves as a template against which tones are compared.

To try to distinguish between the two models, pitch matching experiments were done using special spectra. By changing the spectrum near the mistuned harmonic while maintaining the best fitting template (experiment 1) or by changing the template while maintaining the local spectrum (experiment 2) we were able to show a dominant role for a template. Scaling the frequency range of the matching experiment showed that pitch shifts tend to vanish for mistuned fundamentals near 800 Hz, similar to the limit for virtual pitch formation (experiment 5). This agreement is expected if both the pitch shift and virtual pitch are derived from a template.

Further experiments that maintained the best-fitting template while changing the spectrum near the mistuned harmonic (experiments 3 and 4) indicated a role for the local spectrum, which altered the details of the pitch shifts but did not upset the basic pattern established by the template comparison process. In the end, the dominance of the template is overwhelming.

ACKNOWLEDGMENTS

This work was supported by the National Institute on Deafness and Other Communicative Disorders, Grant No.

DC00181. We are grateful to Brad Rakerd, Tim Vander Velde, Alain de Cheveigné, and Ray Meddis for useful discussions.

[1]Terhardt's semiempirical model is unspecific concerning a physiological mechanism for the pitch shift, but in the low-frequency range the model is consistent in spirit with the idea that pitch is determined by a moment (e.g., centroid) of a neural excitation pattern along the tonotopic axis, and that the effect of masking is to erode the excitation pattern caused by the target component, thereby reducing the statistical weight of target excitation closest to the place of the masker. All the mistuned harmonics in the present paper are in this low-frequency range.

[2]Mathematically, a ''gradient'' is a slope, calculated as the ratio of rise-to-run on a graph. Our measure of the ''pitch shift gradient'' would have been more consistent with this definition had we divided our data by the 16% change in mistuned harmonic frequency [$+8\%-(-8\%)=16\%$]. Doing so, however, would have been misleading because the pitch shift saturates as a function of frequency shift. Doty's 1989 measurements show that pitch shifts measured at $\pm 8\%$ hardly differ from pitch shifts measured at $\pm 4\%$. Nevertheless, we use the term ''pitch shift gradient'' because the frequency change (the putative denominator) was always constant in our experiments, and because this term conveys the idea of a wrinkle in the pitch continuum that causes the pitch shift not to be a constant.

[3]Specific calculations using Terhardt's model show negligible pitch shift gradients for gap 1–3 and pitch shifts that are always positive and decreasing with increasing target frequency for gap 2–4. That occurs because this model gives great weight to low frequency partials. Predicted pitch shift gradients for gaps 2–4 have the wrong sign as usual.

Burns, E. M., and Osterle, E. L. (**1980**). ''Pitch shifts associated with partial masking,'' J. Acoust. Soc. Am. Suppl. 1 **67**, S20.

Darwin, C. J., Ciocca, V., and Sandell, G. F. (**1994**). ''Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic,'' J. Acoust. Soc. Am. **95**, 2631–2636.

de Cheveigné, A. (**1997**). ''Harmonic fusion and pitch shifts of mistuned partials,'' J. Acoust. Soc. Am. **102**, 1083–1087.

Doty (Smith), S. L. (**1989**). ''Pitch perception of mistuned harmonics in a complex tone,'' M.S. thesis, Michigan State University, unpublished.

Duifhuis, H., Willems, L. F., and Sluyter, R. J. (**1982**). ''Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception,'' J. Acoust. Soc. Am. **71**, 1568–1580.

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. (**1996**). ''Pitch periodicity, and auditory organization,'' J. Acoust. Soc. Am. **100**, 3491–3502.

Hartmann, W. M. (**1997**). *Signals, Sound, and Sensation* (AIP Press, Springer-Verlag, New York), Chap. 12.

Hartmann, W. M., and Doty, S. L. (**1996**). ''On the pitches of the components of a complex tone,'' J. Acoust. Soc. Am. **99**, 567–578.

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

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

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

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

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.

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

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

Scheffers, M. T. M. (**1983**). ''Simulation of auditory analysis of pitch: An elaboration on the DWS pitch meter,'' J. Acoust. Soc. Am. **74**, 1716–1725.

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

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

Terhardt, E. (**1979**). ''Calculating virtual pitch,'' Hearing 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.

Webster, J. C., and Muerdter, D. R. (**1965**). ''Pitch shifts due to low-pass and high-pass noise bands,'' J. Acoust. Soc. Am. **37**, 382–383.