

Hearing a mistuned harmonic in an otherwise periodic complex tone^{a)}

William Morris Hartmann

Physics Department, Michigan State University, E. Lansing, Michigan 48824 and Institut de Recherche et Coordination Acoustique/Musique, 31, rue Saint-Merri, F-75004, Paris, France

Stephen McAdams^{b)} and Bennett K. Smith

Institut de Recherche et Coordination Acoustique/Musique, 31, rue Saint-Merri, F-75004, Paris, France

(Received 12 March 1990; accepted for publication 19 June 1990)

The ability of a listener to detect a mistuned harmonic in an otherwise periodic tone is representative of the capacity to segregate auditory entities on the basis of steady-state signal cues. By use of a task in which listeners matched the pitch of a mistuned harmonic, this ability has been studied, in order to find dependences on mistuned harmonic number, fundamental frequency, signal level, and signal duration. The results considerably augment the data previously obtained from discrimination experiments and from experiments in which listeners counted apparent sources. Although previous work has emphasized the role of spectral resolution in the segregation process, the present work suggests that neural synchrony is an important consideration; our data show that listeners lose the ability to segregate mistuned harmonics at high frequencies where synchronous neural firing vanishes. The functional form of this loss is insensitive to the spacing of the harmonics. The matching experiment also permits the measurement of the pitches of mistuned harmonics. The data exhibit shifts of a form that argues against models of pitch shifts that are based entirely upon partial masking.

PACS numbers: 43.66.Fe, 43.66.Hg, 43.66.Jh, 43.66.Ba [WAY]

INTRODUCTION

In everyday circumstances listeners are exposed to sound fields that have been created by several different acoustical sources acting simultaneously. Therefore, a necessary aspect of human hearing is the ability to separate a composite sound field into distinct auditory images or perceptual entities, normally corresponding to the several different sources (Moore, 1982, Chap. 6; McAdams, 1984a, 1989). Because the channels of the auditory system are tuned, it would be natural for the system to segregate entities that are entirely distinct in their frequency ranges. But as Plomp (1976, p. 146) has noted, the system is frequently required to segregate sources with spectra that are interleaved. It is evident that auditory processes more intricate than simple filtering are involved. A review of some of these processes and the corresponding acoustical cues was given by Hartmann (1988).

In segregating the phonemes of speech (part of the cocktail party effect) or the tones of individual instruments in polyphonic music, asynchronous onsets provide a major cue to the auditory system (Bregman and Pinker, 1978; Rasch, 1978). Other temporal differences among sources, e.g., slow amplitude or frequency modulation, are also effective segre-

gators for tones (McAdams 1984a, b, 1989), but perhaps not for speech (Gardner and Darwin, 1986).

There are, in addition, cues to segregation present in steady-state signals. Important among these is the harmonicity or inharmonicity of the spectral components. When spectral components have frequencies that are in a perfect harmonic relationship, those components tend to be integrated by the auditory system into a single perceived entity. The presence of one or more inharmonic components in a spectrum promotes the segregation of two or more entities.

A simple psychoacoustical approach to the study of segregation based upon inharmonicity is the mistuned harmonic experiment. Here, the listener is presented with a complex tone composed of many harmonics, but one harmonic is mistuned from its correct harmonic frequency. If the amount of the mistuning is large enough, the listener hears a sine tone as a distinct entity segregated from the lower-pitched complex tone. The pitch of the segregated entity corresponds approximately to the frequency of the mistuned harmonic.

Early mistuned harmonic experiments (Moore *et al.*, 1985; Hartmann, 1985) were discrimination experiments. The listener heard two tones, one of them with perfectly harmonic components, the other with a mistuned harmonic. The listener was required to say which of the tones was inharmonic.

The discrimination experiment is straightforward and efficient. Unfortunately, the results of this experiment are difficult to interpret because the listener can actually make judgments on the basis of a number of different cues. Hartmann (1988) identified five such cues to inharmonicity;

^{a)} Some of the contents of this article were presented at the 112th Meeting of the Acoustical Society of America in Anaheim, CA [J. Acoust. Soc. Am. Suppl. 1 **80**, S93 (1986)].

^{b)} Present address: Laboratoire de Psychologie Expérimentale (CNRS URA 316), Université René Descartes (EPHE), 28 rue Serpente, F-75006 Paris, France.

their relative salience depends upon the amount of mistuning and upon the stimulus duration. Some of these cues, such as a shift of the low pitch or roughness, do not seem to be related to the segregation of a mistuned harmonic as a distinct entity. Therefore, while the detection of inharmonicity, as revealed in discrimination experiments may be of some interest in itself, the data from these experiments cannot be applied uncritically to the question of segregation of entities.

This conclusion was also reached by Moore *et al.* (1986) when they repeated their 1985 experiments, this time asking listeners to ignore all other cues and to respond positively only when they were sure that they heard out the mistuned harmonic. Not surprisingly, the results of this second series of experiments differed from the first. Thresholds were higher and the dependence upon mistuned harmonic number was much less apparent.

The goal of the experiments described in the present article was the same as in the discrimination experiments described above, namely, to probe the limits of a listener's ability to hear out a mistuned harmonic. The method, however, was different. The present experiments attempted to ensure that a positive response occurred only when the listener indeed heard out the mistuned harmonic by asking the listener to match the pitch of the segregated component. If the listener correctly matched the frequency of the mistuned harmonic, then that was taken as evidence that the mistuned harmonic was segregated because the probability of making a correct match by chance was small.

It was expected that this procedure would give results that are more reliable than either of the experiments by Moore *et al.* This also made it possible to extend the measurements to higher harmonic numbers (higher than the 6th), where interesting effects were expected. Further, we hoped to eliminate the considerable intersubject variability seen by Moore and his colleagues.

I. METHODS

A. Procedure

The listener was seated in a double-walled sound-treated room (Soluna SN 1). Before him was a response box with three push buttons, three lamps, and three potentiometers. When the listener pushed the green button, the computer chose randomly a particular frequency for the fundamental of the complex tone (randomized according to a rectangular distribution with a width of $\pm 10\%$ about the nominal fundamental) and a particular harmonic of the complex tone to be mistuned. When the listener then pressed the yellow button, he heard the complex tone; when he pressed the orange button, he heard a sine tone. Each tone was preceded by a delay. The frequency of the sine tone could be controlled by two ten-turn potentiometers, one a coarse control covering the range from 0–16 kHz, the other a fine control, which was ten times less sensitive.

The listener's task was to adjust the potentiometers so that the pitch of the sine tone matched the pitch of the harmonic heard out of the complex tone. There was no limit to the number of times that the listener could hear either tone. With the third potentiometer, a fader, the listener could ad-

just the level of the sine tone for his own convenience. When the listener was satisfied with his match, he again pressed the green button, and the trial with a particular fundamental frequency and mistuned harmonic was finished. At the end of a trial, the listener received feedback, a white lamp if his match was within ± 0.8 harmonic number of the correct value, a red lamp if his match was lower than this range and a blue lamp if higher.

Each experimental run consisted of 12 or 16 trials (matches) and on each trial a different harmonic of the complex tone was mistuned. The mistuning, measured as a percentage of the harmonic component frequency, was the same for each mistuned harmonic throughout the course of a run. Runs typically lasted 10 min after which the listener could come out of the sound-treated room and rest.

B. Listeners

The three authors B, S, and K served as listeners. They ranged in age from 33–47 years; all were males with normal hearing according to their own reports, and all were experienced in psychoacoustical tasks. They had musical experience and could sing the harmonic series up to the tenth harmonic (jumping octaves of course) or, in one case, the eighth.

C. Stimuli

Mainly, the stimuli for this experiment were a subset of those used by Moore *et al.* (1986). The nominal fundamental frequency was 200 or 400 or 800 Hz, and the duration of the complex tones was 410 or 50 ms. There were 16 harmonics of equal amplitude in the complex tone spectrum, except that there were only 12 when the fundamental frequency was 800 Hz. The low-frequency harmonics were candidates for mistuning: the first 16 for a fundamental of 200 Hz, the first 12 for fundamentals of 400 and 800 Hz. The proportional frequency deviation for mistuned harmonics was constant for a given run and varied parametrically between 0.5% and 8%. All the harmonics, including the mistuned harmonic, started at a positive-going zero crossing (sine phase). The complex tone had a standard level of 60 dB SPL per component; for low-level runs the level was reduced to 30 dB SPL per component. Both equalized and nonequalized tones were used (see the Appendix). Tones were turned on and off with a raised-cosine envelope of 10-ms duration, and the delay preceding each tone was 300 ms for 410-ms tones and 200 ms for 50-ms tones. The matching sine tone had a duration of 300 ms, and it was delayed similarly to the complex tones.

The stimuli were generated by the 4C digital synthesizer at the Institut de Recherche et Coordination Acoustique/Musique (Moorer *et al.*, 1979). The synthesizer was controlled by a PDP 11/34 computer, which also ran the experiment. Stimuli were converted to audio by Tim Orr 16-bit DACs at a sample rate of 32 kHz, and were low-pass filtered from 12.8 kHz (-96 -dB/oct slope). The listener heard the sound diotically through Beyer DT-48 headphones with flat foam cushions. Signal levels were measured at the headphones with a Bruel and Kjaer 2209 sound level meter (A-

weighted) with a flat plate coupler. The computer collected the responses and did some preliminary data reduction so that the listener could see his results after the end of the run. It stored the data for subsequent analysis.

D. Analysis

The raw data for a set of runs can be seen on a "matching plot," as shown in Fig. 1. For a given mistuned harmonic (horizontal axis), the plot shows each match, expressed on a continuous scale of harmonic number, obtained by dividing the listener's matching frequency by the fundamental frequency for the trial. Figure 1 shows the data for listener B when the mistuning was 4%. Although it may appear that only a few points are plotted for low harmonic numbers (e.g., number 1), there are actually ten points, one from each run of the experiment. The straight line shows idealized performance, matching frequencies identical to mistuned harmonic frequencies. Therefore, this line has a slope of 1.04.

In principle, further analysis of the data is a simple matter. If the matching frequency is close to the frequency of the mistuned harmonic, then the match is called a "hit." If the matching frequency is far from the frequency of the mistuned harmonic, then the match is called a "miss." From the pattern of hits and misses for the various experimental pa-

rameters, one can deduce the listener's ability to segregate mistuned harmonics from a complex tone.

In practice, the analysis is not so simple. There is first a problem in identifying a hit, because pitch shift effects are expected for the harmonics of a complex tone (Terhardt, 1971). The systematic departure of the data points from the straight line in Fig. 1 shows that such shifts indeed occurred. Second, the data show a continuum of matching frequencies. One needs a criterion for declaring a match to be a hit or a miss. After some experimenting, we established a twofold criterion. An absolute criterion required that a match be within one harmonic number of the mistuned harmonic ($1/2$ harmonic number for mistuned 1 and 2) in order to be a possible hit. The next requirement was based upon clustering of data, i.e., upon self-consistency. For the matches that survived the absolute criterion, a mean and unbiased ($N - 1$ weight) standard deviation were calculated. If the standard deviation was greater than 2.5% of the mistuned harmonic number, then the data point that differed most from the mean was discarded and a new mean and standard deviation were computed. The procedure was repeated until the 2.5% clustering criterion was met or the data set was exhausted. Figure 2 shows the results of applying the two-stage criterion to the data of Fig. 1. Hits are shown by + symbols; misses are shown by circles. Clearly, the performance of listener B for 4% mistuning was high, at least for low harmonic numbers.

Figure 3 shows a matching plot for listener B when the mistuning was only 0.5%. The number of hits was much

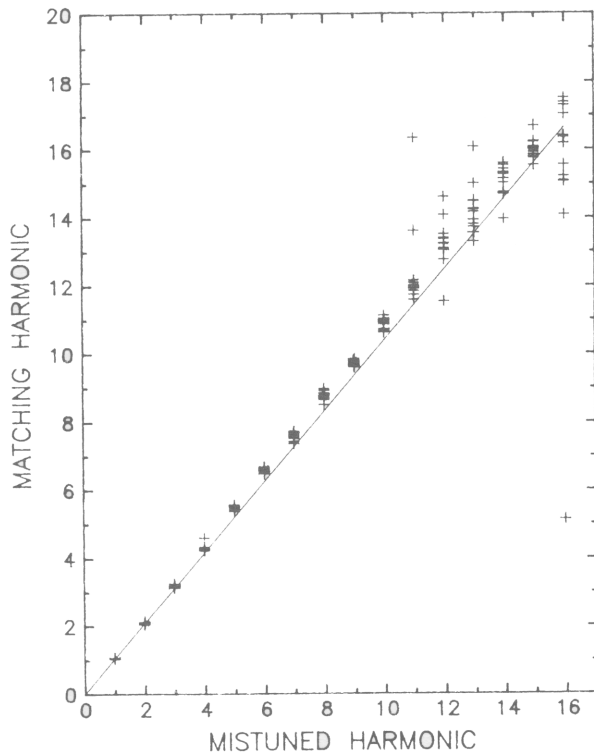


FIG. 1. Plot of the matching harmonic number as a function of the mistuned harmonic number for listener B. The complex tone had 200-Hz nominal fundamental frequency and 16 harmonics, each at a level of 60 dB SPL. The tone duration was 410 ms. The mistuned harmonic was mistuned by 4%. For each mistuned harmonic there were ten matches (symbol +) because there were ten runs. The solid line shows ideal performance where the matching frequency is identical to the frequency of the mistuned harmonic. The slope of that line is, therefore, 1.04.

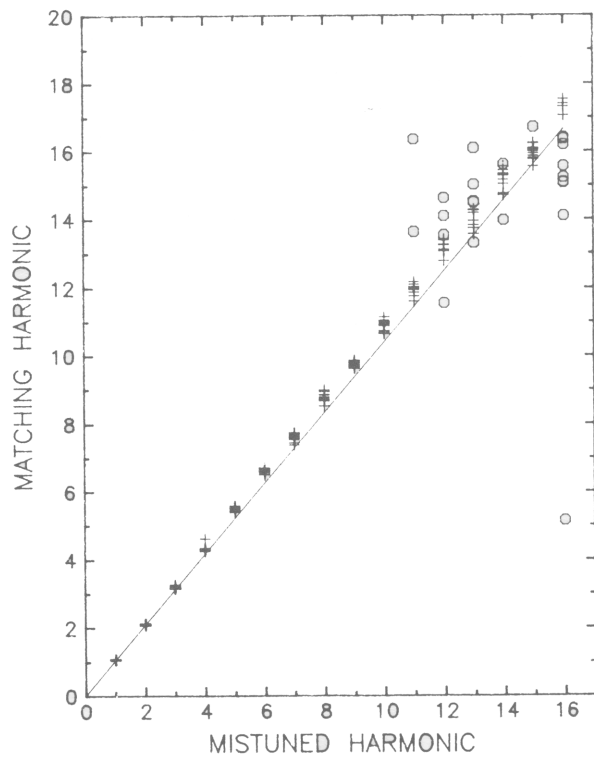


FIG. 2. The data of Fig. 1 (4% mistuning) are replotted with hits (+) and misses (O) identified according to the criteria described in the text.

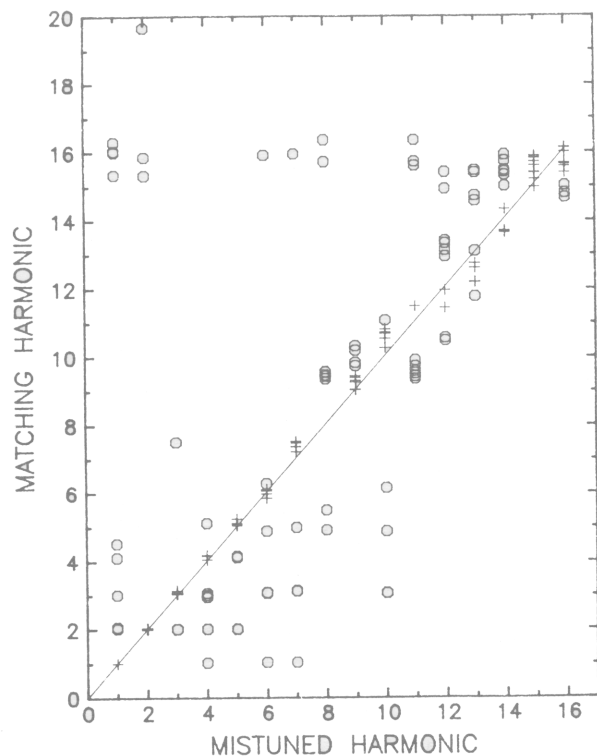


FIG. 3. Same as Fig. 2, except that the mistuning is 0.5%.

reduced. There are interesting clusterings of misses: The listener often matched the 2nd harmonic when it was actually the 1st or the 3rd or the 4th that was mistuned. The matching plot shows that mistuning the 8th or the 11th harmonic produced a dense cluster of misses in the vicinity of harmonic 9.5. There is another cluster of misses near the top of the spectrum between harmonics 15 and 16. It is clear that there are effects in the data that are not preserved by simply counting the number of hits and misses for each mistuned harmonic. However, this simple enumeration does serve to make further quantitative analysis straightforward.

II. HIT RATES AND THRESHOLDS

From plots of hits and misses, as shown above, it was easy to compute the percentage of hits to make plots called "hit-rate" plots. These show the percentage of hits as a function of the harmonic number of the mistuned harmonic. A different hit-rate plot is required for each different percentage of mistuning. Figure 4 shows a hit-rate plot with different mistunings (4%, 2%, 1%, and 0.5%) superimposed. The hit rates shown there have been averaged over the three listeners.

A "threshold" plot is made from the set of hit-rate plots. It shows the percentage mistuning that is required to reach some arbitrarily defined hit rate. It too is plotted against mistuned harmonic number. We found it convenient to take

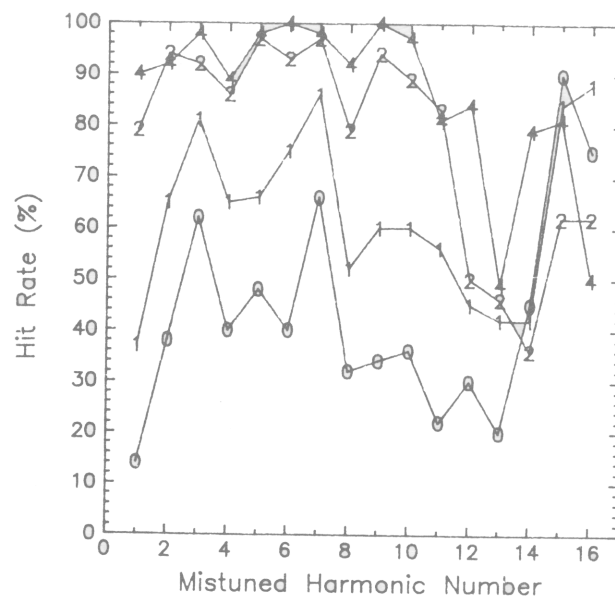


FIG. 4. Hit rates averaged over the three listeners for a nominal fundamental frequency of 200 Hz, a level of 60 dB per component and a tone duration of 410 ms. Equalized and unequalized stimuli have been averaged because no important difference between these conditions could be found. The parameter is the percentage of mistuning, 4%, 2%, or 1%, as labeled, and the symbol 0 indicates 0.5% mistuning.

a hit rate of 60% as the threshold. Normally, it was necessary to interpolate between percentages of mistuning for the 60% hit-rate point to determine a threshold mistuning.

A. Effect of mistuned harmonic number

Threshold plots for individual listeners, B, K, and S, were obtained for a 200-Hz fundamental frequency and a level of 60 dB per component. This was done by interpolating among hit-rate plots for 4%, 2%, 1%, and 1/2% mistuning. The resulting threshold plots showed considerable individual differences. Listeners B and K had higher thresholds for harmonics 4 and 8 and lower thresholds for odd numbered harmonics. This behavior might indicate some special perceptual character for octave relationships. However, listener S had low thresholds for mistuned harmonics 4 and 8 and higher thresholds for some odd harmonics such as 5 and 9.

We believe that the individual differences observed in our experiments are not due to random error but represent genuine idiosyncratic perceptual effects. Listener B, for instance, had difficulty detecting a mistuned 4th harmonic whatever the fundamental frequency or signal level. That there are individual differences of this kind is consistent with the results of Moore *et al.* (1985, 1986). However, one cannot discuss them in terms of a general model. Therefore, to make Fig. 5, we averaged the three individual threshold graphs, realizing that some potentially interesting effects would be averaged away.

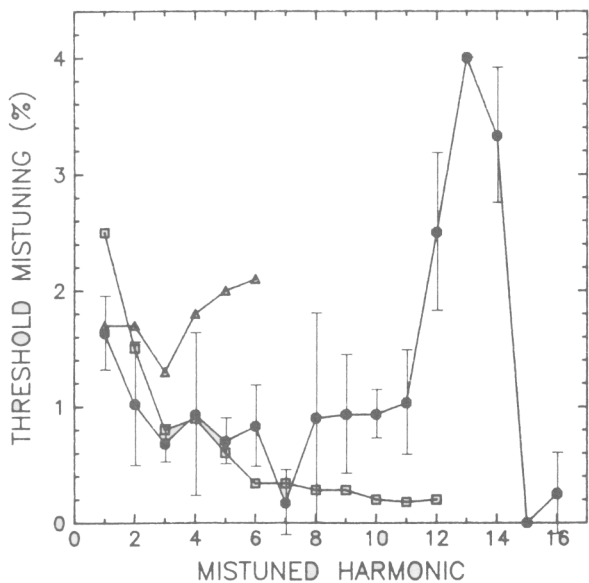


FIG. 5. The closed circles show the threshold for detecting a mistuned harmonic, averaged over three listeners for 200-Hz fundamental, a level of 60 dB per component and a tone duration of 410 ms. The error bars are 2 s.d. in overall length. The open symbols show data from the studies of Moore *et al.*, squares for 1985 and triangles for 1986.

The closed circles in Fig. 5 give the average threshold; the error bars have a total length equal to twice the unbiased standard deviation ($N - 1 = 2$ weight). Error bars are large because of the individual differences.¹

Also plotted in Fig. 5 are the data from the studies by Moore *et al.*, 1985 (squares) and 1986 (triangles). Because our criterion for threshold (60% hit rate) was rather arbitrary, there is little point in comparing the absolute values among different studies, except to note that they are approximately the same. One can, however, compare trends, and this is done below, in order of increasing mistuned harmonic number.

There is general agreement that thresholds decrease as the mistuned harmonic takes on values $n = 1, 2$, and 3, with the threshold for the mistuned fundamental ($n = 1$) particularly high. Subjectively, the case of the mistuned fundamental is unique. It is often detected not so much because it stands out as a separate entity but because it sounds out of tune. We are inclined to regard the higher threshold seen for the mistuned fundamental as reflecting a qualitative difference, not merely a quantitative difference.

As the mistuned harmonic number increases from 3 to higher numbers, the threshold is approximately constant until the dip at the mistuned 7th harmonic. In part, this dip represents a flaw in our experimental technique: Sometimes a listener's hit rate was appreciably greater than the threshold value of 60% even when the smallest mistuning (0.5%) was used. In such cases (there were 6 of them out of 48 threshold values), the threshold for that listener was taken

to be zero. Usually, this had little effect on the average threshold, but, for the mistuned 7th, there were two listeners for whom this occurred, and the threshold for the third listener was also rather low. Therefore, the threshold for the mistuned 7th is not correctly evaluated by our procedure. We can only say that the threshold is unusually low.

Except for the dip at 7, the threshold appears to be essentially constant as the mistuned harmonic number increases from 3 to 11. This behavior disagrees with the decreasing trend seen in the 1985 data of Moore *et al.* It is hard to make the comparison with their 1986 data because these extend only to the sixth harmonic, but in any case, the 1986 data do not show a decreasing trend. Together, the results are consistent with the idea that the 1986 study and the present study correctly represent the threshold for segregating a mistuned harmonic, and that the 1985 data show thresholds for detection of inharmonicity based upon beats and roughness.

For mistuned harmonic numbers 12, 13, and 14, the threshold becomes quite high. The threshold value for the mistuned 13th is only approximately correct because hit rates for two of the three listeners were somewhat less than 60% for the largest mistuning used (4%). The threshold estimate of 4% given in the figure is, therefore, somewhat lower than the true value.

For mistuned harmonics 15 and 16 the threshold is again small. One naturally wonders about the origin of this nonmonotonic behavior. Because harmonics 15 and 16 were at the top of the spectrum, it seemed possible that the low threshold was an artifact of our experimental stimuli. Conceivably, it was an edge pitch (Klein and Hartmann, 1981; Moore and Glasberg, 1989). As a test of this idea, one of us (listener B) did an experiment where the complex tone had 24 harmonics of which the lowest 20 could be mistuned by 2%. The data again showed good performance for mistuned low harmonics, poor performance for the mistuned 13th harmonic, followed by good performance for the mistuned 15th. Performance for 14 and 16 was intermediate. This result suggests that we should accept the low threshold for harmonics 15 and 16 as genuine, and not an artifact.

These peculiar nonmonotonic data suggest that two modes of perception may be involved in the task, their relative importance changing as the harmonic number changes. There are two other pieces of evidence that tend to support this idea. First, introspective listening (with the listener writing notes while unaware of which harmonic is mistuned) agrees: for mistuned harmonic numbers less than 12 the mistuned harmonic is segregated as a separate tone. For mistuned harmonic numbers greater than 13 there is a beating sensation, but the beating takes place in a region of tone height that is narrow enough that one can often make an acceptable pitch match.²

The other piece of evidence for a second mode of perception is the fact that, for harmonics 15 and 16, the expected dependence on percentage of mistuning breaks down. It seems that it is no easier to detect a mistuning of 4% than it is to detect a mistuning of 1% or 0.5%; in fact, Fig. 4 suggests that the reverse is true. This unusual dependence argues for a different mode of detection for the higher harmonics. It is

easy to apply this idea to the 16th harmonic because its only neighbor is the 15th. At these frequencies the critical bandwidth is about 500 Hz. Using the rule (Plomp and Steenen, 1968) that roughness for two sine tones decreases as their spacing increases beyond 1/4 critical bandwidth (125 Hz), one expects that 0.5% mistuning (3000 and 3216 Hz) might sound somewhat rougher than 4% mistuning (3000 and 3328 Hz).

Our conclusion is that our experiment can successfully measure the threshold mistuning required for a listener to hear out a mistuned harmonic in a 200-Hz complex tone and that this threshold is roughly a constant percentage of the mistuned harmonic frequency as the mistuned harmonic number increases from 2–11. Beyond mistuned harmonic number 11, the threshold for segregation increases rapidly. We interpret the low thresholds actually observed for harmonic numbers 15 and 16 as resulting from local roughness and not indicative of segregation of a single component only.

B. Effect of fundamental frequency

The effect of the fundamental frequency on the ability of listeners to hear out a mistuned harmonic was studied by comparing hit rates for given fractional mistunings, without reducing the results to threshold plots. The three listeners each did ten runs with 4% mistuning for each of three fundamental frequencies, 200, 400, and 800 Hz. The average over listeners is shown in Fig. 6, with fundamental frequency as a

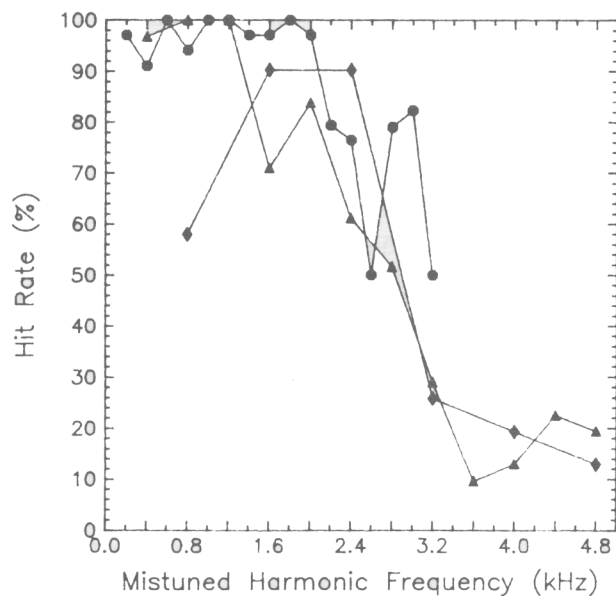


FIG. 6. Hit rate for detecting a mistuned harmonic, mistuned by 4%, averaged over three listeners. The level was 60 dB per component and the tone duration was 410 ms. The fundamental frequency is the parameter: circles for 200 Hz, triangles for 400 Hz, and diamonds for 800 Hz. We suggest that the circles at 2.8, 3.0, and 3.2 kHz should be discounted in comparing the plots, because matching to these harmonics of 200 Hz is presumably mediated by a process that is different from the process involved for lower harmonics of 200 Hz and for all harmonics of 400 and 800 Hz.

parameter. The horizontal axis gives the frequency of the mistuned harmonic, not the harmonic number. Therefore, data points for a fundamental of 400 Hz (triangles) are spaced twice as widely as points for 200 Hz (circles).

The closed circles, indicating 200 Hz, show the peculiar nonmonotonic behavior near the 15th harmonic as expected. Hit rates for the other fundamental frequencies seem to decrease more regularly as the mistuned harmonic number increases. The closed diamonds for 800 Hz show the considerable difficulty of this condition; only mistuned second and third harmonics could be reliably detected, even at 4% mistuning. Figure 7 shows similar results for 200- and 400-Hz fundamentals and 2% mistuning. The plots in Fig. 8 for 1% mistuning have the same trend as those in Fig. 7, but all hit rates are reduced.

Figures 6–8 invite comparison with threshold graphs by Moore *et al.* (1986) for fundamental frequencies of 100, 200, and 400 Hz, where the horizontal axis was chosen to represent mistuned harmonic number. If one imagines that the segregation of mistuned harmonics depends mainly upon auditory filtering with a bank of constant- Q filters, then that way of plotting the data makes good sense. Such a plot should produce a universal curve in that the threshold (or hit rate) should be independent of fundamental frequency, i.e., independent of the frequency spacing of the harmonics. On the other hand, if one imagines that the segregation of mistuned harmonics depends mainly upon neural synchrony and is limited as neural synchrony disappears for increasing frequency, then the kind of plot that we have chosen, with frequency as horizontal axis, is more likely to lead to a universal curve. The data in Figs. 6–8 support the latter interpretation. They suggest that synchrony vanishes rather abruptly as frequency increases from 2 to 3.5 kHz. This

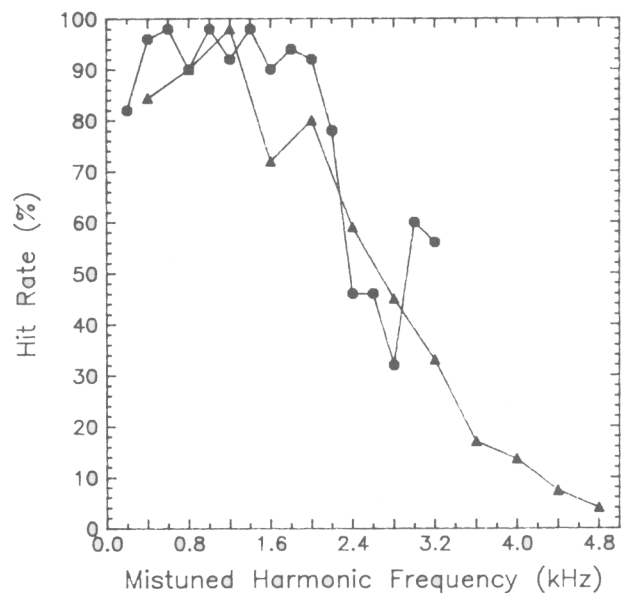


FIG. 7. Same as Fig. 6 except that the mistuning was 2% and there are only two fundamental frequencies, 200 and 400 Hz. Circles at 3.0 and 3.2 kHz should presumably be ignored in the comparison with the 400-Hz data.

frequency limit for useful synchrony is lower than the upper limit of 5 kHz based upon physiological data in cat (Johnson, 1974). Our limiting value does seem consistent with Plomp's (1967) psychophysical limit based upon the beating of mistuned consonances. For instance, one does not hear beats between sine tones at 2000 and 3000 Hz, even though beats at a 2 to 3 ratio are easily detected in lower frequency ranges. Unfortunately, the data of Moore *et al.* (1986) do not contribute on this point because they go no higher than 2.4 kHz.

C. Effect of signal level

The overall level of the signal can be expected to affect the spread of the neural excitation pattern along tonotopic coordinates. If the ability to segregate a mistuned harmonic depends critically upon the spatial distribution of excitation for the mistuned harmonic relative to neighboring harmonics, one might see a level dependence in the hit rate. To test this idea, the listeners did runs at a level of 30 dB per component to be compared with the standard runs at a level of 60 dB per component. Comparisons were made for a fundamental frequency of 200 Hz and signal durations of 410 and 50 ms, and for a fundamental of 400 Hz at a duration of 410 ms.

The results of the experiment for 200 Hz and 410 ms are shown in Fig. 9. This figure includes all the available data, from blocks of runs with 1% and 2% mistuning and for equalized and unequalized spectra (See the Appendix), a total of eight blocks in all. The scatter plot in Fig. 9 shows a difference of hit rates, specifically the hit rate at the 60-dB level minus the hit rate at the 30-dB level. In those cases where hit rates for both levels were 100% or 0%, no data point was plotted.

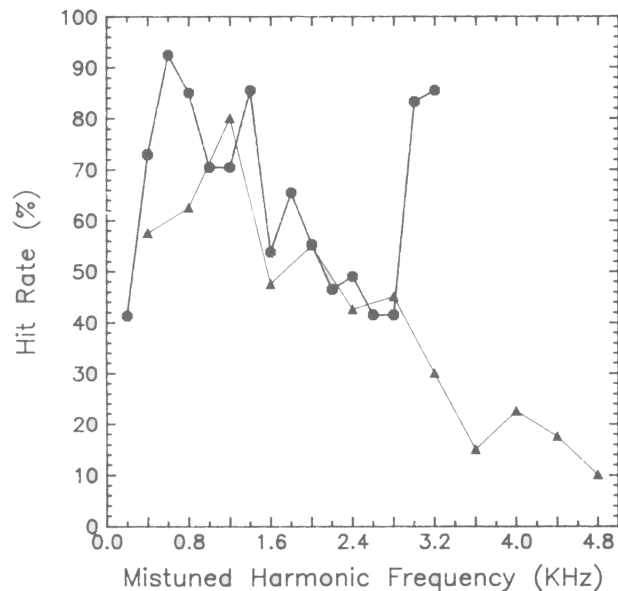


FIG. 8. Same as Fig. 7 except that the mistuning was 1%. The figure shows that hit rates are similar for 200 and 400 Hz. Circles at 3.0 and 3.2 kHz should presumably be ignored in the comparison with the 400-Hz data.

The figure shows that it was more difficult to hear out the mistuned fundamental at the low level, a fact that finds a trivial explanation in terms of equal-loudness contours. At 30 dB a 200-Hz tone can barely be heard.

The figure also shows that there were a few mistuned harmonics that could be heard out better at the lower signal level than at the higher signal level. These were the 4th and 5th. For example, hit rates for 1% mistuning of the 4th harmonic averaged 60% at 60 dB per component but 86% at 30 dB per component (410-ms tone), a difference of -26% .³ This peculiar level effect for the mistuned 4th harmonic was magnified for 50-ms duration (no figure shown), where the difference averaged -70% . A similar level effect for mistuned 4th harmonics in brief tones has also appeared with great strength in discrimination experiments (Hartmann, 1986).

By contrast, the detection of mistuned harmonics in the vicinity of the 7th harmonic was easier for the higher level tone. For 410-ms tones and 1% mistuning the difference in average hit rates for a mistuned 7th was about 20% favoring the 60-dB level over 30 dB. A similar plot made for a fundamental frequency of 400 Hz and duration of 410 ms showed again that hit rates for the 5th harmonic were higher at 30 dB than at 60 dB, but that hit rates for higher mistuned harmonic number were higher at 60 dB.

These results can be compared with the predictions of any model in which detection is based upon overlap of excitation patterns created by neighboring harmonics. If excitation patterns expand with increasing level (Whitfield, 1967), then the comparison leads to a paradoxical result. Consider the mistuned 7th harmonic: We find experimentally that it is more easily detected at 60 dB than at 30 dB.

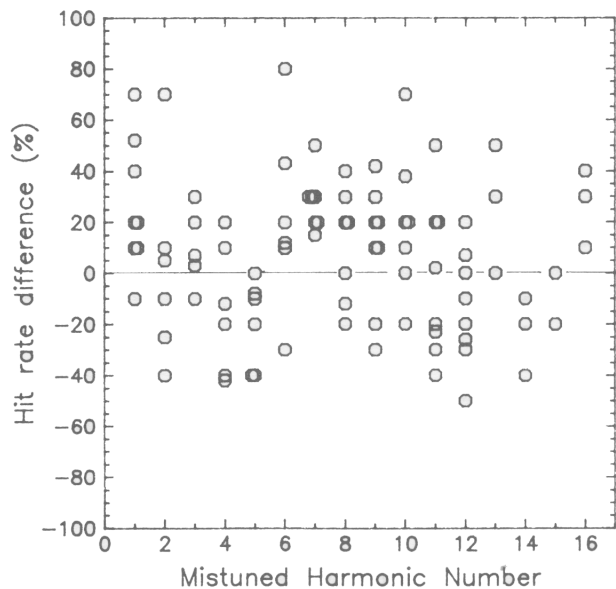


FIG. 9. Each circle represents the difference between the hit rate in runs where the level was 60 dB/component and corresponding runs where the level was 30 dB/component. When the difference is positive, performance was better at the 60-dB level. Some circles have been shifted laterally so that no circles overlap entirely.

Presumably, this means that detection of the mistuned 7th is aided by increasing the width of the excitation patterns; at 30 dB, the amount of overlap is suboptimal. If this argument is accepted, then the prediction for the mistuned 4th harmonic is clear. Near the 4th harmonic the excitation patterns are even more widely separated (on a physiological scale) than in the vicinity of the 7th. Therefore, one expects that again detection performance will be better for increased level. Experimentally, however, the reverse behavior is observed. Actually, to explain our data would require that the relevant aspect of the excitation pattern should broaden along the tonotopic coordinate as the level *decreases*. This paradoxical result is an example of the considerable difficulty involved in developing a comprehensive model for the segregation of mistuned harmonics.

D. Effect of duration

The dependence of performance on tone duration can give clues to the process involved in segregating a mistuned harmonic. If, for instance, the process were identical to frequency discrimination, then one would not expect a significant duration dependence, so long as durations are longer than about 50 ms. This is so because frequency difference limens at intermediate frequencies do not change much as tone duration is increased beyond 50 ms.

1. Hit rates and thresholds

The two experiments by Moore *et al.* (1985) used four durations: 50, 110, 410, and 1610 ms. The different durations led to different functions when the threshold mistuning was plotted against the mistuned harmonic number, n . When the cues were unrestricted, the thresholds for 410 and 1610 ms decreased with increasing n . This behavior is expected if the detection of inharmonicity is mediated by beats or roughness and if the key element in detection is the rate of this kind of modulation. A constant rate, and hence constant sensation according to this explanation, corresponds to a decreasing percentage threshold (decreasing as $1/n$) as n increases. If this explanation is correct then one would expect that, for tone durations as short as 50 ms, where it is hard to detect beats or roughness, the thresholds should not decrease as a function of n . The data from the 1985 study agree on this point. For 50-ms tones there was no systematic tendency for thresholds to decrease as the mistuned harmonic number increased.

One further expects that, if the listeners in the 1986 study by Moore *et al.* were successful in ignoring beats or roughness, as they were instructed to do, then the tendency for thresholds to decrease with increasing n should disappear, even for durations of 410 and 1610 ms. This too was borne out by the data, as shown in Fig. 2 of that article.

Our own experiments on duration dependence used a fundamental frequency of 200 Hz, and levels of 60 dB per component and 30 dB per component. There were two durations, 50 and 410 ms. These conditions are a subset of the conditions used by Moore *et al.*, permitting a direct comparison. Listeners B and S did 10 runs with mistunings of 1%, 2%, and 4%.

Our data showed that the hit rates for 50-ms tones were less than those for 410-ms tones. This tended to be true for all mistuned harmonic numbers; it was especially true for the low-frequency harmonics 1 and 2. Such a reduction is of particular interest if one supposes that the detection of a mistuned harmonic depends upon time-dependent features of the waveform. The elemental temporal feature in the physical description of the stimulus is the dephasing of a mistuned harmonic with respect to other harmonics. A tone of brief duration limits the amount of dephasing that can occur, and this limitation is greatest if the mistuned harmonic has a low frequency. For example, if the second harmonic of 200 Hz is mistuned by 2%, then that harmonic is dephased at an 8-Hz rate, for a total of 144 deg after 50 ms. Hit rates for the first several harmonics were low enough to permit an estimate of a dephasing threshold. The result was 171 deg with a standard deviation of 54 deg, with no particular dependence on level. Apparently, the auditory system can work with rather small values of the dephasing, approximately equal to half a cycle.

Although a dephasing threshold of half a cycle has an intuitive appeal, it is not clear that the concept of a dephasing threshold is very useful. The data of Moore *et al.* (1986) for 410 and 50 ms do not show a particularly large effect at low harmonic numbers. For all the listeners, the thresholds for these two durations are roughly parallel as the harmonic number is increased. To explain this result, one might invoke two limitations imposed by short durations, a dephasing limitation for low-frequency harmonics and a spectral resolution limitation for higher harmonics.

2. The perceived delay of the mistuned harmonic

A peculiar effect occurs when a mistuned harmonic experiment is run at short durations such as 50 ms. Listeners hear the mistuned harmonic segregated from the complex tone, but the mistuned harmonic emerges from the complex tone only after a delay. The effect is striking. It was noticed by the three listeners in this study. It was noticed by all the listeners in Hartmann's 1985 study. Listeners cannot agree whether the mistuned harmonic persists after the end of the complex tone or whether it ends together with the complex tone. They are all agreed that the mistuned harmonic seems to start later than the complex tone.

The perceived delay is consistent with a two-part idea about auditory organization: The first part is that segregation of a tone from a complex requires information and that integration is the default operation in the absence of information. The second part is that information about mistuning is acquired over time, possibly as the mistuned harmonic is dephased with respect to neighboring harmonics. A corresponding two-part idea in the context of stream segregation was proposed by Bregman (1978).

E. Segregation versus resolution

The ability of listeners to hear out a harmonic in a periodic complex tone was studied in experiments by Plomp (1976 pp. 2ff; see also Plomp, 1964 and Plomp and Mimpen, 1968). The experimental technique was straightforward:

The listener heard a complex tone with a dozen harmonics. Then, the listener heard two sine tones; one sine tone had a frequency equal to one of the harmonics, the other sine tone had a frequency that was halfway between two harmonics. The listener's task was to choose which sine tone was present as a harmonic in the complex tone. The data showed that listeners could easily hear out harmonics with low harmonic numbers but experienced increasing difficulty as the harmonic number of the sine probe increased. The harmonic number for which listeners made 75% correct responses was taken to be the threshold for resolution. Such an experiment measures the resolving power of the auditory system, treated as a spectrum analyzer. The fact that the resolving power closely parallels the critical band width, measured by other methods, suggests that the resolving power is essentially a measure of auditory filter bandwidth.

Resolving power is quite different conceptually from the power to segregate a mistuned harmonic. The data are different too. Perhaps the most dramatic illustration of the difference is that Plomp found that the results of his resolving-power experiments were unchanged when the harmonic complex tone was replaced by a complex having inharmonic partials. By contrast, our experiments on segregation depend upon inharmonicity as the major cue. It would be impossible for listeners to do our experiment in the case of inharmonic complex tones. Unlike resolution experiments, it is not at all evident that narrow filter bandwidths are a benefit to the listener in a segregation experiment.

Curiously, it is possible to gain information about auditory frequency resolution from our mistuned harmonic matching experiments. But the information does not come from the hit rate; the information comes from the misses. Figure 3 clearly shows that, when a listener makes incorrect matches near a low harmonic ($n = 1-5$) the matching frequencies are not randomly scattered; instead, they tend to line up at correct harmonic values. This shows that listeners are hearing out the harmonic that they are matching, even if it is the wrong (i.e., not mistuned) harmonic. This idea can be made quantitative by measuring the width of the scatter of matches around harmonic values. If the width is small, then that is evidence that the listener heard out the harmonic as a resolved component.

To do the scatter analysis, we first assigned each miss to a bin, labeled by the harmonic number closest to the miss. For instance, matches at harmonic numbers 4.3 or 3.6 were assigned to the bin for the 4th harmonic. A separate analysis was done for each listener and each experimental condition, i.e., for each block of ten matches. If there were less than four matches in a bin, then that bin was dropped. If there were four or more matches in a bin, then the unbiased standard deviation of the matches was computed, about the (possibly nonintegral) mean.

A convenient reference is the value of the standard deviation for random matching. For example, a random match in the 4th harmonic bin is equally likely to be anywhere between 3.5 and 4.5 harmonic numbers. The value of the standard deviation for random matches is thus equal to the square root of the integral of x^2 between $-1/2$ and $+1/2$, i.e., the square root of $1/12$, namely 0.289. The data showed

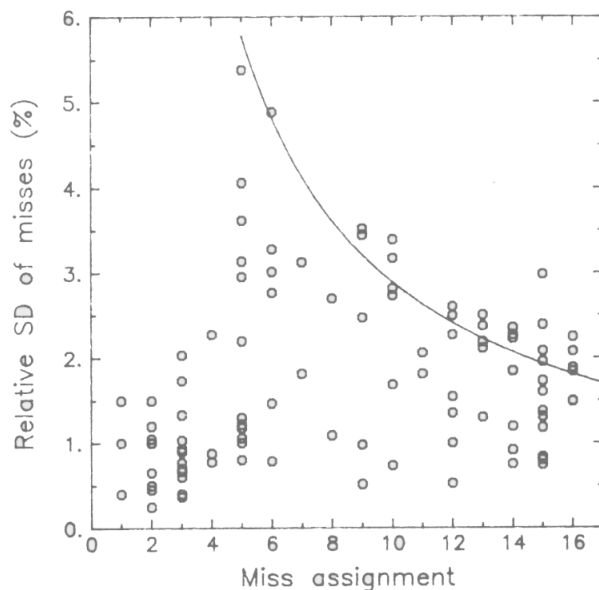


FIG. 10. Each circle represents an average standard deviation (s.d.) of misses assigned to the harmonic number (n) given on the horizontal axis. The s.d. has been divided by n and multiplied by 100 to put the plotted values on a percent basis. The solid line shows the expected result for random matching. The experimental s.d. falls appreciably below the solid line when listeners can hear out a harmonic that is not mistuned (top of graph at 6% = 1 semitone).

that s.d.'s were much less than this value for matches to harmonics 1-4, but a few s.d.'s approached this value for harmonic 5. The data can be displayed more conveniently in relative terms, dividing all the s.d.'s by the assigned harmonic number. A scatter plot showing all the data is given in Fig. 10. For reference, the top of the graph is at 6%, equivalent to a musical semitone.⁴ Also shown is a curve that is the line $0.289/n$ for random matching.

The collection of data on the misses was, by nature, opportunistic and not systematic. Nonetheless, the data do indicate a dramatic change at about the 5th harmonic. This is in near agreement with the results of Plomp and Mimpen for hearing out the harmonics of a 200-Hz fundamental. They actually concluded that the 5th was the highest harmonic that could be resolved.

But although our data can be used, as above, to learn about frequency resolution, the principal purpose of this section of the paper is to emphasize the considerable difference between resolution and segregation. In particular, there is no contradiction when Plomp and Mimpen find that the highest resolvable harmonic of 200 Hz is the 5th and when we find that listeners can successfully segregate a 10th harmonic that has been mistuned by only 1%.

III. PITCH SHIFTS

As noted above, the matching data revealed pitch shifts. The pitch of a sine tone (the mistuned harmonic) embedded

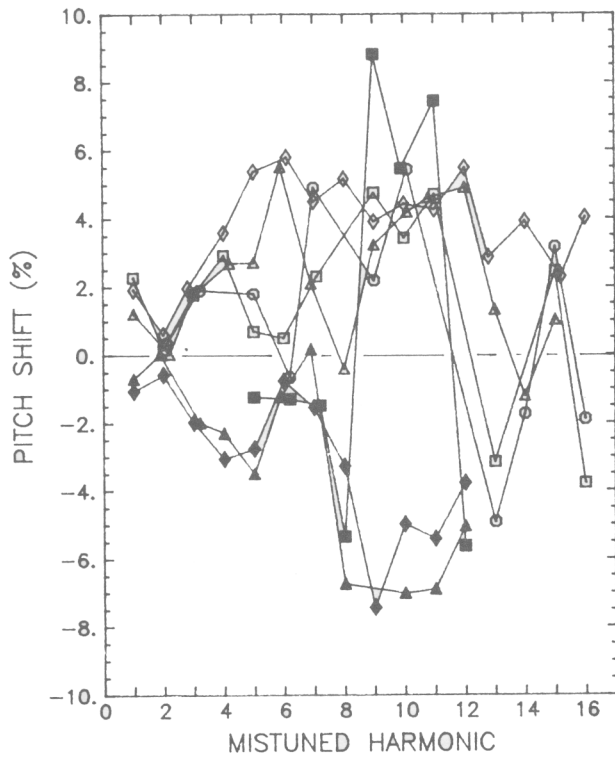


FIG. 11. Pitch shift of a mistuned harmonic as a percentage of the frequency of the mistuned harmonic, for listener B. Fundamental frequency of 200 Hz, level of 60 dB/component and duration of 410 ms. Open symbols are for positive mistunings; filled symbols are for negative mistunings. The amount of mistuning is a parameter: Circles for 0.5% mistuning, squares for 1%, triangles for 2%, diamonds for 4%. Some symbols have been slightly displaced horizontally for clarity.

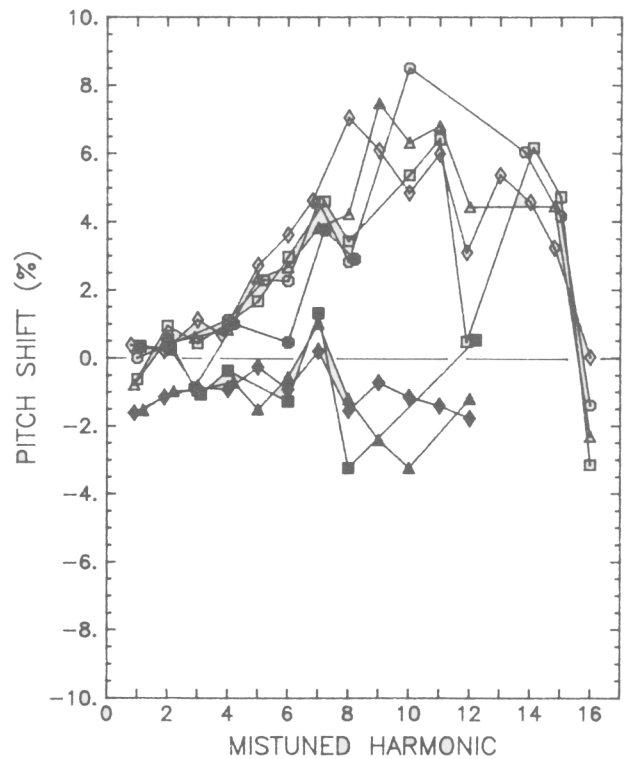


FIG. 12. Same as Fig. 11 but for listener S.

in a complex tone is different from the pitch in isolation. As a result, the frequency of the isolated matching tone that best matches the pitch of the mistuned harmonic is not equal to the frequency of the mistuned harmonic; it is shifted. There have been previous efforts to measure the pitches of the components of a complex tone (Terhardt, 1971; Houtsma, 1981; Peters *et al.*, 1983). The mistuned harmonic technique is one way to make harmonics stand out from the complex so that their pitches can be matched.

The data to be presented here were taken from those pitch matches that were found to be "hits." Therefore, for any mistuned harmonic, the standard deviation of the matches was within the 2.5% self-consistency window. We further required that at least four matches survive the self-consistency test to calculate a mean pitch shift. Because of these requirements, it was not possible to find a pitch shift for every mistuned harmonic number in every condition. The shift itself was calculated as a percentage deviation of the matching frequency from the actual frequency of the mistuned harmonic. Results for listeners B, S, and K are shown in Figs. 11–13, respectively, for the case of a 200-Hz fundamental, 60 dB per component and 410-ms duration. Open symbols in the figures show shifts for runs with positive mis-

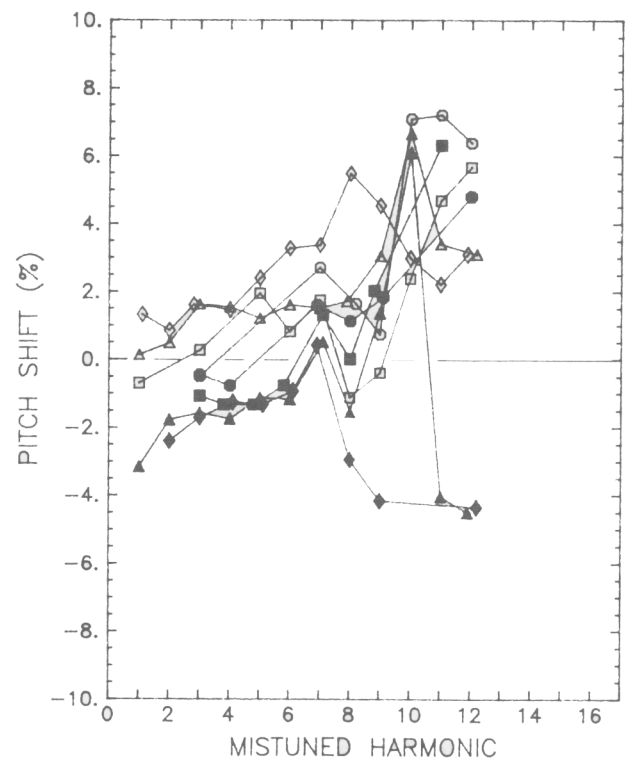


FIG. 13. Same as Fig. 11 but for listener K.

tunings; closed symbols are for runs with negative mistunings. On runs with negative mistuning, only the 12 lowest harmonics were candidates for mistuning.

The figures show that positive mistunings usually led to positive pitch shifts; this tendency is very strong. The figures also show that negative mistunings tended to produce negative pitch shifts. The effect for negative mistunings is not so strong nor consistent as the shift for positive mistunings, but it is clearly present in the data. In other terms, most of the data show a form of "overshoot" effect. When the mistuning is positive, the pitch is shifted to a still more positive value, and, when the mistuning is negative, the pitch becomes even more negative.

The pitch model of Terhardt (1979) and his colleagues (Terhardt *et al.*, 1982) includes an algorithm for the calculation of the spectral pitches of the harmonics. The algorithm has the flavor of a place theory and does not require the partials of a tone to be strictly harmonic. It is, therefore, applicable to our stimuli. According to the algorithm, the spectral pitches are shifted because of partial masking effects. For example, the spectral pitch of the fundamental of a complex tone is shifted downward because the excitation for the fundamental is masked from above by higher harmonics, mainly the second. The predictions of the algorithm, calculated from the formulas given in Terhardt *et al.* (1982), are shown in Fig. 14. The figure shows predictions for -4% and 4% mistuning; the predictions for other values of mistuning actually used in our experiments lie between the curves for -4% and 4% . The algorithm predicts a shift for zero mistuning (harmonic complex), approximately midway between curves for -4% and 4% , although Peters *et al.* (1983) did not find such shifts.

Comparing the predicted shifts with the observed shifts shows that the algorithm correctly predicts the trend of the data when the mistuning is positive. When the mistuning is negative, however, the algorithm fails completely. Experimentally, negative mistuning usually leads to negative pitch shifts. By contrast, the algorithm predicts positive pitch shifts for negative mistuning. Figure 14 shows that, for harmonics 2–7, the predicted pitch shifts for negative mistunings are even more positive than the predicted pitch shifts for positive mistunings. This is a result of the upward spread of masking. For the third harmonic, for example, the major contribution to the pitch shift is the second harmonic. When the third harmonic is mistuned negatively, it comes closer to the second harmonic, and so the pitch shift becomes larger (more positive).⁵ The fact that the algorithm fails so badly for negative mistuning suggests that there is something quite wrong with the idea that pitch shifts are mainly determined by partial masking.

The reliability of pitch shift data can be assessed by comparing the standard deviations with the pitch shifts themselves. The standard deviations were calculated for individual listeners across the matches of different runs. The separate matches that contributed to a standard deviation were those that satisfied the windowing criterion for accepting data, thus, a minimum of four matches and a maximum of ten. Our most complete data are for nominal fundamentals of 200 Hz. Here, the figures for B, S, and K have 102

closed symbols corresponding to negative mistuning. For 65% of those points, the pitch shift is larger than the standard deviation. This is evidence that the negative pitch shift is genuine.

Although the Terhardt algorithm for pitch appears to fail, it still is quite possible that the pitch shifts observed in our experiment result from effects that are tonotopically local. If this is so, then the pitch shifts should be sensitive to the separation between adjacent harmonics, or spectral density. To test this idea, we studied the pitch shifts obtained for nominal fundamental frequencies of 400 and 800 Hz. Only positive mistunings were used. The amount of significant data was limited; pitch shifts were normally greater than the standard deviations for harmonics 3, 5, 6, and 7 for 400 Hz but only for the third harmonic for 800 Hz. Because the spectral density is smaller for an 800-Hz fundamental than for 400 Hz, one expects smaller pitch shifts for the 800-Hz fundamental in any given frequency region. This expectation was borne out by the data: The shift for the third harmonic of 800 Hz was 1.8%, considerably less than the shift for the sixth harmonic of 400 Hz, which was 4.8%.

Similarly, one might expect the pitch shift to decrease with decreasing signal level because a smaller signal level should lead to less interaction between the mistuned harmonic and its neighbors. To test this idea, the three listeners did runs with a 200-Hz fundamental, a mistuning of 2%, and

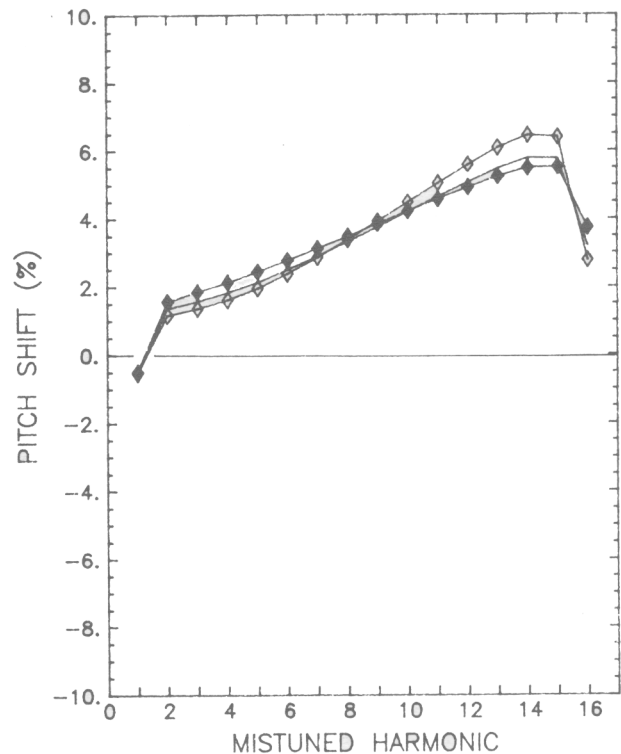


FIG. 14. Predicted pitch shift from the algorithm of Terhardt *et al.* (1982) for a fundamental frequency of 200 Hz and a level of 60 dB per component. The open diamonds are for 4% mistuning; the closed diamonds are for -4% mistuning, and the solid line is for no mistuning. If the algorithm is correct then these predicted shifts should agree with the data in Figs. 11–13.

a duration of 410 ms, using three different signal levels. Per component, these levels were 60, 30, and 12 dB SPL. The goal of the experiment was to compare pitch shifts for a given mistuned harmonic across the different levels. The matches that passed the acceptance tests permitted 36 comparisons of 60 dB with 30 dB, and 15 comparisons of 30 dB with 12 dB. Of these 51 comparisons there were only nine cases that were inconsistent with the rule that lower signal levels lead to smaller pitch shifts.

IV. CONCLUSIONS

The ability of a listener to segregate a mistuned harmonic from a complex tone background is an example of auditory organization based upon steady-state cues. We began our study of this effect with the intention of improving on the data of Moore and his colleagues, by using a matching experiment instead of a discrimination experiment. We found that the matching experiment has a number of advantages, but it is not perfect. First, we were required to discard as artificial successful matches for harmonics above the 12th. We attributed these to localized roughness or beats and not to the segregation of a mistuned harmonic. This problem occurred for a fundamental frequency of 200 Hz; it did not occur for fundamentals of 400 and 800 Hz. Second, using the matching technique did not eliminate the large intersubject variability that seems to be characteristic of mistuned harmonic detection experiments.

The matching technique enabled us to extend measurements to higher frequencies, and this revealed an important absolute frequency limit, between 2.2 and 3.5 kHz, for the segregation of a mistuned harmonic. We understand this result by supposing that segregation is mediated by the detection of asynchrony in neural fibers stimulated by both the mistuned harmonic and neighboring harmonics. The high-frequency barrier then reflects a limit on encoding by neural synchrony as frequency increases.

For frequencies below 2 kHz, segregation thresholds, expressed as a percent of mistuned harmonic number, appeared to be roughly independent of mistuned harmonic number, neglecting special effects for the fundamental and individual differences. To some extent this result reflects our data analysis method. Our criteria for a "hit" included a self-consistency window that was a constant percentage of the mistuned harmonic number. Had we chosen a window with a constant frequency width for all mistuned harmonics then segregation thresholds, expressed as a percent of mistuned harmonic number, would have increased for increasing mistuned harmonic number. However, this methodological arbitrariness is unlikely to have had an important effect on our conclusions because, when detection performance deteriorates at high frequency, it deteriorates dramatically. With either window criterion, the results show clearly that segregation thresholds do not decrease as the mistuned harmonic number increases. By contrast, earlier thresholds from Moore *et al.* (1985), where detection could be based upon any cue, did decrease with increasing harmonic number.

Further experiments found that segregation performance showed paradoxical behavior as a function of overall signal level. A mistuned 7th harmonic was more easily de-

tected at 60 than at 30 dB, but a mistuned 4th harmonic was more easily detected at 30 than at 60 dB. If the detection of a mistuned harmonic requires an overlap between excitation patterns associated with the mistuned harmonic and its neighbor(s), and if excitation pattern widths increase with increasing level, then the level effect, if any, should be opposite to the one that we observed. Experiments with tones of brief duration showed that the amount of waveform dephasing required to segregate a harmonic can be small.

An interesting virtue of the matching method is that the data used to determine segregation performance can also be used to study a quite different matter, namely, the pitches of the spectral components of a complex tone. Our experiments found that there are pitch shifts and that these tend to be exaggerated versions of the mistuning; i.e., positive mistuning leads to positive pitch shift and negative mistuning leads to negative pitch shift. The negative shifts are smaller and less stable than the positive shifts, but they are undeniably real. By contrast, a model whereby pitch shifts are caused by partial masking of excitation strength predicts that pitch shifts should be positive for both positive and negative mistunings.

ACKNOWLEDGMENTS

The work of WMH was supported by the National Institutes of Health and by the US-France Cooperative Science Program of the National Science Foundation and the Centre National de la Recherche Scientifique.

APPENDIX: EXPERIMENT DETAILS

This Appendix includes further details on the experimental procedures and the analysis.

1. Equalized stimuli

Most of the pitch matching experiments, which provide data for this article, were done twice. In the first set, the electrical signal sent to the phones had 16 harmonics with equal amplitudes. For the second set, the harmonic amplitudes of the electrical signal were adjusted so that the amplitudes of the acoustical signal, as measured with a spectrum analyzer and flat-plate coupler, were all equal. Each set comprised ten matches by a listener to each mistuned harmonic for each percentage of mistuning. Although some data points for individual listeners were different for the two stimulus sets, there was little agreement among the data for different listeners with respect to potential effects of equalization. Because there was no systematic effect of equalization, we combined the data for the two stimulus sets. Data from a third set of ten matches, with 16 unequalized harmonics of which only the first 12 might be mistuned, were also added to the set when it was discovered that they did not differ in any important way from the data of the other two sets. Increasing the amount of data in this way had the effect of smoothing the data for individual listeners, a desired result for our purposes.

2. Success by persistence

In a matching experiment of our kind, misleading results may be obtained if a listener persistently makes a particular match. If, for example, the listener always matches the third harmonic of the complex tone, then, on the trials where the third harmonic is actually mistuned, the listener will score 100% correct, even if he cannot identify the third harmonic as the correct answer. This effect is a potential source of error. We discovered only one case of such persistent matching, listener K at 200 Hz, 60 dB/component and 0.5% mistuning. Here, it was necessary to correct the hit rate for the mistuned third by subtracting the rather large percentage of matches to the mistuned third harmonic when harmonics 1,2,4, and 5 were mistuned.

¹ To test our analysis method, we reversed the order of operations: We first found the hit-rate plots for each percentage of mistuning, averaged over the three listeners, given as Fig. 4, and then interpolated for the 60% hit rate to find a single threshold plot. That plot was compared with the plot shown in Fig. 5. For each of the 16 mistuned harmonics, the difference was never greater than 0.3 percentage points, except for the mistuned 7th, where the difference was 0.4 points.

² That listeners can identify frequency ranges where beating takes place in a broadband context was recently shown by Marin *et al.* (1989).

³ The improvement in detecting some mistuned harmonics as the signal level is decreased did not continue indefinitely. Runs done at a level of 15 dB/component showed hit rates that were always less than the rates for 60 dB/component.

⁴ Data for the s.d. of the misses come from ten-run blocks as follows: five unequalized at 0.5% mistuning; five unequalized and two equalized at 1%; five unequalized and three equalized at 2%; five unequalized and one equalized at 4%.

⁵ There is no easy way to explain why the predicted pitch shift becomes larger for positive mistuning than for negative when the mistuned harmonic number is greater than 8. Terhardt's algorithm is semiempirical and owes at least as much to experimental data as to model building.

Bregman, A. S. (1978). "Auditory streaming is cumulative." *J. Exp. Psych.: Human Percept. Perform.* **4**, 380–387.

Bregman, A. S., and Pinker, S. (1978). "Auditory streaming and the building of timbre." *Can. J. Psych.* **32**, 19–31.

Gardner, R. B., and Darwin, C. J. (1986). "Grouping of vowel harmonics by frequency modulation: Absence of effects on phonemic categorization." *Percept. Psychophys.* **40**, 183–187.

Hartmann, W. M. (1985). "Perceptual entities from complex inharmonic tones," *Assn. for Research in Otolaryngology 8th Meeting*, pp. 168 (abs).

Hartmann, W. M. (1986). "Pitch and the perceptual organization of complex spectra," *J. Acoust. Soc. Am. Suppl. 1* **79**, S65.

Hartmann, W. M. (1988). "Pitch perception and the segregation and inte-

gration of auditory entities," in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 623–645.

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," Ph.D. thesis, MIT, Cambridge, MA.

Klein, M. A., and Hartmann W. M. (1981). "Binaural edge pitch," *J. Acoust. Soc. Am.* **70**, 51–61.

Marin, C. M. H., McAdams, S., and Lienard, J.-S. (1989). "The role of beats in concurrent sound segregation based on frequency modulation cues," in *Proceedings of the 5th Annual meeting of the International Society for Psychophysics*, Cassis, France, pp. 120–125.

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

McAdams, S. (1984b). "The auditory image: A metaphor for musical and psychological research on auditory organization," in *Cognitive Processes in the Perception of Art*, edited by R. Crozier and A. Chapman (North Holland, Amsterdam), pp. 183–187.

McAdams, S. (1989). "Segregation of concurrent sounds. I: Effects of frequency modulation coherence," *J. Acoust. Soc. Am.* **86**, 2148–2159.

Moore, B. C. J. (1982). *An Introduction to the Psychology of Hearing* (Academic, London), 2nd ed.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1985). "Thresholds for the detection of inharmonicity in complex tones," *J. Acoust. Soc. Am.* **77**, 1861–1867.

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.

Moore, B. C. J., and Glasberg, B. R. (1989). "Difference limens for phase in normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **86**, 1351–1365.

Moorer, J. A., Chauveau, A., Abbott, C., Easty, P., and Lawson, J. (1979). "The 4C machine," *Comput. Music J.* **3**, 16–24.

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. (1967). "Beats of mistuned consonances," *J. Acoust. Soc. Am.* **42**, 462–474.

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

Plomp, R., and Steeneken, H. J. M. (1968). "Interference between two simple tones," *J. Acoust. Soc. Am.* **43**, 883–884.

Plomp, R. (1976). *Aspects of Tone Sensation—A Psychophysical Study* (Academic, London).

Rasch, R. A. (1978). "The perception of simultaneous notes as in polyphonic music," *Acustica* **40**, 21–33.

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

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

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

Whitfield, I. C. (1967). *The Auditory Pathway* (Arnold, London).