On the Duifhuis pitch effect

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

(Received 12 July 1995; revised 11 July 1996; accepted 14 October 1996)

An effect discovered by Duifhuis [J. Acoust. Soc. Am. **48**, 888–893 (1970)], wherein an omitted high harmonic of a periodic complex tone is found to have an audible pitch, is extended to a variety of new broadband signal conditions. The effect is found to exist for flat spectra and spectra decreasing at 6 dB/octave, independent of phases as long as they are constant. The effect exists for alternating phases and Schroeder phases. It can generate a missing-fundamental pitch. Pitch and loudness matching experiments support the status of the omitted harmonic as an objective tone in the signal. Further experiments using narrower bands challenge the traditional explanation for the effect, which attributes it to short-term frequency analysis by peripheral auditory filters. Instead, the experiments suggest that different peripheral channels must be combined, maintaining some phase information, to generate the effect.
[S0001-4966(97)05202-8]

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

INTRODUCTION

Duifhuis (1970, 1971) described a delightful psychoacoustical effect that might be called the ''pitch of the harmonic that is not there.'' It was created by a periodic train of narrow pulses to which was added a sine tone, phase locked, and adjusted so as to cancel one of the harmonics of the pulse train. The spectrum of the signal therefore looked like Fig. 1(a). When the fundamental frequency was low enough and the cancelled harmonic number was high enough, listeners perceived a pitch corresponding to the cancelled harmonic.

A priori, Duifhuis's effect is surprising. What is heard by listeners is exactly what is *not* present in the signal. On second thought, the effect seems easy to understand. An oscillographic tracing of the waveform, shown in Fig. 1(b) shows a clear periodic component corresponding to the cancelled harmonic. The trace looks like a pulse with added sine. The oscilloscope is, of course, a very broadband system. The auditory system, however, is not broadband in this way. Instead, the auditory system begins with tuned channels, represented as an array of auditory filters. The filters analyze incoming signals into separate bands, and therefore the entire oscillographic representation never appears within the auditory system.

The explanation for the audible effect given by Duifhuis is that when the frequency spacing between harmonics is considerably less than the bandwidth of relevant auditory filters, the tuned auditory channels appear to be broadband too. Then the output of the auditory filter bank retains an impulsive character, and the cancelled harmonic appears in the temporal gap between ringing responses to successive pulses. The cancelled harmonic is present in the output just as though the cancellation tone were an objective sine signal.

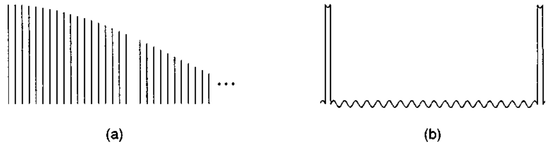

FIG. 1. (a) The amplitude spectrum of a periodic narrow pulse train with the 19th harmonic omitted; (b) waveform generated by the spectrum in (a) with all harmonics having cosine phase.

The present article is a study of the Duifhuis pitch effect (DPE) consisting of two parts. The first uses broadband signals to explore the DPE for a variety of new conditions. The conclusions of these broadband tests are consistent with the idea that the DPE results from an objective added cancellation tone. The second part is a band-narrowing experiment that tests the concept of the cancellation tone as an isolated component appearing in the output of the auditory filter bank during the temporal gap. It is shown that this concept fails. Instead, the DPE appears to be a segregation process that depends both on the cancellation tone in the gap and on the broadband complex tone that establishes a background including multiband phase coherence.

I. BROADBAND DPE EXPERIMENTS

The broadband experiments were undertaken to extend the experiments by Duifhuis to new signal conditions. The first goal was to distinguish between those conditions in which the DPE can be heard and those condition in which it cannot. The second goal was to determine the pitch and loudness of the DPEs.

A. Method

It is well known that changes in the spectrum (turning a harmonic on and off, amplitude modulating a single harmonic, mistuning a harmonic or suddenly shifting its phase) can cause a harmonic to be heard out from a complex tone (Pierce, 1960; Kubovy and Jordan, 1979; McAdams, 1984; Hartmann, 1988; Moore and Glasberg, 1989; Alcantara and Moore, 1995). Such effects occur both for harmonics that are spectrally resolved and for harmonics that are not. They are ''attention'' effects caused by the contrast between ''before'' and ''after'' conditions.

The Duifhuis pitch effect is not such a contrast effect, and to guard against a contrast artifact our formal experiments did not permit the listener to compare omitted-harmonic signals with complete-spectrum signals. Instead they were matching experiments. The listener heard an omitted-harmonic tone and was required to match its pitch with a sine tone. In the end it was concluded that the listener could hear the DPE if the matches were consistent.

1. General procedure

The listener was seated in a sound-treated room, holding a response box that controlled the events of an experimental trial. When the listener pressed a yellow button there was a pause of 200 ms and then a 700-ms complex tone having an omitted harmonic. When the listener pressed an orange button there was a pause of 200 ms and then a 700-ms sine tone with a frequency that could be adjusted by means of a ten-turn potentiometer on the box. The listener's task was to tune the sine tone so as to match the pitch of the DPE. The level of the matching tone could also be easily controlled from the response box. The listener could call up the complex tone or the matching tone as often as he liked. When the listener was satisfied with his match he pressed the green button to finish the trial. The target and matching frequencies were recorded, and then the next trial with a different omitted harmonic began. There was no feedback to the listener.

Trials were blocked into runs. Each experimental run included four trials. It took about 1–5 min for a listener to finish a run. Runs for each of the conditions below continued until the results appeared to be stable. The final data reported here are based on the final five matches.

2. Stimuli

Duifhuis created an omitted-harmonic tone by adding a cancellation sine tone to a pulse train. Therefore, the amplitude spectrum decreased slowly with increasing frequency as shown in Fig. 1(a). Our experiments were similar except that they used a flat amplitude spectrum. The computer added a large number, N, of harmonics with equal amplitudes and cosine phases to make signal $x(t)$,

$$x(t)=\sum_{n=1}^{N} A_n \cos(2\pi n f_0 t+\phi_n), \tag{1}$$

where all the amplitudes A_n, except for the highest two, were equal to 1. The amplitudes of the highest two harmonics in the spectrum were tapered to reduce edge pitches (Klein and Hartmann, 1981) and Gibbs phenomena ($A_N=0.3$ and $A_{N-1}=0.7$). All the phases, ϕ_n, were equal to zero, and the fundamental frequency, f_0, was 50 Hz. To generate the DPE, one harmonic, called n', was omitted by setting the corresponding amplitude $A_{n'}$ equal to zero. Although our technique of signal generation did not explicitly add a cancellation tone to a complex tone, we will continue to refer to the omitted harmonic as the result of an added ''cancellation tone.''

The digital waveform was loaded into a digital-to-analog converter (TDT-DD1) and cycled indefinitely (16 bits, 50 000 samples/s). The signal was low-pass filtered below 20 kHz (−115 dB/octave) and sent to headphones, Sennheiser HD480, where it was heard diotically. The signal level was set so that each component had a level of 45 dB SPL.

The matching tone was generated by a Wavetek VCG116 function generator. Its frequency was controlled by a voltage from the listener's response box and was read by the computer through a reciprocal reading frequency counter on the computer bus. The matching tone was passed through a computer-controlled attenuator (TDT-PA4) so that its level could also be controlled and read by the computer.

3. Listeners

Three male listeners LB, LD, and LJ participated in the experiments. Their ages were 56, 20, and 34, respectively. LJ and LB were the authors. Audiograms showed that LD and LJ had normal hearing through 6 kHz, but that LB had a high-frequency hearing loss typical of males of his age. All the listeners had some training as performers of musical instruments and could perform a standard pitch matching task accurately.

B. Experiment 1: The frequency range of the DPE

To study the frequency range of the DPE, we used a flat spectrum with $N=170$ harmonics. To create the DPE, one harmonic, which could be as low as the 8th or as high as the 145th, was omitted from the complex tone spectrum. The purpose of this experiment was to test the frequency limits of the DPE for a 50-Hz complex tone.

The results are shown in Fig. 2. There are three panels in the figure, one for each subject. The vertical axis gives the matching ratio, defined as the listener's matching frequency divided by the frequency of the omitted harmonic. Therefore, when the ratio is unity the listener has made a perfect match.

The left four data points were tests of the low-frequency end of the DPE. The dot-dashed lines are the edges of the spectral gap created by omitting a harmonic. For example, when harmonic 10 is omitted, the edges of the spectral gap are harmonic 9 (ratio of 0.9) and harmonic 11 (ratio of 1.1). For omitted harmonics with very low harmonic number listeners seem to hear the edges and match them. Our criterion for the existence of the DPE at low frequency is that the listener's match should be closer to the omitted harmonic than to one of the edges. The data show that the three different listeners have different low ends: harmonic 13 for LB, 15 for LD, and 19 for LJ. Duifhuis (1972) found that the low-frequency limit of the DPE decreases with increasing signal level. From his results one would expect a lower limit between harmonics 10 and 20 for the levels used here.

The middle four points test the performance of the listeners at the middle range. All the listeners could make successful matches, although there is a small systematic pitch shift. The shift is positive as expected (Terhardt and Fastl, 1971). Such pitch shifts were also reported in the thesis by Duifhuis (1972). Informal comments by the listeners indicated that in the middle range the DPE was clear and loud, and there was no ambiguity in its pitch.

The four points on the right were intended to test the high-frequency end of the effect. Successful matches as high

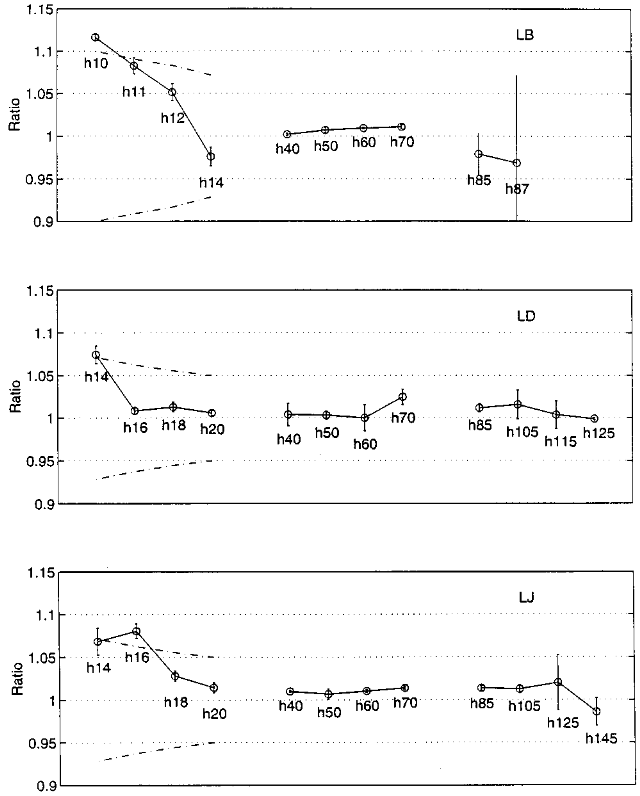

FIG. 2. Frequency matching results for experiment 1 showing matches. Each panel is for one subject. The *Ratio* is defined as the matching frequency divided by the frequency of the omitted harmonic. The symbol ''*h*...'' tells which harmonic is omitted in the complex tone. The left four points are for the low-frequency end of the effect. The dashed lines there show the edges of the spectral gap. The middle four points are in middle range of the effect. The right four points show an attempt to explore the high-frequency end of the effect.

as the 125th harmonic (6250 Hz) for LD and the 145th harmonic (7250 Hz) for LJ indicate a DPE at high frequency. However, matching became difficult and time consuming at these high frequencies, perhaps because of the unavailability of timing information above 5000 Hz (Johnson, 1980; Pickles, 1982; Moore, 1973). In the end, no upper limit was found for these two listeners. Listener LB made large errors for an omitted 87th harmonic (4350 Hz) and could not reliably detect the presence of a DPE for omitted harmonics 89 or 92. This high-frequency failure is possibly related to a high-frequency hearing loss. At 4500 Hz, LB showed a bilateral loss of about 16 dB. However, as shown in the next section, an effective reduction of 16 dB is not, by itself, enough to eliminate the DPE.

C. Experiment 2: Level effect

The purpose of the level effect experiment was to see whether the strength of the DPE is determined by the level of the cancellation tone. In this experiment the listener adjusted both the frequency and the level of a sine tone to match the DPE. The matching level was controlled through two buttons on the response box; pressing one increased the level by 1 dB, and pressing the other decreased the level by 1 dB.

Omitted harmonics at the lower end of the middle range, 25–40, were chosen for this experiment. The top harmonic

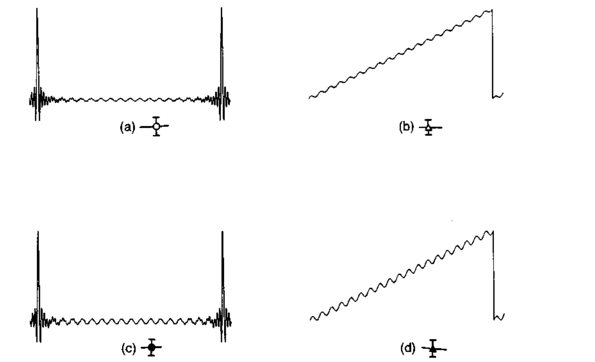

FIG. 3. Waveforms for experiment 2. (a) Waveform of a flat-spectrum cosine-phase complex tone with the 19th harmonic omitted; (b) sawtooth waveform with 19th harmonic omitted; (c) waveform of flat-spectrum cosine-phase complex tone with 19th harmonic phase inverted; (d) sawtooth waveform with 19th harmonic phase inverted. The damped high-frequency oscillations for flat-spectrum cases (a) and (c) are Gibbs phenomena from the upper end of the spectra. The high-frequency ends of the spectra were tapered to reduce this effect. Symbols with the letters are guides to Fig. 4.

number was reduced to 70, which had the effect of increasing the loudness of the DPE compared to the stimulus with 170 harmonics. All phase angles were again zero.

1. Spectra

Four spectra were used in this experiment. The illustrative waveforms are plotted in Fig. 3. Stimulus (a) was 70-component pulse train with one of its harmonics omitted. The amplitude spectrum was flat but tapered at the top to reduce the edge pitch. Stimulus (c) was the same as (a) except that the target harmonic was phase inverted instead of omitted. Therefore, there was no gap in the spectrum; the amplitude spectrum of (c) was identical to the amplitude spectrum of a pulse with all its harmonics equal.[1] For stimulus (c) the cancellation tone was 6 dB greater than in condition (a), and this fact can easily be seen in the waveform in Fig. 3(c). The oscillations between the pulses there have twice the amplitude as in Fig. 3(a). Stimulus (b) was a 70-component sawtooth wave with one of its harmonics omitted. The sawtooth wave is the integral of the pulse waveform (without its dc component), and so the amplitude of a harmonic in a sawtooth wave is inversely proportional to the harmonic number. For example, the 20th harmonic has twice the amplitude of the 40th. We expected to see a corresponding loudness variation in the DPE since the cancellation tone for different harmonic numbers has a different amplitude. Stimulus (d) was the same as stimulus (b) except that phase inversion replaced omission. The sawtooth waveforms were presented at the same level as the pulse train. Therefore, the original sawtooth waveform was given by

$$x(t) = -\sum_{n=1}^{70} \frac{6.552}{n} \sin(2\pi n f_0 t). \qquad (2)$$

2. Results

The level experiment produced both pitch matching data and loudness matching data. The pitch matches were consistent, as expected for this range of harmonic numbers. The

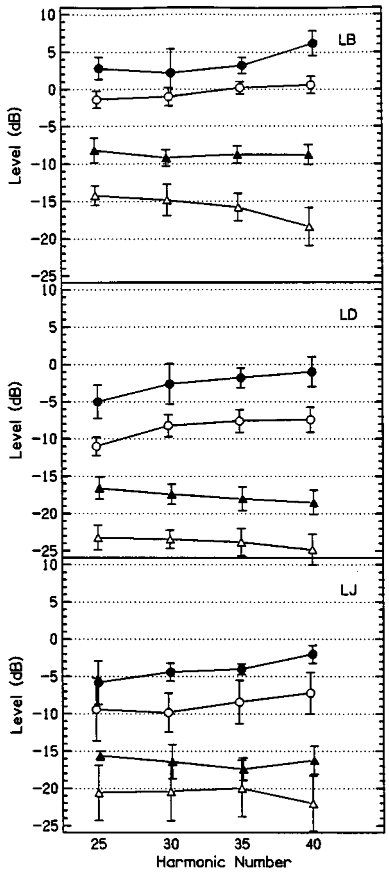

FIG. 4. Results of experiment 2. Each panel is for one subject. Matching levels of the DPE are plotted as a function of the harmonic number of the anomalous harmonic. Symbols identify waveforms *per* Fig. 3. Circle symbols are for a flat-spectrum pulse train, and triangles are for the sawtooth waveform. Open symbols are for omitted harmonics, and filled symbols are for phase-inverted harmonics. (Only relative levels for LD were available. The absolute levels shown here were adjusted *per* the analysis for LJ because early runs by LD and LJ were very similar.)

standard deviations were less than 1% of the matching frequency, and the pitch shifts were less than 2%. Of primary interest are the level matches, and these are shown in Fig. 4. There are three panels in the figure, one for each subject. The level matching scale on the vertical axis is in decibels referenced to the level of one component in a flat-spectrum complex tone, i.e., 45 dB SPL, corresponding to an amplitude of unity in Eq. (2).

For all three listeners, the relative level matches agree with the expected values for an objective cancellation tone. For both the flat spectrum and the sawtooth, the matches made with an inverted harmonic lie about 6 dB above the matches for the omitted harmonic. The average of the 24 differences is 5.4 dB (s.d.=1.5) to be compared with the expected value of 6 dB. Further, the matching levels for the sawtooth waveform agree with expectation when compared to the matching levels for the flat spectrum. The expected level differences (saw−flat) for n' values of 25, 30, 35, and 40 are, respectively: −11.6, −13.2, −14.6, and −15.7 dB. The measured values are −11.4 (1.0), −13.0 (1.9), −14.2 (2.2), and −16.3 (1.9) dB.

The absolute values of the level matches made by LB agreed with the actual level of the cancellation tones. Absolute levels matched by LD and LJ were lower by about 8 dB. The difference is possibly caused by a different treatment of the loudness reduction due to partial masking among the listeners. Quite generally the data from the level experiment support the idea that the loudness of the DPE is determined by the level of the cancellation tone.

3. Further level tests

Informal experiments found that the level of the DPE can be matched by the best beats method, in which the listener adjusts the level of an added sine tone to maximize beats. However, the level range for best beats was found to be rather wide, as large as 4 dB, especially in the omitted-harmonic condition. As Duifhuis (1970) noted, it is significant that an added sine tone can form beats with the DPE because this implies an objective character for the cancellation tone.

The sawtooth waveform has a particular phase spectrum, where all ϕ_n are equal to 90 degrees. Informal experiments showed that a DPE could also be heard for other constant values of the phase. Then the base waveform no longer has the straight sides of the sawtooth, but the base waveform is smooth so that omitted or inverted harmonics are still apparent on an oscilloscope.

The integral of a sawtooth is a parabola wave (Hartmann, 1997), with amplitudes decreasing at 12 dB/octave. The triangle wave has odd harmonics with amplitudes that also decrease at 12 dB/octave; the waveform is composed of straight line segments like the sawtooth. When harmonics of these waveforms were either omitted or inverted, no DPE could be found.

D. Experiment 3: The effect of harmonic phases

The phases of the harmonics play a critical role in the DPE. Duifhuis (1972) found that if the phases are randomized no effect can be heard at all. The purpose of the experiments of this section was to determine the behavior of the DPE under several coordinated changes of phase.

1. Spectra of the stimuli

The formal pitch matching experiments used a 100-component complex tone with four different phase conditions: (a) All the harmonics were cosine phase ($\phi_n=0$). (b) All the harmonics were sine phase ($\phi_n=\pi/2$). (c) All the harmonics had a 45-deg initial phase ($\phi_n=\pi/4$). (d) The harmonics were in Schroeder phase (Schroeder, 1970). For an equal-amplitude complex tone Schroeder phases are given by the equation

$$\phi_n=-\frac{\pi}{N}n^2. \tag{3}$$

To create a DPE one harmonic was omitted. The omitted harmonic numbers were in the middle range, 40–70.

2. Results

The consistency in tuning a matching frequency was used to determine the existence and strength of the DPE. The consistency of matches showed that all three listeners could hear the tone equally well in all four conditions. Frequency matching ratios for LB and LJ were in the range 1.00–1.02; for LD the range was 50% wider. Matching ratios increased somewhat with increasing harmonic number, and this tendency was greatest for the Schroeder phase condition. Otherwise, nothing special could be found in the matching frequencies or their variance to distinguish among the four phase conditions. The standard deviation of the ratio was about 0.005 for LB and LJ and about twice as large for LD.

For conditions (b) and (c), with common phase angles of 90° or 45°, the tone color of the background pulse became ''fuller'' than in condition (a) with zero phase. However, the DPE did not change. Waveforms made with different common phases are similar in that they all are quite spiky. They all have the same envelope with a temporal gap.

Unlike the other three waveforms, there is no time gap in the waveform with Schroeder phases. Therefore, no cancellation tone can be seen in an oscillographic representation of the waveform. Nevertheless, the experiments showed that the DPE exists. An explanation is proposed in the next section.

E. Additional phase effects

In addition to the formal phase experiments of the previous section, some informal phase experiments were done.

1. High and low omitted harmonics

We found it strange that the formal experiments revealed no difference between constant phase and Schroeder phase. Therefore, we performed informal matching experiments with omitted harmonics near 90 and 15 comparing cosine and Schroeder phases. No difference was found near harmonic 15, but the DPE was weaker for the Schroeder phase compared to cosine phase for omitted harmonics near 90, although this difference in strength was hardly noticeable in the matching data.

2. Low-fluctuation pulses

The Schroeder phase condition leads to a waveform with small power fluctuation. An alternative way to generate a waveform with small fluctuation is the low-noise noise method (Pumplin, 1985; Hartmann and Pumplin, 1991). Phase angles from the latter article were taken to make a signal having 36 equal-amplitude harmonics. Exploration with half a dozen values of n' found no DPE for either omitted or inverted harmonics. We conclude that the existence of the DPE for Schroeder phase is not solely the result of the small-fluctuation character of Schroeder-phase signals.

3. Random phases

When all the harmonics had random phases, the three listeners agreed that no DPE could be heard. In order to confirm this quantitatively, LB and LJ tried the matching experiment for ten different random-phase complex tones with omitted harmonics in the range 40–70. The results of their matches had a wide distribution. The average matching frequency for a given omitted harmonic showed no correlation with the frequency of the omitted harmonic, verifying that no DPE exists.

4. Alternating phases

When all even-numbered harmonics had phases of 0 and all odd-numbered harmonics had phases of 90°, or vice versa (*per* Plomp and Steeneken, 1969), a DPE could easily be heard for omitted or inverted harmonics with numbers greater than 33 in informal experiments. We concluded that the change from constant phases to alternating phases affected the low-frequency end of the DPE but not the high-frequency end. Qualitatively such a result is not unexpected because a waveform with alternating phases has two evenly spaced spikes per cycle, hence quiet gaps between ringing responses are half as long.

F. Experiment 4: The missing-fundamental DPE

A famous result in pitch research is that a complex periodic tone in which the fundamental is absent nevertheless has a pitch equal to the fundamental frequency. We attempted to create a missing-fundamental pitch using the DPE. When harmonics 18, 24, and 30 of the 50-Hz fundamental are omitted or inverted, the spectrum of the cancellation tones resembles the 3rd, 4th, and the 5th harmonics of 300 Hz. Informal experiments showed that such stimuli produce a pitch at 300 Hz, clearly suggesting a missing-fundamental DPE. The following experiment was a formal test of this phenomenon using phase inversion. Because the amplitudes were unchanged by this experiment, the amplitude spectrum was simply that of a 100-component equal-amplitude cosine-phase pulse train.

1. Spectra of the stimuli

There were four different conditions, each of which inverted the phases of three harmonics: (a) harmonics 18, 24, and 30; (b) harmonics 27, 36, and 45; (c) harmonics 36, 48, and 60; and (d) harmonics 42, 56, and 70. These harmonics resemble the third, fourth, and fifth harmonics of 300, 450, 600, and 700 Hz, respectively.

2. Results

Results from the three listeners are plotted in Fig. 5. The scale on the vertical axis is the ratio of the matching frequency to the frequency of the missing fundamental. The plot shows that all the listeners could match the missing-fundamental frequency with good reproducibility. The actual matching ratios tend to be less than 1, especially for LB. This result can be compared with the measurements of Terhardt (1971) showing that the pitch of the missing fundamental of a complex tone is about 2% below the pitch of a sine tone with the frequency of the fundamental. LJ and LD show a shift that is less than 2%; LB shows a shift that is larger than 2%.

II. PERIPHERAL FILTER MODEL FOR THE DPE

The explanation for the DPE by Duifhuis (1970, 1971) is that when the frequency spacing between harmonics is considerably less than the bandwidth of relevant peripheral

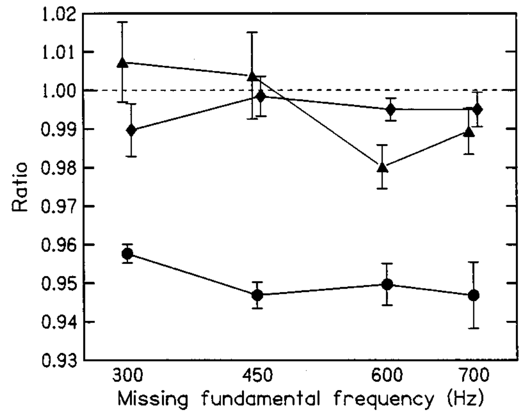

FIG. 5. Results of experiment 4. Matches to the missing fundamental tone are shown for each subject: circles for LB, triangles for LD, and diamonds for LJ. The matches are plotted as the ratio of matching frequency to missing-fundamental frequency.

auditory filters, the peripheral filters can be said to be broadband. A broadband frequency analyzing system has a short impulse response time, substantially less than the period of the input signal. Therefore, when the input is a pulse train with an omitted harmonic, there is a portion of time in each period during which the output of the peripheral filter is dominated by the small oscillations of the cancellation tone. That portion of the output in each period can be detected as a sine tone. A bank of such filters therefore produces a short-term frequency analysis of the input signal.

To get a clearer picture of the short-term frequency analysis, we implemented the idea with a realistic auditory filter bank—the gammatone filter bank (Patterson, 1987). Figure 6 is a three-dimensional plot of the output of a gammatone filter bank. The bandwidths of the filters are determined by the ERBs from Glasberg and Moore (1990). The input to the filter bank is the 50-Hz complex tone with harmonics from 1 to 45. The x axis is a time axis. Its unit is the period of the input signal, 20 ms. Therefore, the plot shows exactly one period of the output. The y axis is the center frequency of a filter in the filter bank. The unit is chosen to be 50 Hz so that harmonics of the signal appear at integer values on this axis. Harmonic n appears at integer value n. The z axis shows the output, normalized to a maximum value of one. The output is half-wave rectified to improve the three-dimensional impression in the figure.

The response shown in Fig. 6 is essentially the impulse response of the auditory filter bank. Filters with a high center frequency, e.g., $26\times 50=1300$ Hz, are relatively broad and have a short impulse response. There is a decided gap in their output. Filters with a lower center frequency, e.g., $10\times 50=500$ Hz, have a longer ring time and no appreciable gap in their output.

Figure 7 is similar to Fig. 6 except that the 19th harmonic (950 Hz) of the input tone is omitted. Figure 7 shows a valley in the ringing region of the output, and a ''peak'' appears in the previously quiet region. This peak is actually the same response as would be obtained had the input been a 950-Hz sine tone. It is this peak that is the traditional explanation of the DPE.

The traditional explanation may also be applied to the DPE heard in the Schroeder-phase condition. The Schroeder phase leads to a frequency sweep (chirp) with the duration of a period. The sweep moves continuously from the highest frequency in the band to the lowest. Therefore, there is some

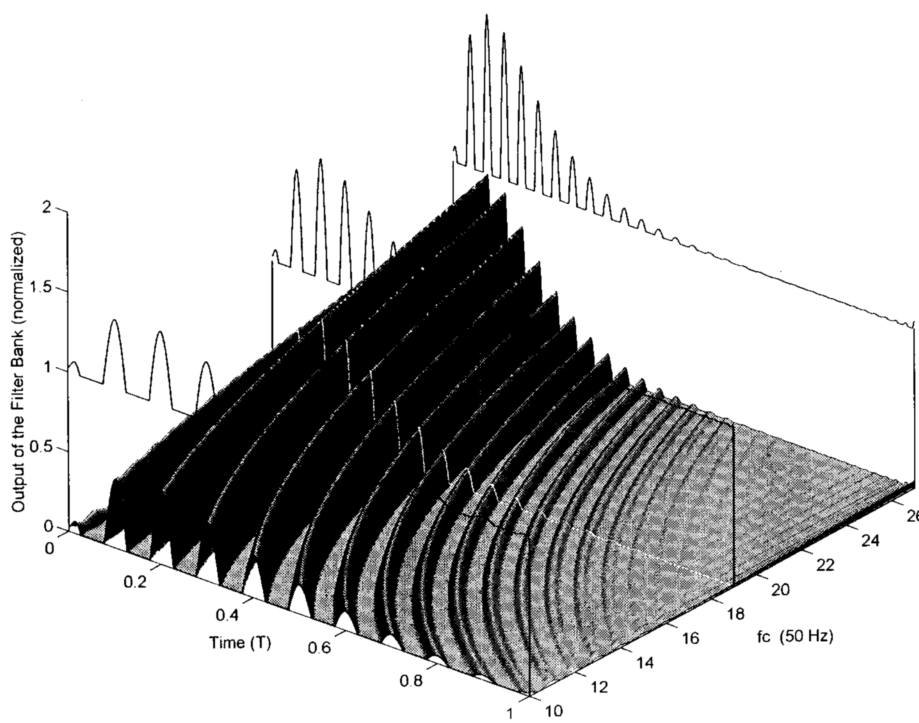

FIG. 6. Three-dimensional plot of the output of the model peripheral auditory filter bank. The input signal is a 45-harmonic flat-spectrum complex tone with a fundamental frequency of 50 Hz. All the harmonics have cosine phase. The elevated functions show the time-dependent outputs of model filters centered on reference harmonic numbers.

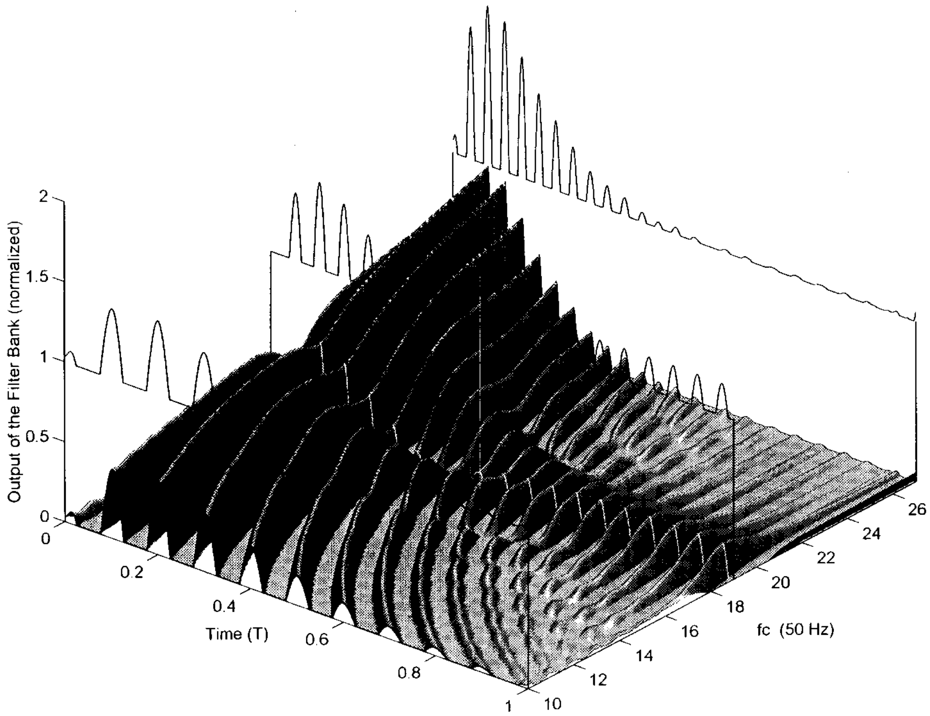

FIG. 7. Three-dimensional plot of the output of the model peripheral auditory filter bank. The input signal is the same as for Fig. 6 except that the 19th harmonic is omitted.

fraction of the period within which the instantaneous frequency of the stimulus differs greatly from the center frequency of any particular auditory filter. Therefore, there is a quiet gap in the temporal output of any auditory filter, but different filters experience the gap at different times within the signal period, as shown in Fig. 8. When a harmonic is omitted or inverted, a peak can appear across the quiet gap in the output of filters having their center frequencies in the range of the anomalous harmonic, as shown for an omitted 19th harmonic in Fig. 8. This transitory peak might well produce a DPE.

III. CHALLENGE TO THE PERIPHERAL FILTER MODEL OF THE DPE

According to the traditional explanation, the DPE is the result of the sine-tonelike peak in the quiet gap in the output of a peripheral filter bank. The challenge to this explanation

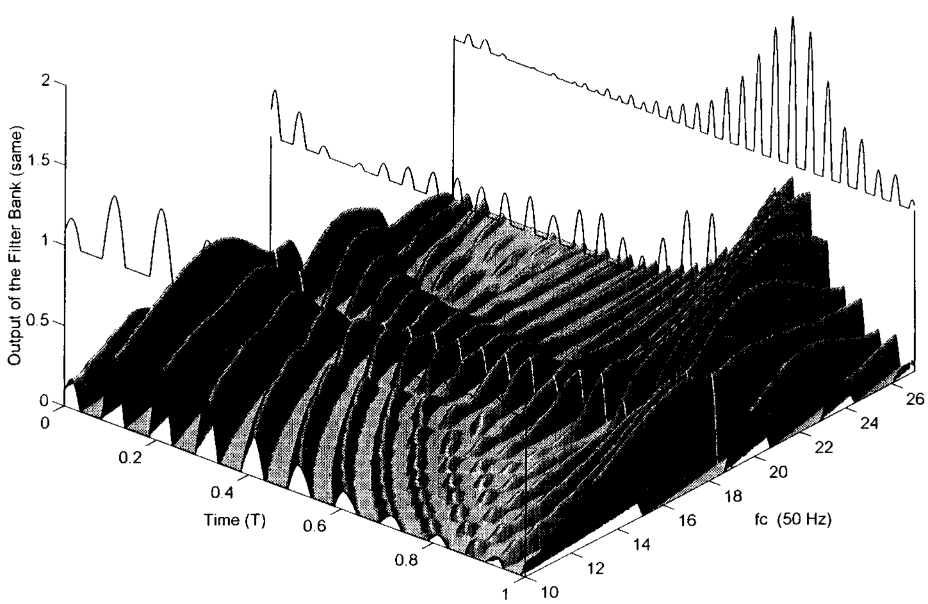

FIG. 8. Three-dimensional plot of the output of the model peripheral auditory filter bank when the harmonics of the complex tone are in Schroeder phase and the 19th harmonic is omitted.

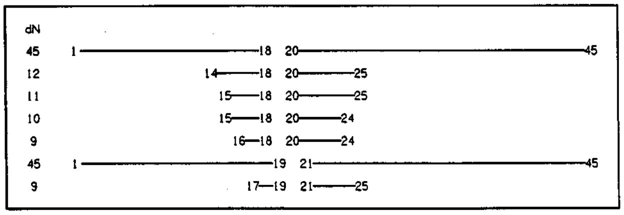

FIG. 9. The seven spectral bands for the band-narrowing experiment. For example, the fourth line shows the spectrum for $dN=10$ harmonic numbers in the band. They are harmonics 15–18 and 20–24, where 19 has been omitted.

begins with the observation that the peak involves only a limited number of harmonics. For the omitted 19th in Fig. 7, for example, only the harmonics in the range 17–22 make an important contribution. Therefore, it follows that if all the harmonics outside this range were removed, the DPE should be unchanged. It might even be stronger because there would be less competition from the other harmonics of the pulse train. The first of our challenge experiments systematically deleted harmonics outside the range of the omitted harmonic.

A. Experiment 5: Band narrowing

The band-narrowing experiment was a matching experiment like those previously described. However, because the central issue in the band-narrowing experiment was the existence of the DPE, we took particular precautions to avoid cuing a harmonic. Therefore, the listener was not allowed to alternate between the complex tone and the matching sine tone. The listener was allowed to listen to the complex tone as many times as he wanted, but once he decided to do the matching and heard a matching sine tone, he was not allowed to listen to the complex tone again. As a further precaution, the sampling rate of the digital to analog converter was randomized by ±5% on each trial to reduce the possibility of cues across trials.

1. Complex tones

There were seven complex tones in the experiment. Five of them had an omitted 19th harmonic, and two had an omitted 20th. Their spectra are described by Fig. 9. Each harmonic had cosine phase and a level of 45 dB SPL as before. Therefore, the narrowed signals were bandpassed versions of the pulse-train signals used previously. The high and low edges of the spectrum were tapered (0.7, 0.3, as before). Figure 10 shows the effect of narrowing the spectrum in the output of the auditory filter bank. The input for Fig. 10 was the signal on the fourth line of Fig. 9 with $dN=10$ harmonics, numbers 15–24 with the 19th omitted. A comparison between the calculated output for Figs. 7 and 10 shows that the sine-tonelike peak in the filter centered on the 19th harmonic changes very little. We expect correspondingly little change in the DPE.

Each of the three listeners did 20 runs of the band-narrowing experiment. Each run required seven matches, one match for each of the complex tones in Fig. 9.

2. Results

The matching data for each listener and each complex tone were plotted in a histogram with its bin width equal to the harmonic spacing (50 Hz). The histograms for the broadband signals (harmonics 1–45) were concentrated on the omitted harmonic, indicating that listeners could easily hear the DPE. As the spectrum was narrowed, the histograms widened considerably, rapidly approaching half the bandwidth of the entire stimulus. The dependence of the 21 histograms in the experiment are here summarized in Fig. 11 by plotting the percentage of matches in the central histogram

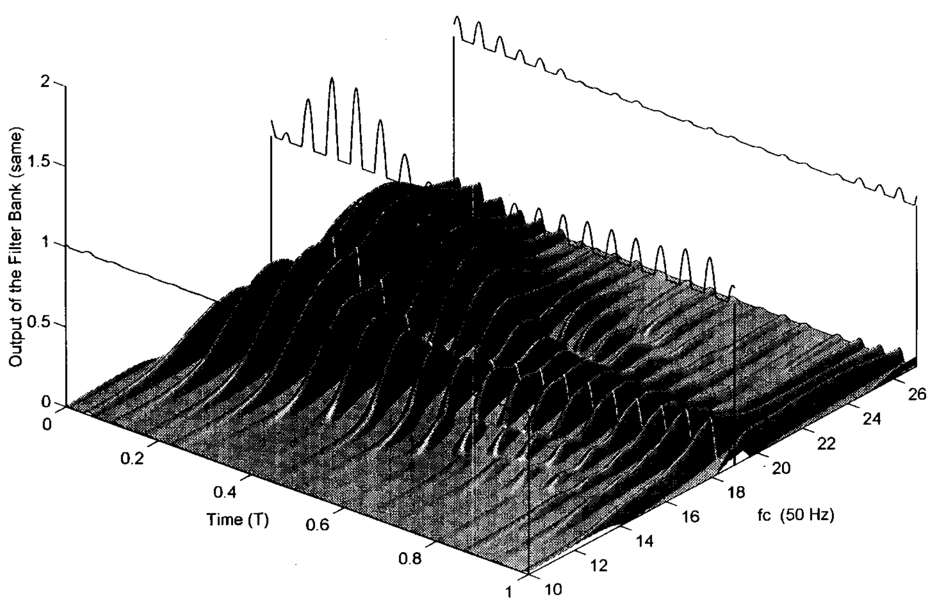

FIG. 10. Three-dimensional plot of the output of the model peripheral auditory filter bank. The input signal is the signal with $dN=10$ from Fig. 9, where all the harmonics, 15–24, have cosine phase, and the 19th harmonic is omitted. Therefore, the input signal was the same as in Fig. 7 passed through a bandpass filter. The region around harmonic 19 can be seen to be similar to Fig. 7, where the signal is broadband.

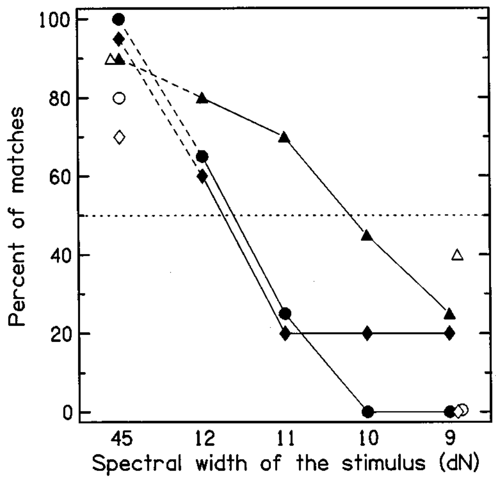

FIG. 11. Plots of the percentages of the matches to a missing 19th harmonic that are in the 19th harmonic bin as a function of the bandwidth of the stimulus. Circles are for LB, triangles are for LD, and diamonds are for LJ. Spectra for different values of dN are defined by Fig. 9. Open symbols are for the missing 20th harmonic.

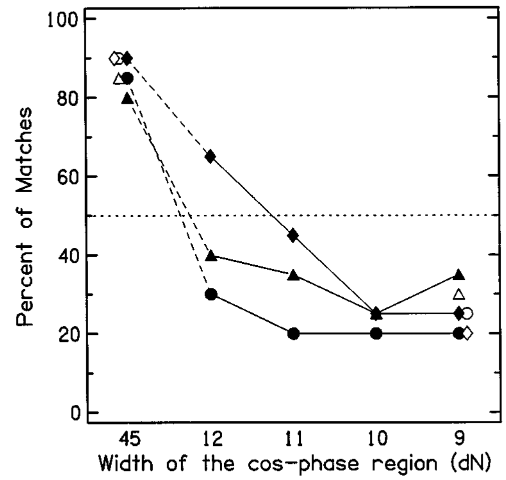

FIG. 12. Plots of the percentages of the matches to a missing 19th harmonic that are in the 19th harmonic bin as a function of the bandwidth of the phase coherence region. Circles are for LB, triangles are for LD, and diamonds are for LJ. Open symbols are for a missing 20th harmonic.

bin (matching ratio=1) for the five complex tones with an omitted 19th harmonic. That is, the percentage of the matches that are closer to the 19th harmonic than to any other harmonic. Open symbols show the percentages for the omitted 20th harmonic, which serve as supplementary data. It is evident from Fig. 11 that as the band was narrowed, the listeners were unable to make accurate matches to the DPE.

For all the listeners the steepest slope in Fig. 11 occurs near the 50% point. Therefore, we define this as the boundary for hearing the DPE, i.e., if the number of matches in the central bin is greater than 50% of the total, the pitch is said to be ''heard,'' otherwise it is ''not heard.'' Applying this criterion to our results, we find that none of the listeners could hear the DPE for dN equal to 9 or 10 (*per* Fig. 10). This result came as no surprise to the listeners themselves; informally they remarked that no DPE was heard for these narrow bands. The conclusion of the band-narrowing experiment is that the peripheral filter model fails the challenge.

B. Experiment 6: Narrowing the phase-coherence region

Because the DPE exists with a broadband tone with constant harmonic phases but does not exist for random phases, there is another way to narrow the effective band of the stimulus. The narrowed-phase coherence-region (NPCR) experiment retained cosine phases for the harmonics near the omitted harmonic, but randomized the phases of all the other harmonics in the signal. Harmonics 1–45 all had unit amplitude, except for an omitted harmonic. The regions of cosine phase were exactly the spectral regions shown in Fig. 9. Thus, the NPCR experiment paralleled the band-narrowing experiment. The three listeners each did 20 runs using the protocol of the band-narrowing experiment.

1. Results

The matching data from the NPCR experiment were plotted in histograms like the data for the band-narrowing experiment. Unlike the matches in the band-narrowing experiment, which were confined because of the narrowed stimulus spectrum, the matches in the NPCR experiment were spread over a very wide frequency range.

The percent of matches in the central bin for the omitted 19th harmonic is plotted in Fig. 12 as a function of the bandwidth of the cosine-phase region, dN. The percentage decreases as the bandwidth decreases. Again, defining 50% as the boundary, we find that none of the listeners could hear the pitch resulting from Duifhuis effect for dN equal to 9, 10, or 11. This result shows that phase information from components that are distant from the omitted harmonic is needed in order for the DPE to exist.

IV. DISCUSSION AND CONCLUSION

The Duifhuis pitch effect (DPE), as studied in this report, is created by starting with a low-frequency complex tone having many harmonics and then omitting a harmonic or inverting the phase of a harmonic. This spectral change is perceived as an added pure tone with a pitch equal to the frequency of the omitted or inverted harmonic. The traditional explanation for the DPE can be regarded as a two-step process. The first step considers the response of a broadband system to the stimulus, and the second step considers the circumstances in which the peripheral auditory system adequately approximates such a broadband system.

The first step essentially considers the waveform of the stimulus as it would appear on an oscilloscope. It establishes the general rule that to generate a DPE, the phases of the original complex tone must be chosen so that the waveform includes a quiescent time interval in which the omitted harmonic can be heard as a cancellation tone. The experiments by Duifhuis used a pulse train for the original complex tone. The experiments of Sec. I in the present report used a variety of original complex tones.

Many of the results from Sec. I were consistent with the cancellation tone concept. The DPE was found to exist when the waveform included some form of quiescent interval: flat-spectrum pulse with constant phase, sawtooth, Schroeder-phase chirp, or alternating-phase pulse. The DPE was found not to exist in the absence of such characteristic time structure, specifically for random phases or phases producing minimum power fluctuation.

The continuation of the DPE well above 5 kHz makes it unlikely that the effect is caused by some aspect of neural timing because jitter causes neural synchrony to vanish above this frequency (Johnson, 1980; Pickles, 1982; Moore, 1973). The generation of a DPE by phase changes alone, while leaving the power spectrum unchanged (inverted harmonic case), rules out an explanation based on simple autocorrelation of the waveform. Instead, the effect appears to result from an objective cancellation tone. The good agreement between the loudness matches and the level of the cancellation tone for the different complex tones in experiment 2 supports this interpretation; so does the generation of a missing-fundamental DPE in experiment 4.

The second step in the traditional explanation of the DPE requires that the omitted or inverted harmonic be in a frequency range where the bandwidth of peripheral auditory filters is considerably greater than the fundamental frequency of the complex tone. Only then do the auditory filters contain enough harmonics to produce a quiescent time interval in which the cancellation tone can be heard. This idea predicts that the effect should exist only for harmonics too high to be resolved.

Experiments by Duifhuis (1972) found a low-frequency limit near the 10th harmonic for 50-dB, 50-Hz tones; experiment 1 of this report suggests a limit near the 15th. Calculations using gammatone auditory filters (Lin, 1996) show that a lower limit near 15 is not unreasonable for a 50-Hz tone. The observation (Duifhuis, 1972) that the low-frequency limit depends on level then suggests that the bandwidth of auditory filters increases significantly with increasing level. There is some evidence that the level dependence of bandwidth is particularly prominent for impulsive signals of this kind (Hartmann and Rakerd, 1993).

The band-narrowing experiments of Sec. III challenged the traditional explanation for the DPE because they showed that no DPE exists for narrowed-spectrum complex tones. By contrast, an auditory filter model shows that the cancellation tone in the quiescent interval is essentially the same for the narrowed-spectrum tones as for the broadband tones for which the DPE is always observed. Thus, although it seems to be *necessary* for auditory filters near the omitted harmonic to isolate the cancellation tone in a quiescent interval, it is not *sufficient*.

The band-narrowing experiments showed that the existence of the DPE depends on the presence of harmonics well outside the range of frequencies passed by auditory filters near the omitted harmonic. The DPE therefore appears to be a broadband effect in which information is combined across different auditory channels. What is especially notable is that this cross-channel combination must include phase information. Experiment 6 showed that the DPE requires phase coherence over a broad frequency region. Cross-channel phase comparison is an unusual idea in auditory theory. However, the phase coordination of separate channels by a broadband impulsive signal does not seem implausible. Possibly even the nonsimultaneous, but regular, temporal pattern generated by Schroeder-phase signals could lead to a coordinated response.

ACKNOWLEDGMENTS

We are grateful to Dr. Hendrikus Duifhuis and Dr. Roy Patterson for helpful discussions. This work was supported by the National Institute on Deafness and Other Communicative Disorders, Grant No. DC00181.

[1]Inversion can lead to amplitude spectrum changes if there is nonlinear distortion. The amplitude spectrum, measured at the headphones with a microphone, revealed no changes caused by inversion within the precision of our measuring process, 0.25 dB.

Alcantara, J. I., and Moore, B. C. J. (**1995**). ''The identification of vowel-like harmonic complexes: Effects of component phase, level and fundamental frequency,'' J. Acoust. Soc. Am. **97,** 3813–3824.

Duifhuis, H. (**1970**). ''Audibility of high harmonics in a periodic pulse,'' J. Acoust. Soc. Am. **48,** 888–893.

Duifhuis, H. (**1971**). ''Audibility of high harmonics in a periodic pulse. II. Time effect,'' J. Acoust. Soc. Am. **49,** 1155–1162.

Duifhuis, H. (**1972**). ''Perceptual analysis of sound,'' thesis, University of Eindhoven, The Netherlands, unpublished.

Glasberg, B. R., and Moore, B. C. J. (1990). ''Derivation of auditory filter shapes from notched noise data,''Hear Res. **47**, 103–138.

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. (**1997**). *Signals, Sound and Sensation* (AIP Press, New York).

Hartmann, W. M., and Pumplin, J. (**1991**). ''Periodic signals with minimal power fluctuations,'' J. Acoust. Soc. Am. **90,** 1986–1999.

Hartmann, W. M., and Rakerd, B. (**1993**). ''Auditory spectral discrimination and the localization of clicks in the sagittal plane,'' J. Acoust. Soc. Am. **94,** 2083–2092.

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

Kubovy, M., and Jordan, R. (**1979**). ''Tone segregation by phase: On the phase sensitivity of the single ear,'' J. Acoust. Soc. Am. **66,** 100–106.

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

Lin, J.-Y. (**1996**). ''Psychoacoustical theory and experiments on human auditory organization of complex sounds and the critical bandwidth,'' thesis, Michigan State University, unpublished.

McAdams, S. (**1984**). ''The auditory image: A metaphor for musical and psychological research on auditory organization,'' in *Cognitive Processes in the Perception of Art* (North-Holland, Amsterdam).

Moore, B. C. J. (**1973**). ''Frequency difference limens for short duration tones,'' J. Acoust. Soc. Am. **54,** 610–619.

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.

Patterson, R. D. (**1987**). ''A pulse ribbon model of monaural phase perception,'' J. Acoust. Soc. Am. **85,** 1560–1586.

Pickles, J. O. (**1982**). *An Introduction to the Physiology of Hearing* (Academic, London).

Pierce, J. R. (**1960**). ''Some work on hearing,'' Am. Scientist **48**, 40–45.

Plomp, R., and Steeneken, H. J. M. (**1969**). ''Effect of phase on the timbre of complex tones,'' J. Acoust. Soc. Am. **46,** 409–421.

Pumplin, J. (**1985**). ''Low-noise noise,'' J. Acoust. Soc. Am. **78,** 100–104.

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

Terhardt, E. (**1971**). ''Die Tonhohe harmonischer Klange und das Oktaveintervall,''Acustica **24,** 126–136.

Terhardt, E., and Fastl, H. (**1971**). ''Zum Einfluss von Stortonen und Storgerauschen auf die Tonhohe von Sinuston,''Acustica **25,** 53–61.