Binaural coherence edge pitch[a]

William M. Hartmann[b] and Colleen D. McMillon[c]

Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

The binaural coherence edge pitch (BICEP) is a dichotic broadband noise pitch effect similar to the binaural edge pitch (BEP). The BICEP stimulus is made by summing spectrally dense sine wave components with random phases. The *interaural* phase angle is a constant (0 or π) for components with frequencies below (or above) a chosen edge frequency, and it is a random variable for the remaining components. The chosen edge frequency is a coherence edge because the noises to the two ears are mutually coherent within any band of frequencies on one side of the edge and they are mutually incoherent in any band on the other side. Pitch-matching experiments show that the BICEP exists for coherence edge frequencies between about 300 and 1000 Hz. It is matched by a pure-tone frequency that differs from the edge frequency by 5% to 10%. The matching frequency lies on the incoherent side of the edge, an important result that is consistent with the way that the equalization-cancellation model has been applied to binaural pitch effects, especially the BEP. The results of BICEP experiments depend upon whether the coherent components are presented in 0 or π interaural phase for some listeners but not for all. The BICEP persists if the noise to one of the ears is delayed, but it becomes weaker and less well matched as the delay increases beyond 2 ms. The BICEP does not depend on whether the component amplitudes are all created equal or are given a Rayleigh distribution. Some reliable pitch sensation exists even when the component amplitudes are entirely independent in the two ears, so long as the phase coherence conditions of the BICEP stimulus are maintained. The existence of the BICEP is a challenge for current models of dichotic pitch because none of them predicts all its features. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1331680]

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

I. INTRODUCTION

The binaural coherence edge pitch (BICEP) is a dichotic noise pitch effect, made with white noise wherein a particular interaural phase relationship causes a pitch to appear. Therefore, the BICEP is a member of the family of effects that began with the Huggins pitch and includes the binaural edge pitch (BEP). Like the Huggins pitch and BEP, the BICEP sounds like a narrow band of noise or like a pure tone embedded in noise.

A. The Huggins pitch

The first, and strongest, tonotopically local dichotic pitch effect is the Huggins pitch (Cramer and Huggins, 1958). It is created by dichotic broadband noise having a transition frequency region over which the interaural phase changes by 360 deg. For example, the interaural phase angle might be zero for all frequencies up to 580 Hz and 360 deg (equivalent to zero) for all frequencies above 620 Hz. Between 580 and 620 Hz is the transition region where the interaural phase changes smoothly from 0 to 360. In the cen-

ter of the transition region, at a frequency of 600 Hz, the interaural phase is 180 deg. Such a stimulus produces a pitch with a frequency of about 600 Hz. The width of the transition region is a critical experimental variable. The width described above, 7% of the center frequency, is about optimal for hearing the effect (Cramer and Huggins, 1958; Hartmann, 1979).

Durlach (1962) explained the Huggins pitch in terms of his equalization-cancellation (EC) model (Durlach, 1972), in which the signals to the two ears can be binaurally added or subtracted in the central auditory system.[1] Applied to the above stimulus, the binaural subtraction operation leads to a central spectrum that is zero everywhere except in the phase transition region. Within the transition region there is a peak centered at 600 Hz, in agreement with the perceived frequency.

Evidence in favor of the EC explanation for the Huggins pitch can be found by creating a monaural analog to the Huggins effect. The analog simulates the supposed action of the binaural system by electronically subtracting the left and right signals of the dichotic stimulus to make a monaural stimulus in which most of the noise components are canceled. Perception of the resulting noise band can be compared with the percept obtained dichotically in Huggins pitch. When the noises are electronically subtracted, the peak in the monaural spectrum leads to a spectral pitch of about 600 Hz as expected. If the width of the transition region is increased, the pitch tends to disappear. As a function of transition region width, the dichotic pitch (Huggins pitch) disap-

pears in the same way as the spectral pitch, electronically generated by subtracting channels (Hartmann, 1979). At every value of the width, the change in pitch strength is about the same for both kinds of presentation.

B. The binaural edge pitch

The binaural edge pitch (BEP) stimulus (Klein and Hartmann, 1980) resembles the Huggins pitch stimulus, but the phase variation is only 180 deg, and the width of the transition region may be zero. Applying the cancellation process from Durlach's model leads to a central spectrum which is either a high-pass or low-pass noise band. For example, if the noise components in the two ears are in phase below the phase transition frequency and 180 deg out of phase above that frequency, then binaurally adding noises in the two ears would lead to a low-pass central spectrum, whereas subtracting noises in the two ears would lead to a high-pass central spectrum. Because noise bands with sharp spectral edges (monaural or diotic) produce pitches near the edges (Small and Daniloff, 1967; Fastl, 1971), it was predicted, by analogy with Durlach's argument, that the central spectrum created by the dichotic stimulus would also have a pitch, namely the BEP. Experiments showed that the BEP did exist, though it was weaker than the Huggins pitch. It is perhaps not extending the analogy too far to say that the BEP is weaker than the Huggins pitch by about the same amount as the spectral edge pitch is weaker than the pitch of the monaural spectral peak generated by subtracting the Huggins pitch channels.

The comparison between BEP and the pitch of a monaural spectral edge is strengthened by the results of pitch-matching experiments. A monaural noise with a spectral edge has a pitch that is close to the edge, but somewhat within the noise band itself. A low-pass noise band leads to a pitch that is a few percent below the cutoff frequency; a high-pass noise band has a pitch a few percent above the cutoff (Klein and Hartmann, 1980; Frijns *et al.*, 1986). Application of this monaural analogy to the BEP involves an ambiguity. Whether the BEP central spectrum most resembles a high-pass or low-pass noise band depends on the EC operation applied by the central auditory system, binaural addition, or subtraction. Klein and Hartmann found evidence that the BEP is bimodal, and they analyzed all their data in that way. The two peaks of the bimodal distribution compared well with spectral edge pitches that were obtained by physically combining the two channels to make high-pass and low-pass bands. Thus, the ambiguity present in the model appeared to be present in the experimental pitch matches as well.

It was conjectured that the pitches of both monaural spectral edges and binaural edges were caused by Mach bands due to lateral inhibition. In the case of a spectral edge, neurons tuned to noise components within the noise band but close to the edge would not be inhibited as much as neurons tuned to frequencies well inside the noise band. The effect of decreased lateral inhibition on these neurons near the edge of the band would produce a peak in the excitation pattern at a tonotopic place corresponding to a sine tone just inside the noise, as observed experimentally. In the case of the *binaural*

edge, central neurons following the EC process would similarly reflect Mach bands. This conjecture for the binaural edge was therefore consistent with the bimodal distribution of the BEP with peaks just above and below the edge frequency. The agreement between experiment and theory was interpreted as support for the EC model. Using the EC model to explain the BEP data depended on the assumption that the binaural system would sometimes use binaural addition of channels and sometimes use subtraction.[2]

C. The binaural coherence edge pitch

A further prediction that follows from the EC model is that a central spectrum with an edge can be created by a binaural coherence edge. For example, if all the spectral components below an edge frequency are identical in amplitude and phase in the left and right channels, and all the components above the edge frequency have uncorrelated phases, then an operation that binaurally subtracts the left and right noises leads to a central edge because the low-frequency components cancel exactly while the incoherent components above the edge survive the subtraction operation. The central spectrum generated by this stimulus should resemble a high-pass noise with a sharp edge leading to a pitch. This predicted pitch is the BICEP. By analogy with monaural high-pass noise this pitch should lie above the coherence edge frequency. Similarly, noise that is coherent above an edge frequency and incoherent below the edge should lead to a pitch that is below the edge frequency.

These then are the predictions for the BICEP. If these predictions hold then the BICEP has a particular advantage over the BEP because the BICEP is predicted to be above or below the coherence edge unambiguously as determined by the stimulus. That is because only the binaural subtraction process leads to a sharp central edge with the BICEP stimulus. In contrast, either binaural subtraction or binaural addition leads to a sharp central edge with the BEP stimulus. Therefore, the BEP depends on a choice made by the central processor; the choice of binaural subtraction or addition determines whether the pitch is below or above the binaural phase edge, and that choice is not experimentally controllable. Because of this theoretical advantage of the BICEP, it is of considerable interest to know whether the BICEP exists and whether it is shifted away from the edge frequency as predicted.

II. EXPERIMENT 1—MIDDLE RANGE

A. Method

Pitch-matching experiments were done to see if the BICEP exists and, if it does exist, to check the predicted pitch shifts. The matching experiments used an alternating sequence of BICEP stimulus and sine wave matching tone. The listener could adjust the frequency of the matching tone to best match the pitch heard in the BICEP stimulus.

B. Stimuli

To make a noise with a binaural coherence edge, the left channel was created by adding 16 384 spectral components

having equal amplitudes and random phases over a 360-deg range. The components for the right channel were identical in amplitude and phase to the components for the left channel for frequencies on the coherent side of the edge. For frequencies on the incoherent side of the edge, the amplitudes were again equal, but the phases were rerandomized over a 360-deg range.

Noise computations were done in a Tucker-Davis (TDT) AP2 array processor. Noises were converted to analog form by the 16-bit DACs on a TDT DD1. The noise stimulus buffers were recomputed, with a new set of 16 384 random component phases, prior to each matching trial. For each channel, the buffer was 32 768 samples in length, and the sample rate was 20 ksps. Therefore, the cycle time was 1.6 s and the components of the noise were separated by 0.61 Hz. Thus, the transition from complete coherence to complete incoherence occurred over a frequency difference of 0.61 Hz. The noises were low-pass filtered at 8 kHz and −115 dB/ octave.

The experiment was run by a microcomputer which also controlled the TDT system. The experimental matching sequence consisted of four intervals: BICEP stimulus, silent gap, matching tone in diotic noise, and silent gap.[3] All intervals were 500 ms long, incommensurate with the buffer cycle time and short enough that listeners were unaware of the 1.6-s repetition. The diotic noise presented with the matching tone had the same level as the BICEP noise, and, by adjusting the matching tone frequency and level, the listener could make the matching interval sound like the BICEP interval. The frequency and level adjustments were made with ten-turn potentiometers on the listener's response box. Potentiometer voltages were read with 16-bit ADCs and used to control a TDT WG2 pure-tone generator. The ten turns of the frequency control allowed a range of two octaves, and push buttons, also on the response box, allowed the listener to move that range up or down by arbitrarily large amounts.

The noise stimuli were presented at a level of 63 dB SPL (24-dB spectrum level) by Sennheiser HD 480 headphones. The level was high enough to produce strong binaural effects while avoiding cross talk. Listeners heard the stimuli while seated in a double-walled sound-treated room.

Matching trials were blocked as experimental runs, consisting of one match to each of five coherence edge frequencies, 550, 600, 650, 700, and 750 Hz, presented in random order. There was no time limit for making a match, and runs typically lasted 3 or 4 min. There were four different conditions: The coherent region could either be above the edge frequency or below the edge frequency, and the coherent components could be either in 0 phase (identical) or π phase (inverted). Each listener did ten runs for each condition.

C. Listeners

There were five listeners, C, J, M, T, and W. Listeners C and M were females, age 19. Listeners J, T, and W were males, ages 21, 21, and 59. Listeners C and W were the authors. All listeners had normal thresholds through 8 kHz except for W, whose thresholds were elevated above 4 kHz as typical for males of that age. Listener J was unusual, having had no musical training at any time in his life.

TABLE I. Monotonicity for five random matches. The table shows the expected percentage for which there are no (0) deviations from monotonically increasing matches with increasing boundary frequency; also the percentage of sets with one deviation, two deviations, three deviations, and four deviations.

Condition	0	1	2	3	4	Deviations
Random matches	0.8	21.7	55.0	21.7	0.8	

D. Results

Because of the predicted pitch shifts, it is not possible to judge the reliability of the pitch matches by direct comparison with the edge frequency. Alternative tests are (1) matches that are monotonic with edge frequency, and (2) matches that are mutually consistent.

1. Monotonicity statistic

The probability that five matching frequencies chosen at random are in a monotonically increasing order is 0.8 percent. Therefore, if the matches obtained experimentally rise monotonically with increasing edge frequency, there is good evidence that the listener heard a reliable pitch depending on the edge frequency. The probability that random frequencies are in ascending order except for one deviation is 21.7 percent. Therefore, if the experimental matches rise monotonically except for one deviation, then that too is evidence for a reliable pitch. This probability argument is the basis of the monotonicity statistic which is used throughout this article. The expected results for random matches are shown in Table I. This table serves as a basis for comparison for all the other tables in the article.

For experiment 1, the monotonicity statistic is given in Table II for the four conditions and five listeners.[4] A plot that corresponds to the average of the four lines of Table II appears as the dark histogram in Fig. 1 for illustration. The data in Table II indicate the presence of a pitch depending on the edge frequency for all stimulus conditions and all listeners. For instance, over all listeners and all conditions Table II shows perfectly monotonic matches for more than 50 percent of the trials, far greater than 1 percent as expected for random matches. Some listeners scored 100 percent. Statistics for monotonicity except for one deviation are similarly out of line with random performance. Although this statistic ignores the magnitude of the deviation from monotonically increasing matches, most deviations were just exchanges of neighboring positions when there was only a single violation. In summary, Table II shows unequivocal evidence for the existence of the BICEP.

TABLE II. Monotonicity for BICEP. Experiment 1. The table shows the percentage for which there were no (0) deviations from monotonically increasing matches with increasing boundary frequency; also the percentage of runs with 1 deviation, two deviations, three deviations, and four deviations. Results are averaged over five listeners, C, J, M, T, and W.

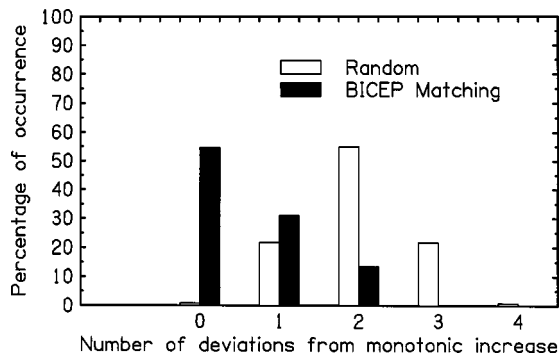| Condition	0	1	2	3	4	Deviations
Zero phase, coherent below	58	32	10	0	0	
Zero phase, coherent above	38	28	32	2	0	
Pi phase, coherent below	80	18	2	0	0	
Pi phase, coherent above	42	46	10	2	0	

FIG. 1. Graphic example displaying the monotonicity statistic. The solid histogram shows average monotonicity statistic (five listeners and four conditions from Table II). The open histogram shows the expected monotonicity statistic for random pitch matches, Table I.

2. Pitch shifts

The pitch-matching results of the four stimulus conditions and five listeners are shown in Fig. 2. Every match is shown, with the five matches of a run connected by a line. The rising character for the lines shows the tendency for the matching frequency to increase with the edge frequency as expected from the monotonicity statistic. There is also a clear tendency for pitches to be above the edge frequency when the noise is coherent below the edge (circles), and for the pitches to be below the edge frequency when the noise is coherent above the edge (diamonds). These results are all in agreement with the original predictions based on the EC model.

Although the data as a whole all give clear evidence of the BICEP and its expected pitch shift, there are individual differences. Listeners M and W with the most musical training show the most consistent matches. Listener J with no musical training is the least consistent. Listeners C and T made more consistent matches to π-phase stimuli than to 0-phase stimuli. A model explanation for this result would be that listeners C and T are more successful in monitoring the binaural summation channel, and less successful in taking binaural differences (see Sec. E below).

E. Monaural analog

According to the EC model for BICEP, the listener binaurally subtracts or adds the noises in the two ears leading to a central spectrum with an edge. The monaural analog experiments performed noise subtraction or addition electronically and sent this noise to both ears diotically. For the purposes of this paper such a diotic noise is called ''monaural'' because each ear receives adequate information to make a pitch match. Otherwise, the monaural experiments were identical to the BICEP experiments above.

1. Noise subtraction experiment

Subtracting the two channels of the 0-phase BICEP stimulus leads to a monaural noise band with a sharp drop (measured to be 50 dB in our apparatus). If the components are coherent below the edge then the noise power vanishes below the edge and the monaural noise is high pass. If the coherent region lies above the edge then the monaural noise is low pass. Listeners C, J, and W matched the pitches of
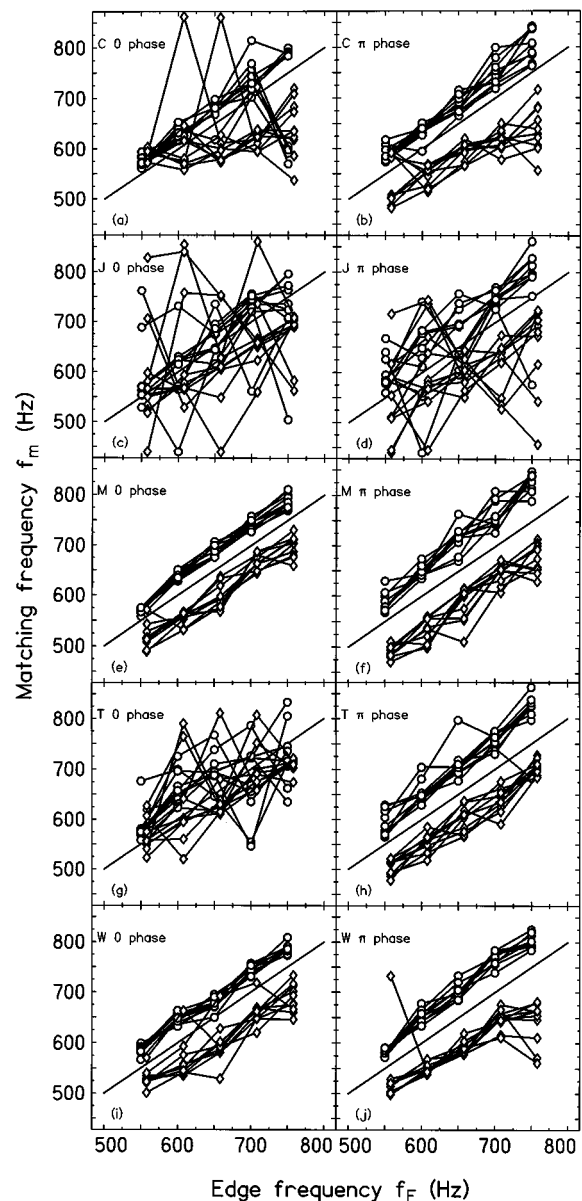

FIG. 2. Experiment 1 BICEP results. Circles show matching frequencies for BICEP noise that is coherent below the edge frequency. Diamonds show matching frequencies for BICEP noise that is coherent above the edge frequency. The diamonds have been slightly shifted to the right for clarity. The coherent noise components were identical in both ears (0 phase) or inverted (π phase). There were five listeners, C, J, M, T, and W.

these noise bands created by subtraction. The results are shown by the monotonicity statistic in rows 1 and 2 of Table III and in Figs. 3(a), (b), (c). It is clear that listeners matched the edge pitches well and that the shifts were similar, in

TABLE III. Monotonicity, for monaural edge pitch. Experiment 1. Monaural noise was made from the difference of the two BICEP channels or the sum of the two BICEP channels with zero-phase coherence above or below the edge. Results are averaged over three listeners, C, J, and W.

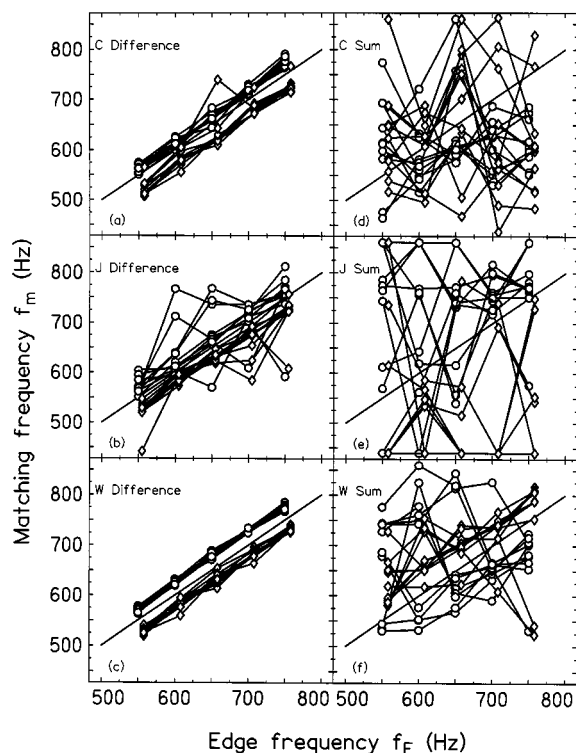| Condition	0	1	2	3	4	Deviations
Difference, coherent below	77	20	3	0	0	
Difference, coherent above	90	10	0	0	0	
Sum, coherent below	3	33	47	17	0	
Sum, coherent above	10	30	43	17	0	

FIG. 3. Experiment 1—Monaural analog to the BICEP. Left column (a), (b), (c): Pitch matches for the *difference* of channels. Circles are for noise coherent below the edge frequency so that the difference noise is high pass. Diamonds are for noise that is coherent above the edge frequency so that the difference noise is low pass. Right column (d), (e), (f): Pitch matches for the *sum* of channels. Circles are for noise that is coherent below the edge frequency so that the summed noise has a 3-dB low-boost shelf below the edge frequency. Diamonds are for noise that is coherent above the edge frequency so that the summed noise has a 3-dB high-boost shelf above the edge frequency.

magnitude and *sign*, to the shifts observed in the BICEP experiment. This agreement suggests a similarity in mechanism between binaural and monaural experiments.

The monaural analog is conceptually similar to the monaural experiment run by Klein and Hartmann in connection with the BEP. However, the comparison of pitch shifts with the BICEP is more impressive than with the BEP. The BICEP comparison shows the auditory system to be choosing binaural addition or subtraction operations to optimize the central edge in the BICEP experiment. In contrast, it is not possible to know what is the optimum operation in the BEP experiment.

2. Noise addition experiment

Adding the two channels of the 0-phase BICEP stimulus leads to a monaural noise with an edge that is only 3 dB high. For the coherent components the amplitudes reinforce in phase, whereas for the incoherent components only the powers add. The difference is 3 dB in favor of the coherent region. Listeners C, J, and W matched the pitches of noises created by noise addition. The results are shown by the monotonicity statistic in rows 3 and 4 of Table III and in Figs. 3(d), (e), (f). Comparison with Figs. 3(a), (b), (c) shows the advantage of cancellation over reinforcement. In fact,

matches to the monaural added stimulus are clearly less consistent than matches to the BICEP stimulus itself.

3. Implications for models for BICEP

In the context of the EC model, the results of the monaural experiments give every reason to believe that the optimum strategy for a binaural system attempting to fabricate a pitch is to cancel the coherent components and not to combine them so as to reinforce. That is the unambiguous binaural strategy elicited by the BICEP stimulus. For this reason it was possible to conclude above that listeners who favor coherent components in π phase, such as C and T, preferentially monitor a binaural addition channel and not a binaural subtraction channel.

III. EXPERIMENT 2—FREQUENCY LIMITS

Generally, binaural effects are strongest in the frequency region around 600 Hz (Licklider *et al.*, 1950). Binaural effects become weaker two octaves below this frequency and especially weaker two octaves above. The experiments of Sec. II were in the favorable frequency region near 600 Hz. The purpose of experiment 2 was to explore alternative frequency regions to determine the limits over which the BICEP could be heard and reliably matched.

A. Method

For the low-frequency range, edge frequencies were 100, 200, 300, 400, and 500 Hz. For the high-frequency range, edge frequencies were 800, 1000, 1200, 1400, and 1600 Hz. Except for the change in range, the experimental methods were identical to experiment 1. Equal numbers of runs were done with the coherent region above and below the edge frequency. Equal numbers of runs were done with the coherent noise in phase (0) and with phase inverted (π), a total of four different edge types. Listener C did ten runs for each edge type in the high-frequency region and four runs in the low-frequency region. Listener W did ten runs for each edge type in both the high-frequency and the low-frequency regions.

B. Results

The results for high- and low-frequency regions are shown in Fig. 4. This figure also includes the data from experiment 1 plotted in the midfrequency region. It is evident that performance was much less reliable outside the midfrequency region. Performance appears to be especially erratic in the high-frequency region, though this is partly the result of the linear frequency scale. In a log–log plot, not shown here, the low-and high-frequency regions appear with similar dispersion.

A second view of the data appears in Fig. 5, which plots the standard deviation of the matching frequencies as a percentage of the edge frequency. Although the standard deviation may be somewhat inflated by octave errors (visible at 100, 200, and 800 Hz in Fig. 4), it provides further data on the consistency of the matches.

Figures 4 and 5 show that matches to edges outside the midfrequency region were not always unreliable. In the low-
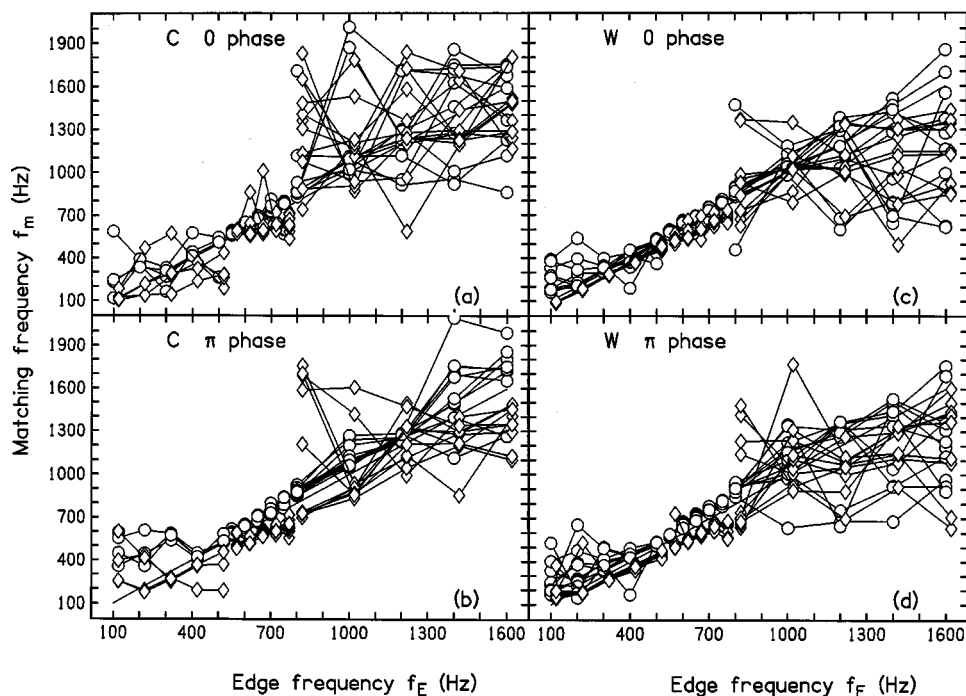
FIG. 4. Experiment 2—BICEP with extended frequency range. See the caption for Fig. 2. The data in the midfrequency range (five edge frequencies from 550 through 750 Hz) are from experiment 1.

frequency region matches by listener W to noise that was coherent *above* the edge were almost as self-consistent as matches in the midfrequency region. Listener W exhibited excellent consistency for the 0-phase coherent noise right down to 100 Hz. For listener C, data in the low-frequency region showed no clear preference for 0-phase or π-phase coherence.

Similarly, in the high-frequency region, matches to π-phase noise that was coherent *below* the edge were consistent at 800 Hz for listener W and for edge frequencies up to 1200 Hz for listener C. The preference shown by listener C for π phase in the midfrequency region reappeared in the high-frequency region.

Because the three different frequency regions of this ex-

periment were done separately, the data in Figs. 4 and 5 exhibit changes at the boundaries between regions that may be more pronounced than would have occurred had all frequency regions been represented in all runs. This apparent artifact was, however, not decisive. Consistent matches for coherence above a 300-Hz edge and for coherence below a 800-Hz edge were seen for both listeners. Evidently the inclusion of many difficult matches in a run did not seriously affect the ability to match at these edge frequencies in the low- and high-frequency regions. Therefore, it seems fair to present all the data in the same figures.

There are two conclusions that can be drawn from experiment 2. The first is that the BICEP is considerably more prominent in the narrow range of frequencies around 600 Hz,
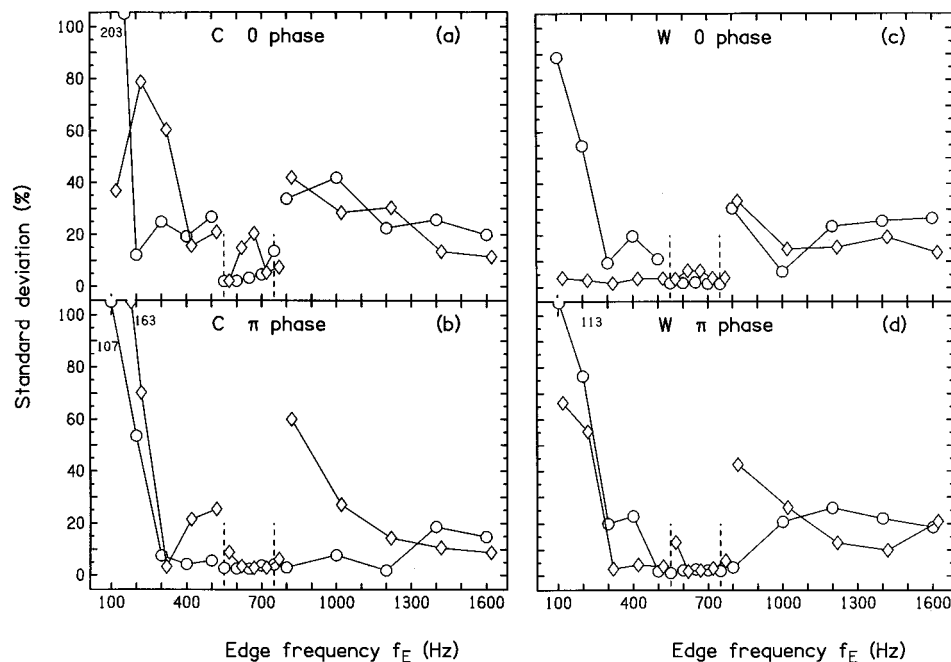

FIG. 5. Experiment 2—BICEP with extended frequency range. The ordinate shows the standard deviation ($N - 1$ weight) of the matching frequencies in Fig. 4. The standard deviation is given as a percentage of the edge frequency shown on the abscissa. A few standard deviations greater than 100 percent are plotted on the top line with actual values given numerically.

where synchrony is well represented in the binaural system. Although there are exceptions, one might generalize from the data and conclude that the BICEP exists for edge frequencies in a range from 300 to 1000 Hz. This range is smaller than the range for the BEP, which can be found as high as 2.5 kHz. In the BEP the noise is coherent on both sides of the edge.

The second conclusion from experiment 2 is that when listeners are required to match edge frequencies outside the midfrequency range then the matches are more successful when the coherent region stretches into the midfrequency range, i.e., coherent above the edge for low edge frequencies and coherent below the edge for high edge frequencies. Thus, coherence in the region of best binaural synchrony seems to be important for the BICEP.

IV. EXPERIMENT 3—TIME DELAY

A. BICEP

The binaural system can obtain a central spectrum with an edge starting with the dichotic stimuli of experiments 1 or 2 by binaurally adding or subtracting the noises to the two ears. According to the standard EC model, the binaural adding operation is accomplished by a 180-deg phase shift followed by a central subtracting process. Thus, there is no central adding process *per se*. The necessary phase shift is normally attributed to an interaural delay line. It is part of the equalization stage in the EC model.

An alternative version of the EC model (Green, 1966; Henning, 1973; Bilsen and Goldstein, 1974) includes a central adding process. No delay line is required for the addition operation. Whether the central process is adding, or subtracting, or possibly both, the internal auditory delay line can be tested by a BICEP stimulus where the noise to one ear is given an external delay. Experiment 3 introduced such an external delay.

An external delay line does not change the incoherence/coherence conditions of the stimulus. For the incoherent region, adding delay maintains the incoherence. For the coherent region, adding the delay means that the (positive or negative) peak of the cross-correlation function is shifted from a lag of zero to a lag equal to the added time delay. However, the magnitude of the peak is still unity.

Delay was added with a BSS TCS-803 digital delay. (Actual values of the delay were verified ± 10 μs by a test that canceled sine tones.) The experiment was otherwise similar to experiment 1 including coherence both above and below the edge. However, the initial phase was always set to zero because it seemed superfluous to introduce a phase shift, such as π, in addition to the delay.

The results of the experiment for listeners C, M, and W are given by the monotonicity statistic in Table IV and the pitch shift plots in Fig. 6. Unlike Figs. 2, 3, and 4 which show actual matching frequencies, Fig. 6 plots each match in terms of its percentage deviation from the coherence edge frequency. Therefore, the 45-deg line in Figs. 2, 3, and 4 corresponds to the horizontal line at zero percent in Fig. 6.

The data indicate that delays as long as 1 ms have little effect on the matches except for a slight increase in variabil-

TABLE IV. Monotonicity for BICEP, delay in one channel. Experiment 3. Results are averaged over three listeners, C, M, and W.

Delay (ms)	0	1	2	3	4	Deviations
Coherent below						
0.0	92	8	0	0	0	
0.5	75	17	8	0	0	
1.0	50	50	0	0	0	
2.0	33	50	17	0	0	
3.0	17	25	33	25	0	
4.0	0	25	58	17	0	
Coherent above						
0.0	58	17	17	8	0	
0.5	50	50	0	0	0	
1.0	50	25	25	0	0	
2.0	8	33	59	0	0	
3.0	0	33	50	17	0	
4.0	0	16	42	42	0	

ity. A delay of 2 ms sometimes disrupts pitch matching but usually does not. A delay of 3 ms produces a more dramatic increase in variability, and a delay of 4 ms increases the trends seen at 3 ms. Delays of 3- and 4 ms lead to a considerable pitch shift, in the direction of lower pitch, for listeners C and W, but not for listener M.

The heavy solid lines in the top panels of Fig. 6 form a
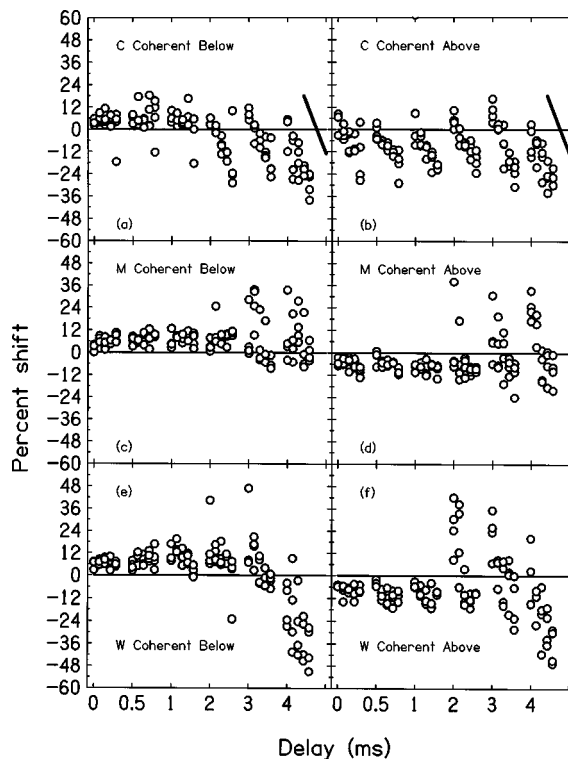

FIG. 6. Experiment 3—BICEP with one channel delayed as shown on the horizontal axis. All matches are shown. There are six blocks of data, corresponding to the six delays shown on the x axis. Within each block, columns, from left to right, indicate edge frequencies of 550, 600, 650, 700, and 750 Hz, the central range. Separate plots show data for stimuli that are coherent below and coherent above the edge. There were three listeners C, M, and W. The heavy lines in the top panels show the average shifts that would occur for random matches with a mean matching frequency of 650 Hz. See the text.

TABLE V. Monotonicity for the disruptive delay. Experiment 3. Results are averaged over three listeners, C, M, and W.

Condition	0	1	2	3	4	Deviations
Sum	0	22	66	12	0	
Difference	0	22	44	34	0	
BICEP	44	44	12	0	0	

reference showing the result expected for random matching with a mean matching frequency of 650 Hz, the center of the stimulus edge-frequency range. The vertical extent of the heavy line, caused by the increasing edge frequency from left to right, is less than the extent seen experimentally for long decays. The slopes of the heavy lines are similar to the slopes of the data for long delays, indicating weak correlation between matching frequency and edge frequency.

What seems most significant about this experiment is that it shows the internal delay line of the binaural system at work. The external delay can be compensated by the internal delay line so that the central spectrum has an edge, leading to a pitch, unchanged from the experiment with no external delay. This demonstration is independent of assumptions about the rule for combination, central addition, or central subtraction. The delay line begins to fail to operate in the expected way somewhere between 2 and 4 ms. This result agrees with the length of the delay line determined by binaural masking experiments (e.g., van der Heijden and Trahiotis, 1999).

B. Monaural analog

A monaural analog to the BICEP stimulus with delay adds the signals for the left and right ears. This leads to a flat noise spectrum in the incoherent region and rippled noise in the coherent region. It can be expected that listeners would have great difficulty in matching the pitch of such a noise. Unlike the binaural case, there is no way to use an internal delay line to compensate for the effects of the external delay. Particularly disruptive is a delay equal to one quarter of the reciprocal of the edge frequency.[5] For example, for an edge frequency of 650 Hz the reciprocal is 1538 μs, and the most disruptive delay is 385 μs.

1. Experiment and results

Listeners C, M, and W did three runs for the sum and three runs for the difference of the monaural delayed task with a delay of 385 μs. This delay is expected to be maximally disruptive because one quarter of its inverse is in the center of our range of edge frequencies. As always, listeners were required to make a match whether or not they could hear a clear pitch. The experiment found that sum and difference stimuli led to indistinguishable results, and both were random. The monotonicity statistic in Table V and pitch-matching plots (not shown) for sum or difference are consistent with random matches. By contrast, Table V and pitch-matching plots show monotonic or nearly monotonic matches for the BICEP with a delay of 385 μs, not different from the results with zero delay. The comparison between binaural and monaural experiments gave clear evidence for the operation of the internal delay line.
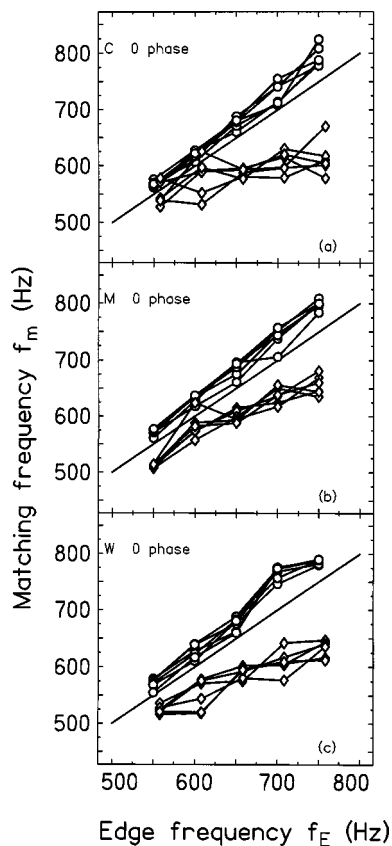

FIG. 7. Experiment 4—Matches for BICEP with random amplitudes—*identical* in both ears. See Fig. 2.

V. EXPERIMENT 4—RANDOM AMPLITUDES

Experiments 1, 2, and 3 used special noise in which every component had the same amplitude, i.e., equal-amplitude random-phase noise. However, for thermal noise the amplitudes are distributed according to a Rayleigh distribution. Experiment 4 replaced the equal-amplitude noise with such a random-amplitude noise. The conditions were otherwise the same as for experiment 1. In particular, the amplitudes of each component were the same in left and right ears. Therefore, in the phase-coherent region a subtraction operation would lead to perfect cancellation as in experiment 1. Because the spectrum is so dense, with many components in any auditory band, one expects that the change from equal amplitudes to random amplitudes should have no effect (Hartmann, 1997, p. 526). Experiment 4 was a brief test to show that this is so.

A. Results

Listeners C, M, and W each did six runs with the BICEP stimulus made with random amplitudes, both coherent below the edge and coherent above. Only the 0-phase condition was tested. The results, shown in Fig. 7, were indistinguishable from the results of experiment 1 with equal amplitudes. The monotonicity statistic in Table VI shows no degradation compared to Table II for equal amplitudes. Both tables show somewhat greater monotonicity for the ''coherent below'' condition. Although the listener populations were somewhat different, the comparison between experiments 1 and 4 gives no reason to suspect that the special choice of equal-

TABLE VI. Monotonicity for BICEP, random amplitudes. Experiment 4. Results are averaged over three listeners, C, M, and W.

Condition	0	1	2	3	4	Deviations
Zero phase, coherent below	100	0	0	0	0	
Zero phase, coherent above	50	33	17	0	0	

TABLE VII. Monotonicity for independent amplitudes. Experiment 5. (1) Monaural noise was made from the difference of the two zero-phase BICEP channels. (2) Monaural noise was made from the sum of the two zero-phase BICEP channels. (3) The BICEP dichotic stimulus. All were coherent below the edge or coherent above. Results are averaged over three listeners, C, M, and W.

Condition	0	1	2	3	4	Deviations
1 difference, coherent below	75	25	0	0	0	
1 difference, coherent above	75	25	0	0	0	
2 sum, coherent below	33	42	25	0	0	
2 sum, coherent above	8	75	8	0	8	
3 BICEP, coherent below	25	25	50	0	0	
3 BICEP, coherent above	9	58	33	0	0	

amplitude noise, as used in experiment 1 and in other experiments, had an important effect on the BICEP.

VI. EXPERIMENT 5—INTERAURALLY DIFFERENT AMPLITUDES

Experiment 4 used Rayleigh amplitudes that were random but were the same in left and right ears. Experiment 5 also chose amplitudes from a Rayleigh distribution, but the choices for left and right ears were made independently. Therefore, in the coherent region, only the phases of components in the left and right ears were identical. In the incoherent region both amplitudes and phases were independent across the ears. The reason for expecting that a BICEP would survive this amplitude randomization can be found in the monaural analog.

A. Monaural analog

Thermal noise is made by adding sine and cosine components with coefficients drawn from a normal distribution. If the normal distribution has variance σ^2 the amplitudes of the spectral components are Rayleigh distributed, and the noise power is half the average value of the squared amplitude or σ^2. Adding or subtracting two statistically identical but incoherent noises leads to twice the power or $2\sigma^2$. In a coherent region, where the noise phases in the two channels are the same, adding (or subtracting) the channels leads to a power given by the summed incoherent power plus (or minus) the square of the average amplitude. For a Rayleigh distribution, this is $2\sigma^2 \pm (\pi/2)\sigma^2$ (Hartmann, 1997, Appendix I). For the plus sign, the noise is 2.5 dB greater than the sum or difference of incoherent noise channels. For the minus sign, the noise is 6.7 dB less than the sum or difference of incoherent noise channels. Therefore, adding or subtracting channels leads to a spectral edge of 2.5 or 6.7 dB, respectively. It seems possible that a reliable pitch might be generated by spectral edges of this kind, especially the 6.7-dB edge created by the difference.

Listeners C, M, and W did four runs of the monaural experiment. The results are given by the monotonicity statistic in the first four lines of Table VII and by the matching graphs in Figs. 8(a)–(f). As expected, the most consistent matching occurred for the edge created by the difference, but the edge created by the sum also led to matches much better than chance.

B. Binaural experiment

Based on the good results for the difference spectrum in the monaural analog, one would expect that listeners could hear a BICEP for left and right channels made from independent noise amplitudes.

Listeners C, M, and W did four runs of this binaural experiment. The results are shown in the monotonicity table, Table VII—last two lines—and by Figs. 8(g)–(i). Evidently the matching performance did not rise to the level expected from the results of the monaural analog. Only listener M approached the consistency achieved with identical amplitudes in the two ears.

The poor performance with interaurally independent amplitudes is somewhat surprising. Because amplitudes are strictly positive numbers, randomizing them does not have the same effect as randomizing phases. With an identical phase spectrum in the two ears, considerable coherence remains even when the amplitudes are randomized. Therefore, good performance was expected, not poor performance. The stimulus for this experiment can be thought of as an identical-amplitude stimulus (as in experiment 4) plus added incoherent noise, albeit a rather special noise. The poor performance suggests that the BICEP is quite vulnerable to added binaurally incoherent noise.
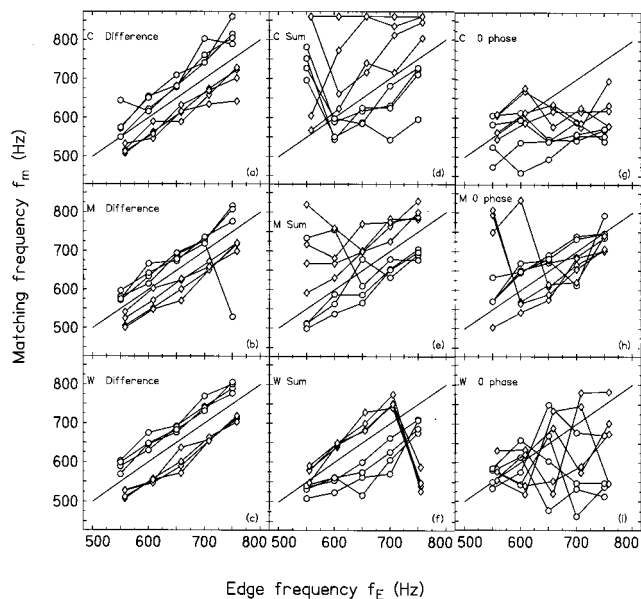

FIG. 8. Experiment 5—Pitch matches for noise with random amplitudes—*independent* in both ears. Circles are for noise coherent below the edge; diamonds are for noise coherent above the edge. Separate columns show the results for (1) Monaural experiment—difference of channels; (2) Monaural experiment—sum of channels; (3) BICEP experiment with coherent components having identical phases.

VII. SUMMARY AND COMPARISON

The most significant results of this work can be summarized in a few sentences: (1) The BICEP exists. (2) Matches to the BICEP show pitch shifts as expected based on BEP experiments and based on the model of central lateral inhibition advanced to explain the BEP (Klein and Hartmann, 1980). Therefore, BICEP and BEP are similar, but BICEP has the great advantage that the pitch shifts can be predicted unambiguously, whereas the pitch shifts for the BEP depend on a choice made by the listener's binaural system. (3) BICEP shifts agree qualitatively with the pitch shifts seen in analogous monaural experiments, concocted to simulate the imagined operation of a hypothetical EC process. Quantitatively the BICEP shifts are somewhat larger and show larger variance. (4) The BICEP continues to exist if the noise to one of the ears is externally delayed, but it becomes unreliable if the delay exceeds 2 or 3 ms. (5) The comparison of π-phase coherence and 0-phase coherence reveals a curious individual difference. Of four listeners who made very reliable matches to π-phase BICEP, only two of them retained that reliability for 0-phase BICEP.

It is natural to compare the BICEP with the BEP. The monaural analogs are essentially identical for both. However, the BEP is coherent on both sides of the edge, whereas BICEP is coherent on only one side. Therefore, the boundaries on a binaural delay-place (ITD-frequency) plot are sharper for the BEP, and, *a priori*, one might expect the BICEP to be weaker than the BEP. Listeners made informal comparisons between the BEP and BICEP in the middle range of edge frequencies, 550–750 Hz. Some of them found the BEP to be stronger and some of them did not. None of them found the difference to be striking. Apparently, what is important is the ability to cancel a coherent spectral region. The spectral region that is not canceled (coherent for BEP and incoherent for BICEP) is apparently less important, at least for favorable values of the edge frequency. It may become more important for edge frequencies that are higher than optimum. Our best evidence is that the BICEP exists over a smaller range of edge frequencies (300 to 1000 Hz) than the BEP, which persists for edges an octave higher. However, a careful head-to-head comparison between BEP and BICEP has not been done.

VIII. DISCUSSION AND CONCLUSION

Besides the EC model, there are other models that attempt to account for dichotic pitch effects like the Huggins pitch, the BEP, and the BICEP. The central activity pattern (CAP) model by Raatgever and Bilsen (1977, 1986), based on the central spectrum model (Bilsen, 1977), is a delay-place plane model for both the pitch sensation and the lateralization of the pitch. The CAP model predicts that dichotic pitches occur at the frequencies of spectral peaks prominent in the central spectrum obtained with advantageous interaural time delays. Recently Culling *et al.* (1998a, b) have developed a modified equalization-cancellation model (mEC) wherein the EC process takes place in frequency bands. The mEC model predicts binaural pitches at frequencies where the interaural phase changes so abruptly with frequency that

there is a residual excitation in a filter channel following the cancellation process. These models can be compared with what we have learned about the BICEP.

On the existence of the BICEP: All three models, EC, CAP, and mEC, as they are applied to binaural pitch effects, predict that a binaural pitch occurs at a place corresponding to a peak in a central representation resembling a central spectrum. However, none of the models actually predicts a peak for the BICEP stimulus, and the most literal interpretation of the situation is that none of the models predicts the existence of the BICEP.

The mEC model has the virtue that it *does* predict a peak for the BEP stimulus. The peak occurs because auditory filters tuned to the edge frequency, where the phase changes rapidly, are the only filters in which good cancellation does not occur. Hence, a peak appears as function of place. By comparison, the original EC model predicts only an edge in the central spectrum. On the other hand, as noted by Culling *et al.*, the mEC model does not obtain a peak for the BICEP stimulus. It predicts only a broadened edge in the residual activation pattern, equivalent to the central spectrum. It was actually this situation that prompted us to resume BICEP experiments after a hiatus of 15 years.

But, although none of the models in their present form predicts a peak in the central spectrum, all of them can predict an edge. It may be reasonably argued that the remaining question is why an edge, either monaural or central, leads to a sense of pitch. Incorporating a central differentiation process with respect to place could account for both monaural and central pitches and would allow any of the binaural models to predict the existence of the BICEP. Differentiation with respect to place is here equivalent to lateral inhibition, enhancing contrast at an edge. The details of how central lateral inhibition could be incorporated into each of the models of dichotic pitch is beyond the scope of this article.

On the pitch shift of the BICEP: The experiments of this article have shown that the BICEP is shifted away from the edge frequency. The results are unequivocal on that matter. We suspect that the BEP is similarly shifted, but it is harder to prove that. Experiments on BEP made with analog stimuli by Frijns *et al.* (1986) did not find pitch shifts, but continued BEP experiments with digital stimuli in our lab continued to find shifts. The unambiguous shifts observed with the BICEP make the otherwise controversial BEP shifts more plausible.

None of the models of dichotic pitch predict the shift of the BEP. However, incorporating central lateral inhibition would predict a shift for both the BEP and the BICEP. Thus, central lateral inhibition not only allows the binaural models to predict the existence of the BEP and BICEP but it also predicts the observed shift. The EC model and the mEC model together with central lateral inhibition predict that the BICEP should be shifted into the incoherent region, in agreement with experiment. Against the hypothesized central lateral inhibition hypothesis are the results of pulsation threshold experiments (Hartmann, 1984b) which were designed to look for the effect but failed to find it. Possibly central lateral inhibition exists but has temporal properties that prevent a pulsation threshold experiment from revealing it. Possibly central lateral inhibition exists but it is such a weak effect

that it can be seen only in a pitch experiment and not in a masking (pulsation threshold) experiment.

On the advantage of π-phase coherence over 0-phase coherence: Figure 2 shows that four listeners made consistent matches to the π-phase condition but only two of them made comparably consistent matches to the 0-phase condition. The EC model and mEC model involve cancellation by central subtraction of left- and right-ear noises following an equalizing delay. As such, they should predict better performance when the coherent region has 0 phase rather than π phase because no internal delay can perfectly cancel all π-phase components within one auditory channel. The fact that two listeners performed better in experiment 1 when the coherent region was in π phase runs counter to the predictions of the EC and mEC models.

By contrast, the CAP model employs a central addition process (rather than subtraction). The apparent advantage of π-phase coherence seen experimentally would be consistent with addition as the fundamental operation because a central addition process can cancel the π-phase noise without the aid of a delay line. This result argues in favor of the CAP model. However, this argument is not consistent with the low-frequency data obtained in experiment 2. If it is really true that the π-phase advantage is caused by the delay line needed for 0-phase coherence then this advantage ought to grow as the edge frequency decreases. The lower the edge frequency the longer the required delay. Experiment 2 includes only two listeners, but both of them produce more consistent matches for 0 phase than for π phase. One of these two listeners is C, who shows an advantage for π phase at higher frequencies.

An additional clue appears in the form of the BICEP data in Fig. 2. For listeners M and W, the good performance seen with 0-phase coherence is not distinguishable from the good performance with π phase. For listeners C and T, performance is considerably worse with 0 phase. This form of the data does not suggest a central addition process that is made noisy by a delay line when the system is required to subtract. Instead, the form of the data suggests a model with two channels, one for central addition and one for central subtraction as in the version of the EC model adopted by Green (1966) and Henning (1973). Listeners may be more successful at monitoring one of these channels. Alternatively, the two channels might correspond to different paths through the delay-place plane as suggested in the CAP model. In the end, the BICEP experiments presented in this article have not enabled us to choose among competing models of dichotic pitch; the results have indicated that all the models require extension, if only to deal with the pitch shifts.

ACKNOWLEDGMENTS

[1]In this article, the term ''addition'' is used in several ways. ''Binaural addition'' means that the auditory system as a whole operates in a way that adds the signals at the two ears. Binaural addition might possibly employ a binaural delay line to shift the phase. ''Central addition'' means that elemental rule of binaural combination is addition. By this definition the combination takes place at a stage that follows all interaural delays. ''Noise addition'' means that the experimenters have added the left and right channels to make a monaural stimulus, intended as a monaural analog to the binaural processes under study. In this article, the term ''subtraction'' is qualified in the same way as addition.

[2]The correspondence between dichotic pitch and the monaural analogy obtained by electronically simulating the operation of the binaural system does not hold perfectly for the BEP. Unpublished experiments show that when the phase transition region is broadened, the BEP survives better than the monaural analog. Although the monaural analog leads to a stronger pitch sensation than the BEP for a narrow transition region, the BEP can become stronger for wider regions (Hartmann, 1984c).

[3]Early experiments used gap intervals consisting of diotic noise. This procedure was based on the assumption that it would be easier to hear the dichotic pitch if the dichotic stimulus entered as an abrupt change in the noise as the stimulus sequence cycled. A similar temporal effect was found by Kubovy (1981) for line spectra. Experiments with a silent gap were introduced to prove that the BICEP was *not* really caused by Kubovy's effect. At that point listeners discovered that silent gaps made the BICEP easier to hear and that noise in the gaps just made the task harder. Because noise in the gaps leads to a potential Kubovy artifact, no further experiments were done with noise in the gaps, and all data reported in this article were collected with silent gaps.

[4]Expanded tables with the data from individual listeners are available from the first author. Please request Report 110SUP.

[5]If the power spectrum of the sum (or difference) of two channels is said to be 1 for the incoherent region, then in the coherent region the power spectrum of the sum of undelayed channels is 2 (because of amplitude addition), and the power spectrum of the difference is 0 (because of perfect cancellation). Delaying one of the channels by T_D before combination produces a power spectrum in the coherent region given by $2 \cos^2(\pi f T_D)$ for the sum and given by $2 \sin^2(\pi f T_D)$ for the difference (Hartmann, 1997, Chap. 15). When fT_D equals 1/4, these functions are both equal to 1. Therefore, if T_D is chosen according to the edge frequency, f_E, such that $f_E T_D = 1/4$ then the delay is maximally disruptive because at the edge, the power spectrum in the coherent region is exactly equal to the value of the power spectrum in the incoherent region. The spectrum in the coherent region now fits smoothly onto the spectrum for the incoherent region and there is no discontinuity, no spectral edge at all.

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

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

Cramer, E. M., and Huggins, W. H. (**1958**). ''Creation of pitch through binaural interaction,'' J. Acoust. Soc. Am. **30**, 413–417.

Culling, J. F., Summerfield, Q., and Marshall, D. H. (**1998a**). ''Dichotic pitches as illusions of binaural unmasking. I. Huggins' pitch and the 'binaural edge pitch','' J. Acoust. Soc. Am. **103**, 3509–3526.

Culling, J. F., Marshall, D. H., and Summerfield, Q. (**1998b**). ''Dichotic pitches as illusions of binaural unmasking. II. The Fourcin pitch and the dichotic repetition pitch,'' J. Acoust. Soc. Am. **103**, 3527–3540.

Durlach, N. I. (**1972**). ''Binaural signal detection—equalization and cancellation theory,'' in *Foundations of Modern Auditory Theory*, edited by J. Tobias (Academic, New York), Vol. 2, pp. 369–462.

Durlach, N. I. (**1962**). ''Note on the creation of pitch through binaural interaction,'' J. Acoust. Soc. Am. **34**, 1096–1099.

Fastl, H. (**1971**). ''Uber Tonhohenmpfindungen bei Rauschen,'' Acustica **25**, 350–354.

Frijns, J. H. M., Raatgever, J., and Bilsen, F. A. (**1986**). ''A central spectrum theory of binaural processing. The binaural edge pitch revisited,'' J. Acoust. Soc. Am. **80**, 442–451.

Green, D. M. (**1966**). ''Signal detection analysis of equalization and cancellation model,'' J. Acoust. Soc. Am. **40**, 833–838.

Hartmann, W. M. (**1979**). ''Some psychoacoustical experiments with all-pass networks,'' Am. J. Phys. **47**, 29–34.

Hartmann, W. M. (**1984a**). ''Binaural coherence edge pitch,'' J. Acoust. Soc. Am. Suppl. 1**75**, S22.

Hartmann, W. M. (**1984b**). ''A search for central lateral inhibition,'' J. Acoust. Soc. Am. **75**, 528–535.

Hartmann, W. M. (**1984c**). ''The dependence of noise edge pitch and binaural edge pitch on the frequency width of the edge,'' J. Acoust. Soc. Am. Suppl. 1 **75**, S22.

Hartmann, W. M. (**1997**). *Signals, Sound, and Sensation* (AIP, Springer, New York).

Henning, G. B. (**1973**). ''Effect of interaural phase on frequency and amplitude discrimination,'' J. Acoust. Soc. Am. **54**, 1160–1178.

Klein, M. A., and Hartmann, W. M. (**1980**). ''Binaural edge pitch,'' J. Acoust. Soc. Am. **70**, 51–61.

Kubovy, M. (**1981**). ''Concurrent-pitch segregation and the theory of indispensable attributes,'' in *Perceptual Organization*, edited by M. Kubovy and J. R. Pomerantz (Erlbaum, Hillsdale, NJ).

Licklider, J. C. R., Webster, J. C., and Hedlum, J. M. (**1950**). ''On the frequency limits of binaural beats,'' J. Acoust. Soc. Am. **22**, 478–483.

Raatgever, J., and Bilsen, F. A. (**1977**). ''Lateralization and dichotic pitch a a result of spectral pattern recognition,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 443–453.

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

Small, A. M., and Daniloff, R. G. (**1967**). ''Pitch of noise bands,'' J. Acoust. Soc. Am. **41**, 506–512.

van der Heijden, M., and Trahiotis, C. (**1999**). ''Masking with interaurally delayed stimuli: The use of internal delays in binaural detection,'' J. Acoust. Soc. Am. **105**, 388–399.