Interaural fluctuations and the detection of interaural incoherence: Bandwidth effects

Matthew J. Goupell[a] and William M. Hartmann
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

(Received 2 May 2005; revised 28 March 2006; accepted 5 April 2006)

One-hundred left-right noise-pairs were generated, all with a fixed value of long-term interaural coherence, 0.9922. The noises had a center frequency of 500 Hz, a bandwidth of 14 Hz, and a duration of 500 ms. Listeners were required to discriminate between these slightly incoherent noises and diotic noises, with a coherence of 1.0. It was found that the value of interaural coherence itself was an inadequate predictor of discrimination. Instead, incoherence was much more readily detected for those noise-pairs with the largest fluctuations in interaural phase or level differences (as measured by the standard deviations). One-hundred noise-pairs with the same value of coherence, 0.9922, and geometric mean frequency of 500 Hz were also generated for bandwidths of 108 and 2394 Hz. It was found that for increasing bandwidth, fluctuations in interaural differences varied less between different noise-pairs and that detection performance varied less as well. The results suggest that incoherence detection is based on the size of interaural fluctuations and that the value of coherence itself predicts performance only in the wideband limit. © *2006 Acoustical Society of America.* [DOI: 10.1121/1.2200147]

PACS number(s): 43.66.Pn, 43.66.Ba, 43.66.Qp [AK] Pages: 3971–3986

I. INTRODUCTION

Interaural coherence is a measure of the similarity of signals in a listener's two ears. It is derived from the interaural cross-correlation function, $\gamma(\delta)$, which is a function of the interaural lag δ,

$$\gamma(\delta) = \frac{\int_0^T x_L(t) x_R(t+\delta) dt}{\sqrt{\int_0^T x_L^2(t_1) dt_1 \int_0^T x_R^2(t_2) dt_2}}, \tag{1}$$

where x_R is the signal in the right ear, and x_L is the signal in the left. The cross-correlation is bounded by $-1 \leqslant \gamma \leqslant 1$.

With respect to perception, interest normally centers on the peak of $\gamma(\delta)$. The value of δ for which the peak occurs is the relevant interaural time difference (ITD) cue for the location of the sound image. This value of δ was given a place representation in the famous binaural model by Jeffress (1948). The height of the peak is thought to determine the compactness of the image. If the sounds in the two ears are identical except for an interaural delay and/or a fixed level difference then γ assumes its maximum value of 1, and the image is expected to be maximally compact. If the height of the peak is less than 1, the image is broader or more diffuse (Barron, 1983; Blauert and Lindemann, 1986).

The height of the cross-correlation peak is known as the coherence. To be psychologically relevant, the peak must occur in a range of lag values, δ, that the binaural system can use for localization, and this requirement limits the number of peaks in the cross-correlation function for which the concept of coherence is meaningful, though the quantitative limits are somewhat uncertain. As applied in architectural acoustics, the coherence is called the IACC (interaural cross-correlation), and it refers to the height of $\gamma(\delta)$ for δ in the range $-1 \leqslant \delta \leqslant 1$ ms (Beranek, 2004). In recent years this objective architectural measure has been divided into two measures. One is the apparent source width (ASW), based on coherence within 80 ms of the onset of a sound (Barron and Marshall, 1981). The other is listener envelopment (LEV), determined by the coherence of later arriving sound, as measured after 80 ms (Bradley and Soulodre, 1995; Barron, 2004). Together, the ASW and LEV greatly influence the spatial impression of a sound in a room. Normally, the architectural measurements are made with directional microphone techniques and not with artificial heads.

Perceptual aspects of cross-correlation and interaural coherence have been studied by psychoacousticians, usually with bandpass filtered noise as a stimulus. Noise provides an abstraction of real-world sounds that is devoid of meaningful information, and it affords many opportunities for parametric variations. Using broadband noise, Pollack and Trittipoe (1959a, b) found thresholds for changes in cross-correlation as a function of the reference correlation, either 1.0 (i.e., No) or −1.0 (i.e., Nπ). They explored the effects of duration, sound level, frequency range, and interaural level difference.

Listeners are particularly sensitive to deviations from a reference correlation of 1.0. Using narrowband noise, Gabriel and Colburn (1981) found that listeners could easily distinguish between noise with a coherence of 1.0 and noise with a coherence of 0.99. They also reported the somewhat counterintuitive result that as the bandwidth of the noise increases, the just-noticeable difference (jnd) also increases. One might have expected the jnd to decrease instead given that a wider bandwidth generally offers the listener more information.

[a]Electronic mail: goupellm@msu.edu

A reference coherence of 1.0 is also of interest in connection with the masking level difference (MLD). Wilbanks and Whitmore (1967) and Koehnke *et al.* (1986) concluded that the threshold signal-to-noise ratio in the NoSπ condition is essentially determined by the ability to detect the incoherence introduced by the out-of-phase signal.

The present article is also concerned with incoherence detection starting with perfectly coherent noise as a reference. Its working hypothesis is based on the suspicion that the extreme sensitivity shown by listeners to small amounts of incoherence is not properly described by the long-term cross-correlation function. The reason is that when a small amount of incoherence is added to an otherwise perfectly coherent noise, the image of the noise acquires lateral fluctuations that are not present when the coherence is 1.0. The hypothesis continues with the observation that whereas long-term coherence is a measure that is averaged over time, the fluctuations that are hypothesized to be at the basis of coherence discrimination are dynamic. Thus, although the coherence measure may be a mathematically useful characterization of the similarity or dissimilarity of signals in the two ears, this measure may not be the most perceptually relevant characterization. Instead, it is possible that some measure that specifically considers fluctuations is better. A similar point of view with respect to the MLD was taken by Breebaart and Kohlrausch (2001) and in unpublished work by Isabelle and Colburn (2003). The rest of this article describes experiments that test this hypothesis.

II. EXPERIMENT 1: INCOHERENCE AND INTERAURAL FLUCTUATIONS

The purpose of Experiment 1 was to determine whether the height of the peak of the cross-correlation function (the coherence) adequately describes incoherence detection given a reference coherence of 1.0. The experiments employed reproducible noises, as have been used in MLD experiments by Gilkey *et al.* (1985), Isabelle and Colburn (1991), and Evilsizer *et al.* (2002). The noise stimuli, consisting of left- and right-ear channels, are referred to as "noise-pairs." In Experiment 1, the different noise-pairs all had the same value of long-term interaural coherence. If the coherence is an adequate measure of perception, then all the noise-pairs will have approximately the same detectability. That is, all pairs will be about equally distinguishable from perfectly coherent noise.

A. Stimuli

A collection of 100 noise-pairs with reproducible amplitudes and phases was created for Experiment 1. The process began with two waveforms, A and B, written as a sum of cosines in the form

$$x_A(t) = \sum_{n=1}^{N} C_n^A \cos(\omega_n t + \phi_n^A) \quad (2)$$

and

$$x_B(t) = \sum_{n=1}^{N} C_n^B \cos(\omega_n t + \phi_n^B), \quad (3)$$

where C and ϕ are the amplitudes and phases of the spectral components. The narrowband noises were generated with components having random phases over a frequency range of 490–510 Hz and with a frequency spacing of 2 Hz. Components between 495 and 505 Hz had equal amplitudes of unity. Frequencies below 495 Hz had a raised-cosine window applied to the amplitude spectrum of the form

$$C_n^A = C_n^B = \sin^2\left[\frac{\pi(f_n - 490)}{10}\right] \quad \text{for } 490 \leqslant f_n \leqslant 495 \quad (4)$$

to minimize spectral edge effects in the noises. The amplitudes of components from 505 to 510 Hz were similarly windowed. The 3-dB bandwidth of the noise was therefore 14 Hz.

For each noise in a collection of 100 reproducible noise-pairs, the B noise (x_B) was orthogonalized to the A noise (x_A) by the Gram-Schmidt orthogonalization procedure. The orthogonalized B noise is here denoted as x_B'. This was done to ensure that signals were uncorrelated and that the final value of the cross-correlation after mixing would be precise. The orthogonalization procedure is outlined by Culling *et al.* (2001).[1]

The two perfectly uncorrelated noises were then mixed, with mixing factor α, to create the final left and right (L and R) noise-pairs to be sent to the listeners,

$$x_L = x_A, \qquad x_R = \sqrt{(1-\alpha^2)}x_A + \alpha x_B'. \quad (5)$$

The coherence, as defined by Eq. (1), is

$$\rho = \sqrt{1-\alpha^2}. \quad (6)$$

Because the mixing factor used in all of the experiments was α=0.125, the interaural coherence of all the noise-pairs was ρ=0.9922.

After mixing, each noise was given a time interval shape with a total duration of 500 ms. A temporal window, $s(t)$, was applied such that there were raised-cosine edges with rise/fall times of 30 ms and a full-on duration of 440 ms. Consequently, the stimuli were made with both spectral and temporal windows. Both had raised-cosine edges. The application of the temporal window could change the coherence of the noise; therefore, the value of the computed stimulus waveform coherence was measured after the window was applied so that ρ=0.9922±0.0001. Noise-pairs that did not meet this criterion were rejected. For the 100 noise-pairs accepted in this experiment, 875 were rejected.

It should be noted that the coherence value of 0.9922 is a long-term value computed over the 500-ms stimulus. It differs from the short-term coherence—typically averaged over tens of milliseconds—in some models of binaural detection (e.g., Kollmeier and Gilkey, 1990). The long-term coherence is the relevant *stimulus* cross-correlation in incoherence detection experiments, such as those by Gabriel and

Colburn (1981), and in MLD experiments that are described in terms of signal-to-noise ratios.

To determine the time-dependent interaural phase difference (IPD) and interaural level difference (ILD), the analytic signals were found. By eliminating the negative frequencies, the analytic signal for either x_L or x_R is

$$\hat{x}(t) = s(t)\sum_{n=1}^{N} C_n \exp[i(\omega_n t + \phi_n)], \tag{7}$$

where $s(t)$ is the temporal window, C_n and ϕ_n are left or right amplitudes and phases as required, and i is the square root of -1.

By Euler's relation, the analytic signal becomes

$$\hat{x}(t) = s(t)\sum_{n=1}^{N} C_n[\cos(\omega_n t + \phi_n) + i \sin(\omega_n t + \phi_n)]$$
$$= F(t) + iG(t). \tag{8}$$

The phase and envelope of the analytic signal as functions of time are

$$\phi(t) = \arg[G(t), F(t)] \tag{9}$$

and

$$E(t) = \sqrt{F^2(t) + G^2(t)}, \tag{10}$$

where the arg function is the arctangent with possible quadrant correction. The time-dependent IPD (rad) and ILD (dB) of the analytic signal are then defined as

$$\Delta\Phi(t) = \phi_R(t) - \phi_L(t) \tag{11}$$

and

$$\Delta L(t) = 20 \log_{10}\left[\frac{E_R(t)}{E_L(t)}\right]. \tag{12}$$

Equation (11) yields a positive value of $\Delta\Phi(t)$ for signals that lead in the right ear. Similarly, Eq. (12) gives a positive value of $\Delta L(t)$ for a signal that has a larger level in the right ear. The interaural phase, $\Delta\Phi(t)$, was required to remain in the range between $-\pi$ and π rad at every point in time, and was corrected by adding or subtracting 2π when necessary. Because of the narrow bandwidth, the IPD is proportional to the ITD, which can be calculated by dividing the IPD by the center angular frequency of $2\pi \times 500$.

B. Signals

The IPD fluctuation over time is defined as

$$s_t[\Delta\Phi] = \sqrt{\frac{1}{T}\int_0^T [\Delta\Phi(t) - \overline{\Delta\Phi}]^2 dt}, \tag{13}$$

where T is the duration of the noise, 500 ms, $\overline{\Delta\Phi}$ is the mean IPD computed over time (normally very close to zero), and the t subscript indicates a standard deviation computed over time. The ILD fluctuation over time is defined as

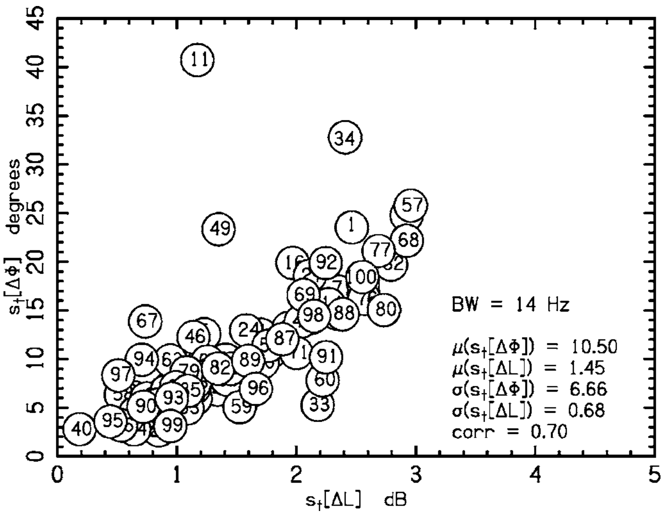

FIG. 1. Fluctuations of IPD vs fluctuations of ILD for the collection of 100 reproducible noise-pairs having a 14-Hz bandwidth, as used in Experiment 1. Each noise-pair is labeled by a serial number indicating only the order of creation. The means, standard deviations, and the correlation between $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ (denoted by corr) are reported.

$$s_t[\Delta L] = \sqrt{\frac{1}{T}\int_0^T [\Delta L(t) - \overline{\Delta L}]^2 dt}. \tag{14}$$

Figure 1 shows $s_t[\Delta\Phi]$ plotted against $s_t[\Delta L]$ for the 14-Hz bandwidth noise-pairs. The mean, standard deviation, and correlation of $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ are reported in Fig. 1. It is interesting to see that the correlation between phase fluctuations and level fluctuations across different noise-pairs is fairly strong, 0.70. To determine whether the statistics of our 100-noise ensembles were typical, we computed similar statistics for ensembles of 5000. These computations are described in the Appendix, which shows that a correlation of 0.70 is somewhat low compared to a typical value for this bandwidth, namely 0.80.

To perform Experiment 1, the five noise-pairs with the greatest fluctuations in IPD and the five noise-pairs with the smallest fluctuations in IPD were selected to form a *phase set* of ten noise-pairs. Similarly, the ten noise-pairs with the greatest and smallest fluctuations in ILD were selected to form a *level set*.

The two channels of noise were computed and downloaded by a Tucker-Davis AP2 array processor (System II) and converted from digital signals by 16-bit DACs (DD1). The buffer size was 4000 samples per channel and the sample rate was 8 ksps; thus the duration of the noises was 500 ms. The noise was lowpass filtered with a carrier frequency of 4 kHz and a −115 dB/octave roll off. The noises were presented at 70 dB SPL±3 dB with levels determined by programmable attenuators (PA4, prior to lowpass filtering) operating in parallel on the two channels. The levels of the two channels were equal and were randomized over a range of ±3 dB for each of the three intervals within a trial in order to discourage the listener from trying to use level cues to perform the task.

TABLE I. The p-values from a two-sample one-tailed t-test for phase and level sets with different bandwidths—Experiments 1, 2, and 3.

	14 Hz		108 Hz		2394 Hz	
Listener	Phase	Level	Phase	Level	Phase	Level
D	<0.001	<0.001	0.024	0.007	0.907	0.375
M	0.002	0.002	0.019	0.047	0.981	0.080
P	0.011	<0.001	0.187	0.187	0.683	0.654
W	0.001	<0.001	0.002	0.026	0.947	0.866

C. Procedure

Listeners were tested individually, seated in a double-walled sound attenuating room and using Sennheiser HD414 headphones. Six runs were devoted to listening to a set of ten reproducible noise-pairs. A noise-pair could be presented either incoherently–the dichotic presentation of x_L and x_R–or it could be presented coherently–the diotic presentation of x_L. A run consisted of 60 trials, where each of the ten noise-pairs in a set was presented incoherently a total of six times. Thus, listeners heard an individual noise-pair incoherently a total of 36 times (six runs times six presentations per run).

On each trial the listener heard a three-interval sequence. The first interval was the standard interval, which was always a coherent noise. The second interval was randomly chosen to be either incoherent or coherent. The third interval was the opposite of the second (e.g., if the second interval was coherent, the third interval was incoherent). The two coherent presentations were randomly selected from the remaining nine reproducible noises in the set except that they were required to be different from the incoherent "odd" interval and to be different from one another. The interinterval duration was 150 ms.

D. Data collection

Listeners used a four-button response-box to make decisions. Four buttons were used so that listeners could respond to the correct interval with a confidence estimate. The buttons from left to right were 2!, 2, 3, and 3!, representing confident second interval, second interval, third interval, and confident third interval respectively. Listeners were instructed to use a confident response only if there was no uncertainty as to which interval was incoherent. If more than one incorrect confident response occurred during a run, the run was terminated immediately, and the listener was obliged to begin again. After a decision was made by the listener, the next trial began following an intertrial duration of 900 ms.

There were two reasons to introduce the confidence measure in this experiment. The first was that for a given coherence, it proved to be noticeably easier to detect incoherence in some stimuli. Thus, extra weighting was wanted for identifying obviously incoherent sounds. The second was to increase the "dynamic range" of the experiment. The dynamic range problem arose because it was necessary to use the same waveforms, with a coherence of $\rho=0.9922$, for all the listeners, and it was further desired to use the same coherence for waveforms with different bandwidths. However, some listeners were better at the task than others, as shown by the percentage of correct (P_c) responses, and some bandwidths led to better performance than others. The confidence measure increased the dynamic range by preventing ceiling effects for the most successful listeners and easiest bandwidths.

The data collection procedure kept track of both the percentage of correct responses, which ignored the confidence estimate (e.g., responses of 2 and 2! were not treated differently), and a confidence adjusted score (CAS). The CAS is defined as the number of times the listener responded correctly plus the number of times that the listener was confident about the correct response. Since an individual noise-pair was heard 36 times, it was possible for a listener to get a score of 72 if the listener was able to respond correctly and confidently for all 36 presentations. This article will report only the CAS. In comparison with P_c, the use of CAS improved interlistener correlation, moved p-values of t-tests to greater significance, and improved the agreement that was achievable by models of binaural processing. Further justification of this technique together with the P_c results are in the first author's thesis (Goupell, 2005).[2]

E. Listeners

This article employed four male listeners, D, M, P, and W. Listeners D, M, and P were between the ages of 20 and 30 and had normal hearing according to standard audiometric tests and histories. Listener W was 65 and had a mild bilateral hearing loss, but only at frequencies four octaves above the center frequency used in this experiment. Listeners M and W were the authors.

F. Results

Figure 2(a) shows the CAS values for the selected phase set in Experiment 1. The five smallest $s_t[\Delta\Phi]$ noise-pairs are to the left of the vertical dashed line. The five largest $s_t[\Delta\Phi]$ noise-pairs are to the right of the dashed line. The dashed line thus represents a gap of 90 unused noise-pairs. All four listeners usually show a greater CAS for the five largest $s_t[\Delta\Phi]$ noise-pairs compared to the five smallest. In a two-sample t-test, all four differences were significant at the 0.02 level. The individual p-values are shown in the left-most block of Table I. Figure 2(a) also shows that listeners tend to agree about which noise-pairs are easy and which are difficult. The left-most block of Table II shows the correlations between listeners; all of them are above 0.77.

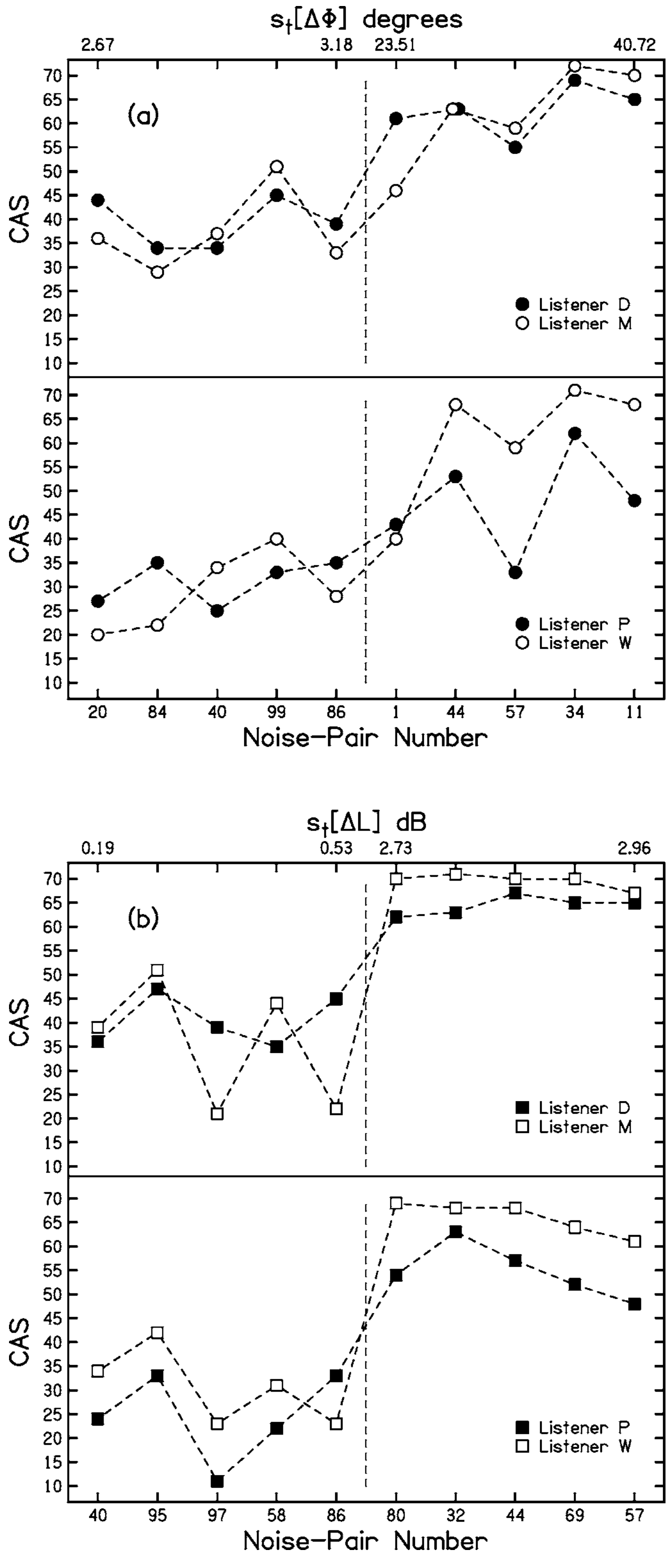

FIG. 2. (a) The confidence adjusted scores (CAS) for four listeners for the *phase set* of Experiment 1, the 14-Hz noise-pairs. The noise-pairs were chosen to have the five smallest and the five largest values of $s_t[\Delta\Phi]$ in the collection of 100 noise-pairs. The noise-pairs are rank ordered by increasing $s_t[\Delta\Phi]$ along the horizontal axis. The vertical dashed line represents 90 unused noise-pairs. The CAS values are higher for noise-pairs with the largest $s_t[\Delta\Phi]$ than for noise-pairs with the smallest $s_t[\Delta\Phi]$ for all listeners. The plots of CAS scores vs noise-pair serial number show a large measure of agreement among listeners. (b) The CAS for four listeners for the *level set* of Experiment 1, the 14-Hz noise-pairs. The noise-pairs were chosen to have the smallest and largest $s_t[\Delta L]$ in the collection of 100 noise-pairs.

Figure 2(b) shows the CAS values for the level set for Experiment 1. Figure 2(b) exhibits results similar to Fig. 2(a) for the phase set. As shown by Tables I and II, the p-values are all significant at the 0.02 level, and the correlations between the individual listeners are all greater than or equal to 0.86.

The percent-correct values were computed to compare with the CAS. In terms of P_c, seven of eight t-tests were significant at the 0.05 level, and two were significant at the 0.02 level. The average interlistener correlation in P_c was 0.56 for the phase set and 0.64 for the level set. Therefore the P_c results also show significant differences among the noise-pairs, but the P_c results are less robust than the CAS results due to ceiling effects.

G. Discussion

Experiment 1 shows that the peak of the cross-correlation function is not an adequate predictor of the detectability of incoherence for narrowband noise. Instead, the fluctuations in interaural phase and interaural level as measured by the standard deviation are clearly related to incoherence detection.

The interlistener correlation of 0.77 or greater indicates that listeners agree in detail about the kinds of fluctuations that are easy or hard to detect. This correlation can also be seen in Figs. 2(a) and 2(b). The correlation between listeners is clear, even when the detection results are out of line with expectation based on interaural fluctuations. For example, Fig. 2(a) shows that the incoherence in noise-pair #57 was difficult to detect compared to other large-fluctuation noise-pairs of the phase set, and Fig. 2(b) shows that incoherence in noise-pair #95 was easy to detect compared to the small-fluctuation noise-pairs of the level set.

The anomalous results obtained with noise-pairs such as #57 and #95 indicate that the phase sets and level sets were not assembled using optimum criteria. For example, when the envelope is small, dramatic phase fluctuations can occur because the arg function in Eq. (9) becomes unstable. However, when the envelope is small, such phase fluctuations would be hard to detect. Nevertheless these fluctuations are not discounted in the criteria. A serious attempt to optimize criteria will be the subject of a following article. For the present, it is enough to note that the criteria as applied represent the phase and level fluctuations well enough to lead to highly significant differences.

Last, it should be reported that some noise-pairs [e.g., #97 in Fig. 2(b)] led to values of percent correct that were well below chance. Through informal listening experiments, it was found that these noise-pairs had very little roughness in their waveforms. It seems possible that when the fluctuations of IPD and ILD are barely detectable, listeners may sometimes mistake roughness for incoherence. They would then be reluctant to say that a particularly smooth sounding noise-pair is incoherent, which could lead to a P_c less than 50%. (See Sec. VI.)

TABLE II. The correlations between listeners for phase and levels sets with three different bandwidths—Experiments 1, 2, and 3.

Listener pair	14 Hz		108 Hz		2394 Hz	
	Phase	Level	Phase	Level	Phase	Level
D-M	0.90	0.86	0.42	0.54	0.57	0.24
D-P	0.85	0.93	0.05	−0.04	−0.06	0.14
D-W	0.87	0.93	0.67	0.58	0.43	0.13
M-P	0.77	0.89	0.31	−0.03	0.17	0.04
M-W	0.97	0.97	0.33	0.30	0.38	−0.68
P-W	0.79	0.93	0.49	0.31	0.03	−0.49

III. EXPERIMENT 2: INCREASED NOISE BANDWIDTH

Experiment 2 was identical to Experiment 1 except that the bandwidth was increased from 14 to 108 Hz. The value of the coherence remained fixed at 0.9922.

A. Method

As for Experiment 1, the geometric mean frequency was 500 Hz and the spectral spacing was 2 Hz. The spectral shaping again had 5-Hz edges so that components 444–449 and 555–560 Hz were shaped with raised-cosine functions, and components between 449 and 555 Hz had unity amplitude. The sample rate was 8 ksps and the analog output was low-pass filtered with a 4-kHz corner frequency.

Figure 3 shows $s_t[\Delta\Phi]$ vs $s_t[\Delta L]$ for the 108-Hz bandwidth noise-pairs. The black dots represent the noise-pairs with 14-Hz bandwidth from Experiment 1 for comparison. The means of the distributions of fluctuations increased only slightly when compared to the 14-Hz bandwidth waveforms. The insensitivity of the mean fluctuation to the bandwidth has been seen before, e.g., Fig. 3 of Breebaart and Kohlrausch (2001). By contrast, the standard deviations across different waveforms decreased. The standard deviation of the phase fluctuations decreased from 6.66° to 3.72°. The standard deviation of the level fluctuations decreased from 0.68 to 0.40 dB. The correlation between $s_t[\Delta\Phi]$ and $s_t[\Delta L]$, evaluated over the ensembles of 100 waveforms, was essentially the same, 0.73 compared to 0.70. The Appendix shows that a value of 0.73 is in line with expectation based on statistics for a large number of noise-pairs.

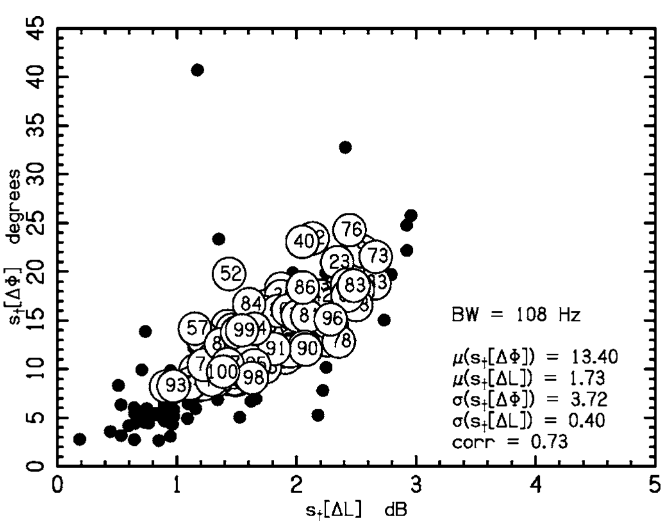

FIG. 3. Fluctuations of IPD vs fluctuations of ILD for the 100 reproducible noise-pairs with a 108-Hz bandwidth. Each noise-pair is labeled by a serial number. The means, standard deviations, and IPD-ILD correlation of the distributions are reported. The closed dots represent the noise-pairs with 14-Hz bandwidth from Experiment 1, as shown in Fig. 1. The means remained about the same as in Fig. 1, but the standard deviations and IPD-ILD correlation decreased.

B. Results

Figures 4(a) and 4(b) show the CAS values for the four listeners for the phase and level sets for the 108-Hz noise-pairs of Experiment 2. Table I shows that three of the four t-tests from the phase set were significant at the 0.05 level. That means that the differences between the CAS values for the five largest fluctuations and the CAS values for the five smallest fluctuations were significant for three listeners. Also, three of the four t-tests from the level set led to differences that were significant at the 0.05 level. In both cases, it was Listener P who did not show a significant difference. Further, as shown in Table II, the correlation between listeners was smaller for the wider bandwidth for all listener pairs. The correlation between listeners was 0.38 on average for the phase set and 0.28 on average for the level set. Listener P showed negative correlations with Listeners D and M.

C. Discussion

The standard deviation of the fluctuations computed across the 100 different waveforms of the ensemble decreased by more than 40% when the bandwidth was increased from 14 to 108 Hz. Since these fluctuations were found to correlate with incoherence detection from Experiment 1, it was not surprising to find that there was less variation in the listeners' ability to detect incoherence at the wider bandwidth.

The bandwidth of 108 Hz may be of special interest because this bandwidth approximately corresponds to a critical bandwidth at 500 Hz, and critical band noise has often been used in binaural experiments. For instance, Koehnke *et al.* (1986) used a noise with a bandwidth of 114 Hz centered on 500 Hz, and Evilsizer *et al.* (2002) used a bandwidth of 100 Hz. Our experiments found that the tests comparing detection performance with the size of the phase and level fluctuations led to a significant difference at the 0.05 level on six of eight of the tests. This can be compared to all of the tests

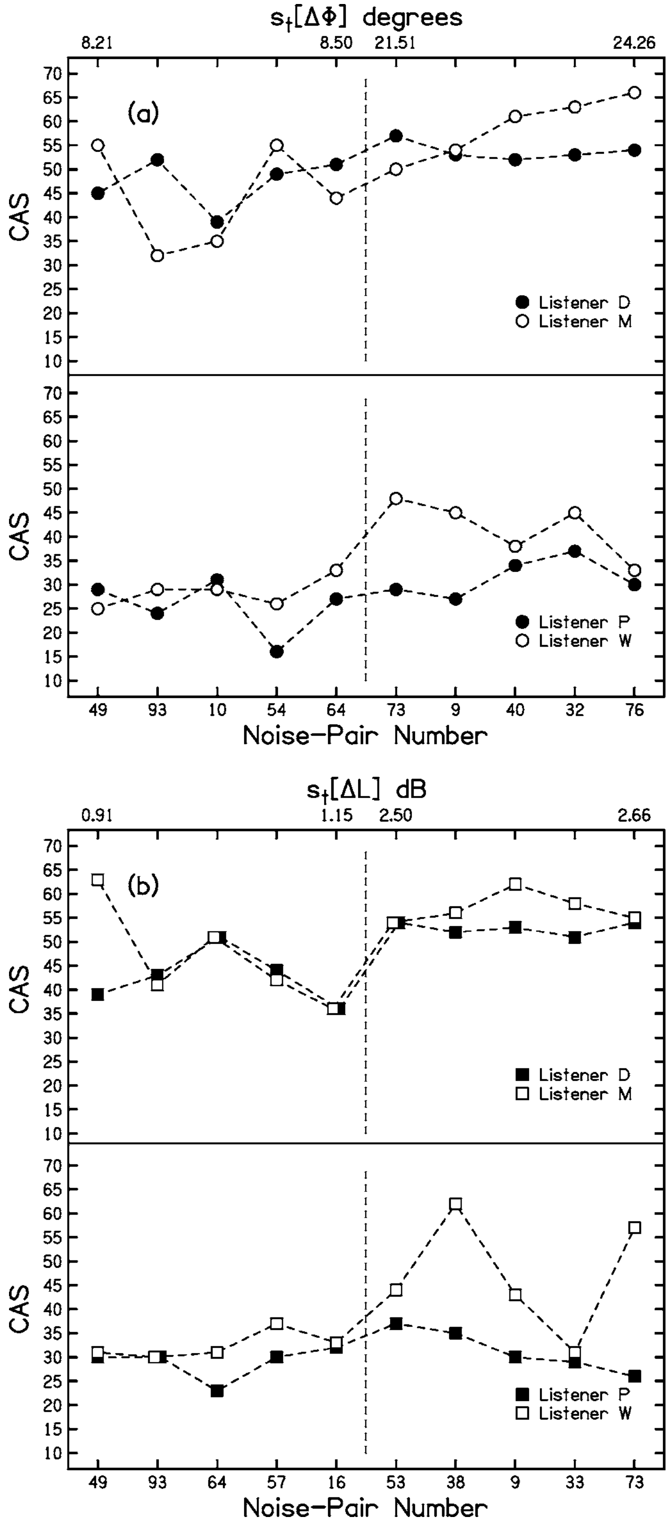

FIG. 4. (a)The CAS for the phase set of Experiment 2, the 108-Hz noise-pairs. The noise-pairs to the left and right of the dashed vertical line were chosen to have the smallest and largest $s_t[\Delta\Phi]$ in the collection of 100 noise-pairs. On average, the CAS is smaller than for the 14-Hz noise-pairs shown in Fig. 2(a). The difference between the left and right noise-pairs is less dramatic than in Fig. 2(a). Also, the plots of CAS vs noise-pair serial number show less agreement among listeners than in Fig. 2(a). (b) The CAS for the level set of Experiment 2, the 108-Hz noise-pairs. The noise-pairs to the left and right of the dashed vertical line were chosen to have the smallest and largest $s_t[\Delta L]$ in the collection of 100 noise-pairs. On average, the CAS is smaller than for the 14-Hz noise-pairs as shown in Fig. 2(b). The difference between the left and right noise-pairs is less dramatic than in Fig. 2(b). Also, the plots of CAS vs noise-pair serial number show less agreement among listeners than in Fig. 2(b).

being significant at the 0.02 level in Experiment 1. It was expected that if the bandwidth were further increased the variation in fluctuations would continue to decrease and the incoherence detection performance would be approximately the same for all the waveforms in the ensemble. In that case, the ability to detect incoherence would be only a function of the incoherence measure itself. That expectation led to Experiment 3.

IV. EXPERIMENT 3: WIDE BAND

A. Method

Experiment 3 was identical to Experiments 1 and 2 except that the bandwidth was increased to 2394 Hz. The geometric mean frequency of 500 Hz and the coherence of 0.9922 remained the same. Spectral components between 105 and 2495 Hz had unity amplitude and components in the ranges of 100–105 and 2495–2500 Hz were shaped by the raised-cosine spectral envelope. The sample rate and lowpass corner frequency were again 8 ksps and 4 kHz.

For Experiment 3, with a band more than four octaves wide, several features of the narrower-band stimuli in Experiments 1 and 2 no longer apply. Interaural phase differences are no longer equivalent to interaural time differences, and the temporal variations can no longer be neatly separated into fine structure and envelope. The procedures for selecting waveforms in Experiment 3 followed those for Experiments 1 and 2 in that noise-pairs were selected according to their IPDs (not ITDs), and the mathematical definition of the envelope was formally extended to the wider bandwidth without change. Because binaural processing is expected to occur within tuned channels, a selection of noise-pairs according to criteria based on the entire bandwidth does not have an evident perceptual significance. However, following the procedure of Experiments 1 and 2 did lead to uniformity of treatment.

Figure 5 shows the distribution of interaural phase and level fluctuations for one-hundred 2394-Hz bandwidth noises. The corresponding values for Experiments 1 and 2 are shown by the small and large black dots, respectively. The means of the distribution remained about the same as in Experiment 2 in that the ensemble average phase fluctuation was about 13° and the level fluctuation was approximately 1.7 dB. However, the standard deviations of the phase and level fluctuations decreased dramatically when the bandwidth was increased to 2394 Hz, as shown by the σ values in Fig. 5, respectively 1.07° and 0.11 dB. The correlation between level and phase fluctuations decreased to 0.27. The dramatic decrease in the variability of interaural fluctuations approaches the wideband limit where fluctuations are entirely determined by the coherence with negligible variation for individual waveforms.

B. Results

The results of the incoherence detection experiments, expressed as CAS values, are shown in Figs. 6(a) and 6(b) for the phase and level sets respectively. The CAS values in the figures are remarkably flat compared to Figs. 2(a), 2(b), 4(a), and 4(b). With four listeners and two sets there were

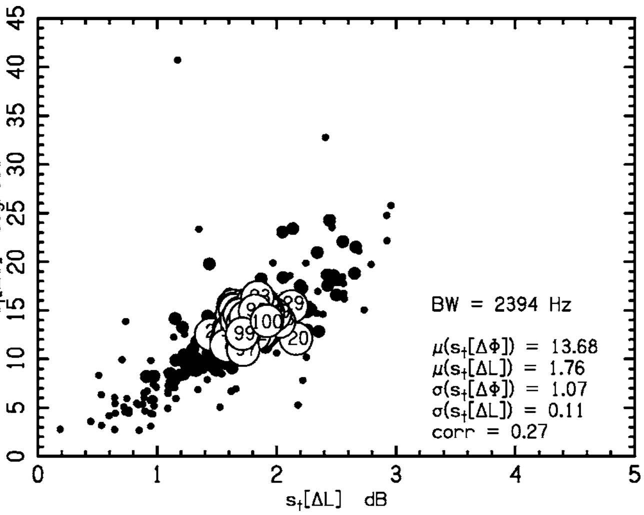

FIG. 5. Fluctuations of IPD vs fluctuations of ILD for the 100 reproducible noise-pairs with a 2394-Hz bandwidth. Each noise-pair is labeled by a serial number. The means, standard deviations, and IPD-ILD correlation of the distributions are reported. The small closed dots represent the 14-Hz noise-pairs from Experiment 1 and the larger closed dots represent the 108-Hz noise-pairs from Experiment 2, as shown in Figs. 1 and 3. The means are about the same as those in Figs. 1 and 2 but the standard deviations and correlation are smaller.

eight possible significance tests, and Table I shows that none of them led to a significant difference at the 0.05 level. Further, Table II shows that the interlistener correlations were smaller on average than for the other two bandwidths.

C. Discussion

The stimuli for Experiments 1, 2, and 3 show clear trends in the ranges of fluctuations over waveforms. The broadband stimuli of Experiment 3 approach the wideband limit. The incoherence detection results for Experiments 1, 2, and 3 also show clear trends. For narrow bands, detection depends on the details of interaural fluctuations, and it is not possible to predict detection performance if one knows only the value of coherence. As the bandwidth increases, the coherence becomes a better predictor of detection performance. In addition to the wideband limit for the stimulus, one can imagine a wideband *perceptual* limit in which incoherence detection is entirely determined by the coherence with negligible variation for individual waveforms.

The wideband perceptual limit was apparently reached for the bandwidth of 2394 Hz because no systematic differences appeared between those noise-pairs with the largest fluctuations and those noise-pairs with the smallest. However, this evidence is suspect because it is not clear that those large and small fluctuations, which were computed over the entire band of noise, are perceptually relevant. Instead, following binaural theory, one would expect that fluctuations computed over a critical band would be relevant. For instance, the MLD experiments by van de Par and Kohlrausch (1999) indicated that listeners detect an $S\pi$ signal in noise by monitoring a critical band. This difficulty has two parts. The first is in the interpretation of the data from Experiment 3. The second is in understanding the nature of the wideband perceptual limit.

If listeners monitor a critical band, in agreement with binaural theories, then our choices of noise-pairs in Experiment 3 were incorrectly made and there is no reason to expect better detection performance for noise-pairs to the right of the vertical dashed line compared to those on the left. Of course, given a bandwidth of 2394 Hz there is no way to know which critical band listeners might monitor, but it is usual to assume that listeners learn to monitor a band that is useful to them in performing the task. Therefore, it is expected that listeners would monitor a band somewhere in the 500-Hz neighborhood, where IPD fluctuations are of greatest use. Consequently the fluctuations in a critical band would resemble those for the 108-Hz bandwidth in Experiment 2. Possibly the range of fluctuations might be somewhat larger than in Experiment 2 because critical bands near 500 Hz are somewhat narrower than 108 Hz. In that case, the uniformity of performance across the 17 noise-pairs of Experiment 3 poses a problem. Whether or not the noise-pairs were well chosen, the odds are good that the relevant fluctuations, whatever they might be, would be uncharacteristically strong in at least one of the 17 noise-pairs. Thus it would be expected that at least one pair would stand out from the rest, but there is no evidence of that experimentally.

A second difficulty posed by our data for a model incorporating critical band listening arises from the overall level of detection performance. Listeners performed much less well for the wide band in Experiment 3 compared to the 108-Hz band of Experiment 2. Comparing CAS values in Figs. 4(a) and 6(a) (phase sets) and comparing values in Figs. 4(b) and 6(b) (level sets) show that the most successful listeners, D and M, had higher scores even for the most difficult noise-pairs at 108 Hz than for *any* noise-pairs with wide bands. Listeners P and W also scored consistently better at 108-Hz for the level sets. If listeners were able to take advantage of the slower and potentially larger fluctuations in a critical-band portion of the wide band one would have expected that some listener would have scored well for some one of the noise-pairs, contrary to the results of Experiment 3.

The data of Experiment 3 are more easily understood from an alternative view which says that in a broadband incoherence detection experiment, listeners cannot monitor a critical band but must monitor a much wider band. There is evidence in favor of this view in the detection results of Gabriel and Colburn (1981) where incoherence jnds generally increased for increasing bandwidth and continued to increase as the bandwidth went from 1000 to 4500 Hz. There is informal evidence in favor of the wider analysis bandwidth in the fact that the spatial fluctuations seem to the authors to be faster for the 2394-Hz bandwidth than for 108 Hz.

If the operative perceptual bandwidth were to be as large as the stimulus bandwidth, then our choice of noise-pairs would be correct and Experiment 3 would have demonstrated the wideband perceptual limit. Although it is hard to argue for a perceptual bandwidth as large as the stimulus bandwidth, the data do suggest that the perceptually relevant frequency range extends beyond a critical band. A similar conclusion from a MLD experiment was reached by Evilsizer *et al.* (2002), who found that $NoS\pi$ detection was *not*

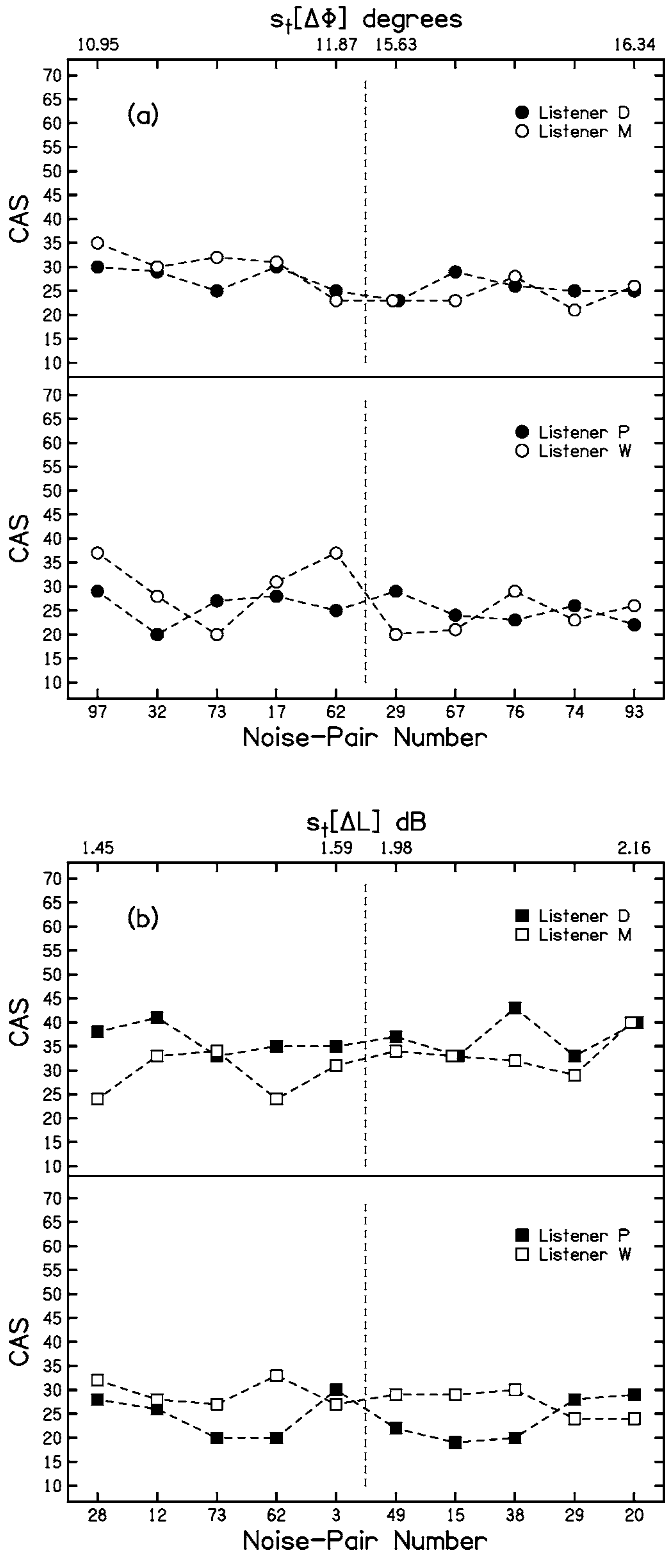

FIG. 6. (a) The CAS for the phase set of Experiment 3, the 2394-Hz noise-pairs. The noise-pairs were chosen to have the smallest and largest $s_t[\Delta\Phi]$ in the collection of 100 noise-pairs. Compared to Figs. 2(a) and 4(a), the differences between CAS values for different noise-pairs are less dramatic and the correlation among individual listeners is smaller. (b) The CAS for the level set of Experiment 3, the 2394-Hz noise-pairs. The noise-pairs were chosen to have the smallest and largest $s_t[\Delta L]$ in the collection of 100 noise-pairs. Compared to Figs. 2(b) and 4(b), the differences between CAS values for different noise-pairs are less dramatic and the correlation among individual listeners is smaller.

solely determined by the masker spectrum within a critical band centered on the target.

The wideband perceptual limit is defined as the condition in which incoherence detection performance is predictable from the coherence alone. As noted previously, the odds are that Experiment 3 actually reached that limit. There remains the problem, however, of understanding what that limit represents. If the perceptual bandwidth is as large as the stimulus bandwidth then the wideband perceptual limit is not different in character from the wideband limit for the stimulus. The perceptual limit is reached when there are so many spectral components in the stimulus that, for a given coherence, the sizes of fluctuations are about the same for all different noises.

If the perceptual bandwidth is as small as a critical band, large fluctuations are available to the listener at a peripheral level, but for some reason listeners cannot take advantage of them in detecting incoherence. Perhaps the wideband limit should be thought of as a multiband limit wherein fluctuations in some bands serve as distractors that detract from the detection of fluctuations in other bands. Incoherence detection in broadband noise raises questions of optimum filter choice similar to questions that arise in MLD experiments with respect to off-frequency listening strategies. However, in the case of incoherence detection, there would seem to be no systematic approach to these questions because what corresponds to the signal varies with individual waveforms.

V. EXPERIMENT 4: THE ROLE OF BANDWIDTH

Experiments 1, 2, and 3 demonstrated that as the bandwidth increases two effects occur. First, the ranges of fluctuations in IPD and ILD become narrower. Second, the ability of listeners to detect incoherence depends less on the individual noise-pairs and is better determined by the value of coherence itself. According to our hypothesis, the second effect is the direct result of the first, and the main effect of a variation in bandwidth is to alter the distributions of interaural variances. Experiment 4 was designed to test this idea.

A. Method

To test the hypothesis, we assembled a subset of noise-pairs from a new collection of 1000 noise-pairs with a bandwidth of 14 Hz to make a "matched set" whose members were selected to best match the fluctuations in the noise-pairs from Experiment 2, which had a bandwidth of 108 Hz. Experiment 2 included 20 noise-pairs, ten for the phase set and ten for the level set, as determined by the five largest and five smallest fluctuations. However, phase and level fluctuations tend to be correlated and five of the noise-pairs were common to the phase and level sets. Therefore, there were only 15 distinct noise-pairs in Experiment 2. For each of these 108-Hz bandwidth noise-pairs a 14-Hz bandwidth noise-pair was selected that best matched the fluctuations in phase and level. The selection is illustrated by the 15 open and filled circles in Fig. 7. Phase and level sets using the matched noise-pairs formed the stimuli for Experiment 4, which was otherwise identical to the other experiments of this article.

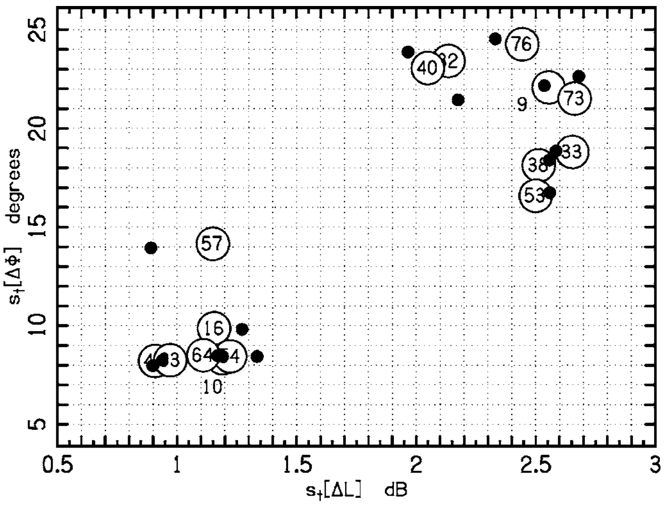

FIG. 7. The fluctuations in IPD vs the fluctuations in ILD for 15 noise-pairs used in Experiment 2 (bandwidth of 108 Hz) are shown by open circles. The best matching noise-pairs from Experiment 4 (bandwidth of 14 Hz) are shown by closed circles. These noise-pairs are called the "matched set." Incoherence detection results for these corresponding pairs were compared to determine the role of bandwidth *per se* when fluctuations were held constant.

B. Results

According to our working hypothesis, the detection scores (CAS) for the matched 14-Hz sets from Experiment 4 ought to be identical to the detection scores in the 108-Hz sets from Experiment 2. The results of the comparison are shown in Figs. 8(a) and 8(b) for the phase and level sets, respectively. The phase set data in Fig. 8(a) show that the values of CAS are comparable for the two bandwidths when the IPD fluctuations are matched for Listeners D and M. However, for Listeners P and W performance is better for the 14-Hz bandwidth than for the 108-Hz bandwidth.

The level set data in Fig. 8(b) show comparable CAS values for the two bandwidths for all the listeners, though Listeners P and W still tend to show better performance at the smaller bandwidth—better on 14 of 20 possible comparisons. Therefore, the raw data, shown in Fig. 8, offer modest support for the hypothesis that bandwidth should be unimportant if the sizes of fluctuations are matched.

The hypothesis can be further tested by examining the relative detectability of the noise-pairs with the largest interaural fluctuations versus the noise-pairs with the smallest interaural fluctuations. These are, respectively, to the right and to the left of the vertical dashed line in Fig. 8. A t-test of the hypothesis that CAS scores are higher for the five noise-pairs on the right led to the p-values in Table III. There, it can be seen that three of eight p-values are significant at the 0.05 level for the 14-Hz matched sets, and that six of eight p-values are significant at that level for the targeted 108-Hz sets. By comparison, all eight p-values were significant at the 0.02 level for the 14-Hz bandwidth in Experiment 1 and none of the p-values were significant even at the 0.05 level for the wide bandwidth in Experiment 3. Thus, the matched phase and level sets appear to have approximately matched the relative difference in performance between the noise-

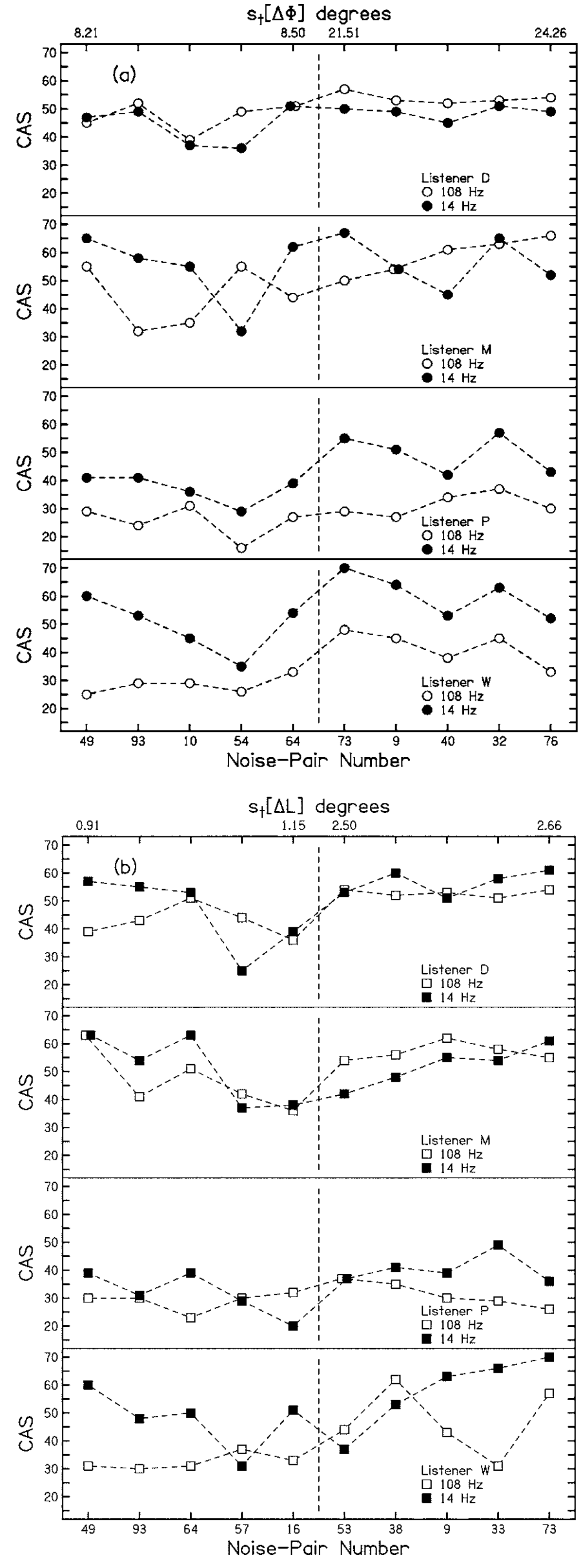

FIG. 8. (a) The CAS for Experiment 4, using the matched phase set selected from the 14-Hz noise-pairs are shown by closed circles. They are plotted together with the data from Experiment 2, the 108-Hz noise-pairs, shown by open circles, repeated from Fig. 4(a). (b) The CAS for Experiment 4, using the matched level set selected from the 14-Hz noise-pairs are shown by closed squares. They are plotted together with the data from Experiment 2, the 108-Hz noise-pairs, shown by open squares, repeated from Fig. 4(b).

TABLE III. The p-values from a two-sample one-tailed t-test for the 108-Hz phase and level sets in Experiment 2 (repeated from Table I) and for the matched 14-Hz phase and level sets in Experiment 4.

	14 Hz		108 Hz	
Listener	Phase	Level	Phase	Level
D	0.103	0.077	0.024	0.007
M	0.384	0.315	0.019	0.047
P	0.006	0.038	0.187	0.187
W	0.042	0.117	0.002	0.026

pairs in the Experiment 2 sets, consistent with the hypothesis that the size of interaural fluctuations determines the detection of incoherence.

However, there are differences between the results of Experiments 2 and 4. As noted above, there is a tendency for listeners P and W to score better on the 14-Hz bandwidth sets (Experiment 4). More impressive, a comparison of the interlistener correlations in Table IV shows that all 12 of the correlations are higher for the 14-Hz matched sets than for the 108-Hz bandwidth sets. Averaged over listener pairs and over phase and level sets, the interlistener correlation was 0.64 (sd=0.11) for the 14-Hz bandwidth and only 0.33 (sd =0.23) for the 108-Hz bandwidth.

C. Discussion

Experiment 4 attempted to construct a set of noise-pairs with a 14-Hz bandwidth that would lead to the same patterns of detection performance that had been seen in Experiment 2, which used noise-pairs with a 108-Hz bandwidth. This was done by best-matching the values of $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ of the Experiment 2 noise-pairs with noise-pairs having a 14-Hz bandwidth. Experiment 4 was partially successful in that the number of significant p-values was comparable in Experiments 4 and 2, and quite different from Experiments 1 and 3. It is clear that the fluctuation matching procedure caused the 14-Hz bandwidth noise-pairs of Experiment 4 to behave more like the 108-Hz noise-pairs of Experiment 2 than like 14-Hz noise-pairs of Experiment 1.

Although the values of CAS for the 108-Hz bandwidth (Experiment 2) and the 14-Hz bandwidth (Experiment 4) were interleaved for two of the listeners, the values of CAS were higher for the 14-Hz bandwidth for the other two, and interlistener correlations were higher overall for the 14-Hz bandwidth. Thus, there seems to be an advantage for the smaller bandwidth, whether the fluctuations are large or small. An advantage for smaller bandwidth was also found by Gabriel and Colburn (1981) where the jnd for detecting incoherence decreased with decreasing bandwidth. The impressive listener correlation for the 14-Hz bandwidth compared to 108-Hz may find a straightforward explanation in terms of the number of instances of salient fluctuations. For a 14-Hz bandwidth, fluctuations tend to be slow and within a duration of 500 ms there are relatively few strong fluctuations. Because of the absence of many alternatives, listeners are likely to detect these fluctuations in similar ways. For a 108-Hz bandwidth, the number of fluctuations is about eight times greater and different listeners may use different fluctuations to perform the task.

TABLE IV. The correlations between listeners for the 108-Hz phase and level sets of Experiment 2 (repeated from Table II) and for the matched 14-Hz phase and level sets of Experiment 4.

	14 Hz		108 Hz	
Listener pair	Phase	Level	Phase	Level
D-M	0.68	0.71	0.42	0.54
D-P	0.60	0.66	0.05	−0.04
D-W	0.75	0.68	0.67	0.58
M-P	0.66	0.56	0.31	−0.03
M-W	0.80	0.69	0.33	0.30
P-W	0.91	0.46	0.49	0.31

VI. EXPERIMENT 5: MONAURAL CUES

During the course of our numerical study of interaural fluctuations we compared the interaural fluctuations in noise-pairs with the envelope fluctuations in the left-ear signal itself. (Left and right signals were so similar that it did not matter which we chose.) *A priori* it seemed possible that noise-pairs with especially large (small) interaural fluctuations might often be derived from individual noise tokens with especially large (small) envelope fluctuations.

To make the comparison we calculated the envelope fluctuation,

$$s_t[E] = \sqrt{\frac{1}{T}\int_0^T [E(t) - \bar{E}]^2 dt}, \tag{15}$$

where $E(t)$ is the envelope defined in Eq. (10), and $\bar{E}$ is the average over the time of the stimulus. Calculations with 1000 noise-pairs with 14-Hz bandwidth–stochastically identical to the pairs of Experiment 1–showed a positive correlation between envelope fluctuations and interaural fluctuations. Specifically, the correlation between $s_t[E]$ and $s_t[\Delta L]$ was 0.48. The correlation between $s_t[E]$ and $s_t[\Delta\Phi]$ was 0.43. The latter correlation would be hard to understand were it not for the strong correlation between the standard deviations of interaural phase and level differences, as shown in Figs. 1, 3, and 5.

We next made the same calculations for the 20 noise-pairs actually used in Experiment 1. The interaural phase fluctuations and the interaural level fluctuations correlated with the monaural envelope fluctuations at levels of 0.59 and 0.65, respectively. Evidently, those waveforms with interaural fluctuations that are especially large or small particularly owe their binaural character to the envelope of the original generating noise token.

Given the positive correlation between interaural and monaural fluctuations, it seemed possible that there might be information in the monaural signals that was used by listeners in performing the experiments of this article. Because of the evident importance of fluctuations in Experiment 1, our attention centered on the stimuli used there, with a bandwidth of 14 Hz.

A. Diotic experiment method

Experiment 5 was identical to Experiment 1 except for the important difference that the left-ear signal of Experiment 1 was the signal for both ears in Experiment 5. In a second difference, the listeners in Experiment 5 had all completed Experiments 1–4 and therefore were highly experienced. The listeners were given three-interval sequences as before and were asked to apply the same strategy that they had used in the previous experiments. There was reason to believe that this approach might be successful because all the listeners volunteered that in the previous experiments they based their decisions on a sense of width, choosing the interval–either two or three–with the larger width. It seemed possible that the sense of roughness or other "action" associated with a diotic stimulus having large fluctuations could be associated with a sense of width. Consequently, we expected that each trial of Experiment 5 would constitute a comparison between apparent "widths" for a particular identified noise and a different, randomly-chosen, noise from the set of ten. The stimulus sets were the phase set and level set from Experiment 1.

B. Results

Listeners made a negligible number of "confident" responses. Apparently the strong sensation of width elicited by some of the dichotic stimuli did not occur with any of the diotic noises. Therefore, the CAS had little value and results of Experiment 5 were analyzed in terms of the percentage of the trials on which a given noise was selected over other noises. The statistic will be called P_s, percent selected. It can be compared with P_c, the percent correct in the dichotic experiment.

1. Large fluctuation comparison

Particular interest centered on the five noise-pairs for which the interaural fluctuations were the greatest. By examining the scores from the diotic experiment we expected to gain insight into the role that envelope fluctuations may have played in the dichotic experiment. The average values of P_s for those five noise-pairs for listeners D, M, P, and W were, respectively: for the phase set(%): 54, 61, 54, 63; for the level set (%): 63, 58, 72, 70.

Evidently, in the diotic experiment listeners chose the noises that had led to the largest interaural fluctuations clearly more than half the time. These numbers can be compared with the values of P_c in the dichotic experiment (Experiment 1) which averaged 88%.

2. Agreement between listeners

The agreement among the listeners was assessed by comparing values of P_s against noise serial number for listeners taken in pairs. Interlistener correlations for D-M, D-P, D-W, M-P, M-W, and P-W were as follows: for the phase set: 0.59, −0.38, 0.38, −0.22, 0.88, 0.17; for the level set: 0.70, 0.53, 0.76, 0.67, 0.89, 0.76.

The strongest correlation was between M and W. Listener P was responsible for the only negative correlations–both in the phase set. Correlations were clearly larger in the level set than in the phase set. The strong correlation indicates that listeners tended to agree about which fluctuations were salient.

3. Comparison with envelope fluctuation

A comparison between the listener selection of noises and fluctuation was assessed by comparing P_s with $s_t[E]$ as a function of the noise-pair serial number. Correlations for listeners D, M, P, and W were, respectively: for the phase set: 0.38, 0.84, −0.03, 0.88; for the level set: 0.85, 0.66, 0.65, 0.70.

Again, P is responsible for the only negative correlation. The positive correlation indicates that the choices that listeners make can be predicted based on the physical envelope fluctuations, as measured by the standard deviation of the envelope over time, especially for the level set.

4. Comparison with Experiment 1

A comparison between the results of the corresponding diotic and dichotic experiments was made by comparing P_s on Experiment 5 with P_c on Experiment 1, both as functions of the noise-pair serial number. Correlations for listeners D, M, P, and W were, respectively: for the phase set: 0.24, 0.53, −0.07, 0.44; for the level set: 0.82, 0.66, 0.81, 0.80.

Again, correlations are larger for the level set.

C. Discussion

The above-mentioned correlations are fairly impressive, with some exceptions for listener P. These include the correlations between P_s in Experiment 5 and P_c in Experiment 1 as well as the correlations between P_s and monaural and dichotic fluctuations. There are several possible interpretations of these correlations.

Possibly the correlation between P_c and P_s scores only represents a chain of stimulus circumstances. For a narrow bandwidth like 14 Hz, every kind of stimulus fluctuation seems to correlate with every other kind. Interaural phase fluctuations correlate with interaural level fluctuations and both correlate with noise envelope fluctuations. In a dichotic experiment probing the detection of interaural incoherence listeners attend to the interaural fluctuations. In a diotic experiment probing an evaluation of roughness or other stimulus action listeners attend to the envelope fluctuations. The results of the two experiments, as functions of the stimulus serial number, are similar because the interaural and monaural fluctuations behave similarly with respect to serial number.

Alternatively, it is possible that the correlation between P_c and P_s scores arises because listeners in a dichotic experiment are misled by monaural envelope fluctuations that are particularly large or particularly small. Given the enormous difference in the average P_c and P_s values for the five noise-pairs with large fluctuations, as noted in Subsection 1, it seems highly unlikely that monaural fluctuations *per se* contribute to listener judgments in the dichotic experiment when the detection of interaural incoherence is easy. But when detection of interaural incoherence is difficult, or impossible, the cues from monaural envelope fluctuations (or the lack of them) may influence judgments and are probably responsible for P_c values in a dichotic experiment that are less than chance.

VII. DISCUSSION AND CONCLUSION

Listeners are sensitive to small amounts of interaural incoherence in an otherwise diotic noise. Given a diotic noise as a comparison, listeners can detect a coherence change of 0.01, i.e., they are sensitive to the difference between 1.00 and 0.99 (Gabriel and Colburn, 1981). The goal of the present article was to understand the origin of this remarkable sensitivity.

A. Detection of incoherence

Experiment 1 selected stimuli from an ensemble of 100 reproducible left-right noise-pairs, all of which had a bandwidth of 14 Hz and an interaural coherence of 0.9922. It was found that those pairs that had a large fluctuation in IPD or large fluctuation in ILD were much more readily recognized as not perfectly coherent compared to pairs with small fluctuations. This result led to the conclusion that, for bandwidths as narrow as 14 Hz, the interaural coherence is not an adequate predictor of the ability to detect incoherence. Instead, the size of the interaural fluctuations tends to dominate.

Experiments 2 and 3 progressively increased the bandwidth and found that the ranges of fluctuations in IPD and ILD among different noise-pairs in an ensemble decreased with increasing bandwidth (see the Appendix). This observation led to the expectation that the detectability of incoherence would exhibit less variation for different noise-pairs with these wider bandwidths. Detection experiments similar to Experiment 1 showed an increasing uniformity in detectability for the incoherence in noises with increasing bandwidths as expected.

It was conjectured that the only reason that detection performance for 14-Hz bandwidth was different from performance for 108-Hz bandwidth was that more extreme values of average interaural fluctuations (both very small and very large) were available in the ensemble with the narrow bandwidth. In Experiment 4, a comparison was made between performance on noise-pairs with 108-Hz bandwidth and performance on a matched set of noise-pairs with 14-Hz bandwidth. Noise-pairs in the matched set were selected to have approximately the same interaural fluctuations as the pairs of the 108-Hz set. The comparison showed that detectability differences among different noise-pairs in the 14-Hz matched set were reduced to about the same level as for the 108-Hz noises-pairs, consistent with the conjecture. The overall performance on the 14-Hz matched set was approximately equal on the 108-Hz set for two of the listeners; it was consistently higher for the other two listeners. These two results from the comparison suggest that differences in interaural fluctuations are responsible for differences in the detectability of incoherence for different noises, but that for some listeners fluctuations of comparable size are more easily detected when the bandwidth is narrow. A possible explanation for the advantage of narrow bands is that fluctuations are slower. The role of fluctuation speed will be addressed in a following article.

Our conclusions differ from those of Breebaart and Kohlrausch (2001) who dismissed a specific role for IPD and ILD fluctuations in binaural detection because the distributions of those fluctuations, with or without a signal, failed to show a bandwidth dependence. By contrast their NρSπ detection experiments did show such a dependence. We agree that the $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ have mean values, averaged across waveforms, that are insensitive to bandwidth, as noted in Figs. 1, 3, and 5, and as shown in the widths of the distributions plotted by Breebaart and Kohlrausch (2001). However, the *variations* in $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ among different waveforms depend strongly on bandwidth, and these variations are responsible for the large detectability differences that we are reporting.

B. Caveats—duration and bandwidth

All the stimuli used in this article had a duration of 500 ms. The stimuli were all constructed to have an interaural coherence of 0.9922 computed over this duration. Because different noise-pairs exhibited very different detectabilities, this article reached a main conclusion that the coherence measure is inadequate to predict detection for narrow bands. A problem that arises with this conclusion is that one expects that the temporal analysis windows that are appropriate for binaural perception are considerably shorter than 500 ms. Therefore, it is possible that, after all, the coherence is a perfectly adequate statistic to predict incoherence detection, but that the coherence must be calculated over the correct (shorter) time interval(s). This point of view would say that it was a mistake to compute a fixed coherence over an interval as long as 500 ms and expect it to be perceptually valid because the coherence varies from one momentary analysis window to the next.

There are two responses to the above criticism. First, we have performed model calculations based on a running short-term cross-correlation, averaged over temporal windows ranging from 0 to 150 ms. The calculated results were compared with the results of incoherence detection experiments with 100 randomly chosen reproducible noises with bandwidths of 14 Hz and durations of 500 ms. It was found that short-term coherence is inadequate to account for the detectability of incoherence. Second, we have done a series of experiments with 14-Hz bandwidth and progressively shorter durations. Coherence values calculated over those short durations were again required to be 0.992. Stimuli were again selected based on large or small fluctuations over the duration, and listeners did discrimination experiments similar to Experiment 1. It was found that the results for durations of 100 and 50 ms were the same as the results for 500 ms in that noise-pairs with large fluctuations led to CAS values that were significantly greater compared to noise-pairs with the smallest fluctuations. Both these responses to the duration criticism appear in Goupell's thesis (2005) and will be discussed in articles to follow.

A second criticism pertains to the wideband noise-pairs investigated in Experiment 3. Specific noise-pairs were chosen with large or small fluctuations as computed over the entire bandwidth. However, it is expected that the perceptually relevant bands would be determined by auditory filters similar to critical bands. If so, then the basis for choosing the pairs was faulty and it would not be surprising to find that there is no significant difference in CAS values for the different pairs. However, as noted in Sec. IV C, a model that begins with critical-band filtering runs into difficulties in explaining both the poor performance and the uniformity of performance in Experiment 3. It would become necessary to add some mechanism of cross-band interference to explain the data.

C. Binaural processing

The psychoacoustical literature often connects binaural capabilities, particularly binaural release from masking, with interaural cross-correlation (Wilbanks and Whitmore, 1967; Domnitz and Colburn, 1976; Koehnke *et al.*, 1986). Wilbanks (1971) cited historical articles by Cherry, Licklider, and Jeffress "…supporting the notion that the binaural system is, logically speaking, a correlational detector."

The experiments of the present article agree with that conclusion, but only as a statistical approximation that gains validity as the bandwidth becomes larger. A correlational model adequately reflects the fluctuations only in the wideband limit. Wideband noise signals tend to be ergodic wherein the statistical properties of an ensemble of noises become manifest in any given noise sample as that particular noise evolves in time. Then the fluctuations in any given sample of wideband noise are not much different from those in any other sample, and the size of the fluctuations appears to be a simple function of the coherence. The transition to that limit was seen in the stimuli for Experiments 1–3 where, with increasing bandwidth, the variance among different noises of IPD and ILD decreases.

The wideband limit for signals has a psychological parallel in the wideband perceptual limit, where incoherence detection becomes a simple function of coherence. The wideband limit is approached but not reached for a bandwidth of 108 Hz, which is close to a one-third octave or critical bandwidth for 500 Hz. With a bandwidth that wide, a psychoacoustical experiment might easily miss the inhomogeneity of noises unless it were specifically designed to look for it.

For a narrow bandwidth, such as 14 Hz, each individual noise has so few spectral components that the interaural properties of an individual noise-pair can differ greatly from the ensemble mean properties. Improbable variations of this kind are responsible for the fact that a small amount of incoherence may be difficult to detect for one sample of noise but easy to detect for a another one, as reported in this article. It is likely that this effect was present in many historically important studies of binaural effects in narrow bands. In their coherence discrimination experiment Gabriel and Colburn (1981) found the just-noticeable difference to decrease with decreasing bandwidth given a reference coherence of 1.0. The decrease may result from particularly favorable noise samples that occur for narrow bands.

Parallels to incoherence detection appear in MLD experiments in connection with the detection of a low-frequency signal in the NoSπ condition. Van de Par and Kohlrausch (1999) found that as the bandwidth decreased, while remaining narrower than a critical band, threshold signal-to-overall-noise ratios remained approximately constant. The results of listeners D and M in Experiments 1 and 2 seem similar in that CAS values do not change much as the bandwidth changes from 14 to 108 Hz. By contrast, Zurek and Durlach (1987) found that the threshold signal to noise ratio decreases as the bandwidth decreases, a result that seems consistent with listeners P and W in Experiments 1 and 2 presented here. Zurek and Durlach (1987) interpreted the decrease in threshold that they observed as the result of binaural sluggishness and a consequent advantage for the relatively slow fluctuations in narrow bands. The results of our Experiment 4, using stimuli that matched the IPD and ILD fluctuations but not the fluctuation rate, support this interpretation too, at least for listeners P and W. However, our computational modeling for overall results (Goupell, 2005) does not support a role of binaural sluggishness. Instead, it supports a binaural system that is capable of responding to rapid fluctuations from a stable diotic standard stimulus as suggested by Hall *et al.* (1998) and references cited therein. For instance, Green (1966) found essentially no difference in binaural advantage for signals that were 1 s long and 10 ms long, and what difference there was suggested better binaural processing for 10 ms.

Although there are parallels between interaural incoherence detection and MLD tasks, there are also important differences. An NoSπ stimulus with a sine signal and noise masker may include both static and dynamic interaural cues. An interaural incoherence detection experiment, as presented here, involves dynamic fluctuations only. Results can be different too. Using multiplied noise maskers and sine signals Breebaart *et al.* (1999) found that cross-correlation accounted for their signal detection data better than various measures of interaural fluctuations, a result that is entirely contradictory to our experience with incoherence detection. Further, the experimental data obtained by Breebaart *et al.* were best fitted by a model that used the energy in the binaural difference signal as a decision variable. For stimuli such as ours, constructed from orthogonalized waveforms all having the same coherence, the energy in the difference signal is the same in all waveforms. Such a model would predict no difference in a listener's ability to detect the incoherence in any of our noise-pairs, contrary to our experiments.

In a MLD experiment the cues for detecting the signal are reported to be different depending on the bandwidth. For narrow bands an Sπ signal contributes a width to the binaural image. For wide bands an Sπ signal increases the tone-like strength (Evilsizer *et al.*, 2002). By contrast, in an incoherence detection experiment only the width cue occurs, whatever the bandwidth, though the rate of the fluctuations that establish the width does vary with bandwidth.

Experimental evidence in favor of this bandwidth effect on cues is that in a MLD experiment, performance for NoSo is correlated with performance for NoSπ for wide bands, but the correlation is significantly less for narrow bands (Evilsizer *et al.*, 2002.) (It should be noted that the bands called "narrow" by Evilsizer *et al.* are 100 Hz in width, equivalent to our "critical-band" noise-pairs.) See also Gilkey *et al.* (1985) and Isabelle and Colburn (1991) for similar evidence, as described by Evilsizer *et al.* Presumably this difference in correlation arises because listeners are using the tonal cue for wide bands, which is similar for So and Sπ, but are using the width cue for narrow bands, which has a binaural contribution only for Sπ.

Similarly, Evilsizer *et al.* found that interlistener correlation was strong for NoSo and also for wideband experiments for any combination of interaural noise and signal phases. Interlistener correlations dropped and became negative for NoSπ when the band was narrow (100 Hz). This result is consistent with the idea that the tonal cue is similarly detected by different listeners, but that there are marked individual differences when the cue becomes an image width for narrow bands (Bernstein *et al.*, 1998). The situation for incoherence detection is just the opposite. Interlistener correlations are strongest for narrow bands where the width cues vary greatly among different noise-pairs, and interlistener correlations are weak for wide bands where the width cues are similar for different pairs. A corollary to the above argument is that Evilsizer *et al.* would have found better correlation among listeners for Sπ in narrow bands had they used bands as narrow as ours.

In a MLD experiment with noise bands that are not narrow, a listener might listen in different subbands where the values of interaural coherence will be very different. In an incoherence detection experiment the stimulus is simpler because the coherence is spectrally homogeneous and advantageous sub-band listening possibilities are limited to special cases for individual noise samples as deviations from the ensemble mean statistics. In our view, the incoherence detection experiment has the advantage that it extracts the essential binaural element from a MLD experiment and exposes it for observation.

Experiment 1 for narrow-band stimuli clearly showed the importance of the dynamic fluctuations in interaural parameters, in contrast to the coherence measure of interaural differences. Experiment 1 gained its power from a selection of stimuli based on values of the standard deviation of IPD and ILD over time. These measures were plausible guesses about what is important to the binaural system, but there is no reason to expect them to be optimum measures of interaural fluctuations. In fact, it is clear that they are not optimum. Stimulus #57 in the 14-Hz phase set had one of the largest phase fluctuations and yet Fig. 2(a) shows that all four listeners had a relatively difficult time with it. Another article, to follow, explores alternative models for the detection of incoherence in narrowband noises and attempts to determine the best characterization of fluctuations from a perceptual point of view.

ACKNOWLEDGMENTS

The authors are grateful to Dr. H. S. Colburn and Dr. C. Trahiotis for useful discussions about coherence. Dr. N. I. Durlach and Dr. J. F. Culling provided important comments as did Associate Editor Dr. A. Kohlrausch. This work was supported in part by the National Institute on Deafness and Other Communicative Disorders, Grant No. DC00181.

APPENDIX: VARIATION AS FUNCTION OF BANDWIDTH

The mean values and standard deviations of $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ computed across noise-pairs are given in in the text blocks of Figs. 1, 3, and 5 for a collection of 100 noise-pairs and for bandwidths of 14, 108, and 2394 Hz respectively. It can be seen that as bandwidth increases, the standard deviation decreases. The purpose of this appendix is to provide more precise values of the means and standard deviations and to fit a functional form to the variation of $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ as a function of bandwidth.

For each value of bandwidth, 5000 noise-pairs were generated using a mixture of two orthogonalized noises as described in Experiments 1–3. These noise-pairs included the same spectral envelope and temporal windowing as in the experiments. The results are given in Table V.

A power regression was used to find a power law describing the variation of $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ as a function of bandwidth. The power regression was of the form $y=mx^p$ where the bandwidth is the x-variable and the standard deviation of $s_t[\Delta\Phi]$ or $s_t[\Delta L]$ is the y-variable. The line of best fit was

$$y = 18.6x^{-0.37} \tag{A1}$$

for $s_t[\Delta\Phi]$ and

$$y = 1.93x^{-0.36} \tag{A2}$$

for $s_t[\Delta L]$. Both fits described over 99% of the variance of the points. In summary, the standard deviations of the IPD and ILD computed over 500 ms follow distributions that have standard deviations computed over 5000 different noise-pairs that vary approximately the as inverse cubed-root of the bandwidth.

TABLE V. Values of the mean and standard deviation of $s_t[\Delta\Phi]$ and $s_t[\Delta L]$ for noise-pairs with three bandwidths: 14, 108, and 2394 Hz. Correlation between the standard deviations is also given. Each value is based on 5000 noise-pairs.

BW (Hz)	$\mu(s_t[\Delta\Phi])$ (deg)	$\mu(s_t[\Delta L])$ (dB)	$\sigma(s_t[\Delta\Phi])$ (deg)	$\sigma(s_t[\Delta L])$ (dB)	Correlation
14	10.78	1.45	6.95	0.77	0.80
108	13.19	1.71	3.40	0.35	0.73
2394	13.56	1.75	1.04	0.12	0.41

[1]In the description of the orthogonalization procedure by Culling *et al.* (2001), the denominator of Eq. (A4) is missing a factor of N.

[2]In the CAS measure, a confident response is given a weight of 1, the same weight as a correct response. The P_c measure corresponds to a confidence weight of zero. Goupell's thesis (2005) studied the effect of varying the weight and showed that interlistener correlation is insensitive to the value of the weight in the vicinity of a weight of 1. Further, the results of significance tests do not change as the weight increases past 1. This insensitivity tends to validate the CAS technique as applied here.

Barron, M. (**1983**). "Objective measures of spatial impression in concert halls," Proceedings of the Sixth International Congress on Acoustics, Paris, Vol. **7**, pp. 105–108.
Barron, M. (**2004**). "The current status of spatial impression in concert halls," Proceedings of the 18th International Congress on Acoustics, Kyoto, Th2.B1.1, Vol. **IV**, pp. 2449–2452.
Barron, M. and Marshall, A. H. (**1981**). "Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure," J. Sound Vib. **77**, 211–232.
Beranek, L. L. (**2004**). *Concert Halls and Opera Houses—Music, Acoustics, and Architecture* (Springer, New York).
Bernstein, L. R., Trahiotis, C., and Hyde, E. L. (**1998**). "Inter-individual differences in binaural detection of low-frequency tonal signals masked by narrowband or broad-band noise," J. Acoust. Soc. Am. **103**, 2069–2078.
Blauert, J., and Lindemann, W. (**1986**). "Auditory spaciousness: Some further psychoacoustic analyses," J. Acoust. Soc. Am. **80**, 533–542.
Bradley, J. S., and Soulodre, G. A. (**1995**). "Objective measures of listener envelopment," J. Acoust. Soc. Am. **98**, 2590–2597.
Breebaart, J., and Kohlrausch, A. (**2001**). "The influence of interaural stimulus uncertainty on binaural signal detection," J. Acoust. Soc. Am. **109**, 331–345.
Breebaart, J., van de Par, S., and Kohlrausch, A. (**1999**). "The contribution of static and dynamically varying ITDs and IIDs to binaural detection," J. Acoust. Soc. Am. **106**, 979–992.
Culling, J. F., Colburn, H. S., and Spurchise, M. (**2001**). "Interaural correlation sensitivity," J. Acoust. Soc. Am. **110**, 1020–1028.
Domnitz, R. H., and Colburn, H. S. (**1976**). "Analysis of binaural detection models for dependence on interaural target parameters," J. Acoust. Soc. Am. **59**, 598–601.
Evilsizer, M. E., Gilkey, R. H., Mason, C. R., Colburn, H. S., and Carney, L. H. (**2002**). "Binaural detection with narrowband and wideband reproducible noise maskers. I. Results for human," J. Acoust. Soc. Am. **111**, 336–345.
Gabriel, K. J., and Colburn, H. S. (**1981**). "Interaural correlation discrimination. I. Bandwidth and level dependence," J. Acoust. Soc. Am. **69**, 1394–1401.
Gilkey, R. H., Robinson, D. E., and Hanna, T. E. (**1985**). "Effects of masker waveform and signal-to-masker phase relation on diotic and dichotic masking by reproducible noise," J. Acoust. Soc. Am. **78**, 1207–1219.
Goupell, M. J. (**2005**). "The use of interaural parameters during incoherence detection in reproducible noise," Ph.D. dissertation, Michigan State University.
Green, D. M. (**1966**). "Interaural phase effects in the masking of signals of different durations," J. Acoust. Soc. Am. **39**, 720–724.
Hall, J. W., Grose, J. H., and Hartmann, W. M. (**1998**). "The masking level difference in low-noise noise," J. Acoust. Soc. Am. **103**, 2573–2577.
Isabelle, S. K., and Colburn, H. S. (**1991**). "Detection of tones in reproducible narrow-band noise," J. Acoust. Soc. Am. **89**, 352–359.
Isabelle, S. K., and Colburn, H. S. (**2003**). "Binaural detection of tones masked by reproducible noise: Experiments and models," J. Acoust. Soc. Am. (submitted).
Jeffress, L. A. (**1948**). "A place theory of sound localization," J. Comp. Physiol. **41**, 35–49.
Koehnke, J., Colburn, H. S., and Durlach, N. I. (**1986**). "Performance in several binaural-interaction experiments," J. Acoust. Soc. Am. **79**, 1558–1562.
Kollmeier, B., and Gilkey, R. H. (**1990**). "Binaural forward and backward masking: Evidence for sluggishness in binaural detection," J. Acoust. Soc. Am. **87**, 1709–1719.
Pollack, I., and Trittipoe, W. J. (**1959a**). "Binaural listening and interaural noise cross correlation," J. Acoust. Soc. Am. **31**, 1250–1252.
Pollack, I., and Trittipoe, W. J. (**1959b**). "Internal noise correlation: Examination of variables," J. Acoust. Soc. Am. **31**, 1616–1618.
van de Par, S., and Kohlrausch, A. (**1999**). "Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters," J. Acoust. Soc. Am. **106**, 1940–1947.
Wilbanks, W. A. (**1971**). "Detection of a narrow-band noise as a function of the interaural correlation of both signal and masker," J. Acoust. Soc. Am. **49**, 1814–1817.
Wilbanks, W. A., and Whitmore, J. K. (**1967**). "Detection of monaural signals as a function of interaural noise correlation and signal frequency," J. Acoust. Soc. Am. **43**, 785–797.
Zurek, P. M., and Durlach, N. I. (**1987**). "Masker-bandwidth dependence in homophasic and antiphasic tone detection," J. Acoust. Soc. Am. **81**, 459–464.