The influence of interaural stimulus uncertainty on binaural signal detection

Jeroen Breebaart
IPO, Center for User-System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands

Armin Kohlrausch
IPO, Center for User-System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands and Philips Research Laboratories Eindhoven, Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands

(Received 24 August 1998; revised 15 June 1999; accepted 31 August 2000)

This paper investigated the influence of stimulus uncertainty in binaural detection experiments and the predictions of several binaural models for such conditions. Masked thresholds of a 500-Hz sinusoid were measured in an NρSπ condition for both running and frozen-noise maskers using a three interval, forced-choice (3IFC) procedure. The nominal masker correlation varied between 0.64 and 1, and the bandwidth of the masker was either 10, 100, or 1000 Hz. The running-noise thresholds were expected to be higher than the frozen-noise thresholds because of stimulus uncertainty in the running-noise conditions. For an interaural correlation close to +1, no difference between frozen-noise and running-noise thresholds was expected for all values of the masker bandwidth. These expectations were supported by the experimental data: for interaural correlations less than 1.0, substantial differences between frozen and running-noise conditions were observed for bandwidths of 10 and 100 Hz. Two additional conditions were tested to further investigate the influence of stimulus uncertainty. In the first condition a different masker sample was chosen on each trial, but the correlation of the masker was forced to a fixed value. In the second condition one of two independent frozen-noise maskers was randomly chosen on each trial. Results from these experiments emphasized the influence of stimulus uncertainty in binaural detection tasks: if the degree of uncertainty in binaural cues was reduced, thresholds decreased towards thresholds in the conditions without any stimulus uncertainty. In the analysis of the data, stimulus uncertainty was expressed in terms of three theories of binaural processing: the interaural correlation, the EC theory, and a model based on the processing of interaural intensity differences (IIDs) and interaural time differences (ITDs). This analysis revealed that none of the theories tested could quantitatively account for the observed thresholds. In addition, it was found that, in conditions with stimulus uncertainty, predictions based on correlation differ from those based on the EC theory. [DOI: 10.1121/1.1320472]

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

I. INTRODUCTION

For a period of more than 50 years, the phenomenon of the binaural masking level difference (BMLD) has intrigued psychoacousticians. It has been shown that the interaural correlation of both the masker and the signal are important parameters influencing binaural detection thresholds. For example, when a low-frequency out-of-phase sinusoid is added to an in-phase broadband noise masker (NoSπ condition), the threshold of audibility is up to 15 dB lower compared to that for an in-phase sinusoidal signal (i.e., NoSo condition, cf. Hirsh, 1948; Hafter and Carrier, 1969; Zurek and Durlach, 1987). If the signal has an interaural correlation of +1 and an out-of-phase masker is used (i.e., an NπSo condition), BMLDs of up to 12 dB are reported (Jeffress *et al.*, 1952, 1962; Breebaart *et al.*, 1998).

In experiments where the masker correlation was varied between −1 and +1 using Sπ signals, Robinson and Jeffress (1963) found a monotonic increase in the BMLD with increasing interaural correlation. Small reductions from +1 of the interaural masker correlation in an NρSπ condition led to a large decrease of the BMLD, while for smaller correlations, the slope relating BMLDs to interaural correlation was shallower. The stimuli used by Robinson and Jeffress (1963) were composed by adding interaurally correlated noise with an interaurally uncorrelated noise. The relative intensities of both sources determined the *mean* interaural correlation. The consequence of this method for generating the stimuli is that for the masking noise alone, the interaural cues (i.e., interaural time- and intensity differences) fluctuate randomly. Moreover, because finite-length masker samples are used, the actual correlation within an observation interval can deviate considerably from the adjusted mean correlation. Thus, in terms of binaural cues, the masker contains uncertainty. The addition of the Sπ signal results in a change in the mean of the interaural cues but does not reduce the randomness of the interaural cues.

Analogous to monaural conditions (Lutfi, 1990), binaural masking can be attributed to two different sources. The first results from the limited resolution of the binaural auditory system and has been termed energetic masking by Lutfi. In models of binaural processing, this source of masking is included as internal noise. For example, the EC theory summarizes the internal errors of timing and amplitude represen-

tation in the factor *k*, which is directly related to the BMLD. The second source of masking results from the uncertainty associated with the trial-to-trial variation of the binaural cues used to detect the signal (called informational masking by Lutfi). This source of masking has not, as far as we are aware, been addressed explicitly so far in binaural models. In standard MLD conditions like NoSπ, the masker contains no uncertainty in terms of binaural cues: the interaural correlation is always exactly one, the energy of the difference signal between right and left masker is zero, and the interaural differences in time and intensity are always exactly zero.

Because it is well-known that the auditory system can benefit from the presence of binaural cues in a detection task, it is interesting to study the influence of uncertainty in these cues and the extent to which uncertainty limits detection. One of the possibilities to remove stimulus uncertainty is by using *frozen* noise. Thresholds for frozen binaural maskers would thus reflect the amount of energetic masking. The difference in detection performance between running noise and frozen noise indicates the amount of informational masking and this is the main topic of the present paper. Data will be presented that were measured under conditions with and without stimulus uncertainty. Three common theories for binaural processing (the change in interaural correlation, the EC theory and processing of interaural intensity differences, or IIDs, and interaural time differences, or ITDs) will be discussed for their ability to predict these data. We selected these theories because they are often used to explain binaural processing. In addition, they have been discussed recently for their ability to predict the amount of energetic masking in binaural conditions with non-Gaussian maskers (Breebaart *et al.*, 1999). Because, as we stated above, stimulus uncertainty has not been analyzed so far in terms of these models, we provide such an analysis in the following section.

II. STIMULUS UNCERTAINTY

A. Interaural correlation

It is often assumed that a change in the interaural correlation induced by adding a signal to a masker can be used as a detection cue in binaural masking experiments. Various mathematical details have been published treating changes in the interaural correlation for different experimental paradigms. For example, Domnitz and Colburn (1976) argued that models based on interaural correlation and models based on interaural differences yield similar predictions for NoSπ conditions with Gaussian noise. Breebaart *et al.* (1999) presented data with non-Gaussian noise maskers for which this close correspondence between the change in the cross correlation and the size of the interaural differences is no longer found. Durlach *et al.* (1986) determined an analytical expression for the interaural correlation in an NoSπ condition. Analytical expressions for the interaural waveform correlation and the interaural envelope correlation were derived by van de Par and Kohlrausch (1995) for NoSπ and later also for NoSm (van de Par and Kohlrausch, 1998). Bernstein and Trahiotis (1996) showed that for NoSπ stimuli, the interaural correlation of the stimuli after peripheral preprocessing did account for their NoSo vs NoSπ discrimination results for a wide range of center frequencies. Because this correlation approach is widely accepted, we will discuss stimulus uncertainty first in terms of this concept.

An interaurally partially correlated noise can be generated by adding an interaurally correlated noise $[N_0(t)]$ and an interaurally out-of-phase noise $[N_\pi(t)]$. In the following we assume that these two independent noise sources have the same rms value. To end up with a long-term normalized interaural correlation of ρ, the left-ear signal $L(t)$ and the right-ear signal $R(t)$ are combined as follows:

$$\begin{aligned} L(t) &= \tfrac{1}{2}\sqrt{2}\sqrt{1+\rho}N_0(t) + \tfrac{1}{2}\sqrt{2}\sqrt{1-\rho}N_\pi(t), \\ R(t) &= \tfrac{1}{2}\sqrt{2}\sqrt{1+\rho}N_0(t) - \tfrac{1}{2}\sqrt{2}\sqrt{1-\rho}N_\pi(t). \end{aligned} \tag{1}$$

Because both N_0 and N_π stem from random processes, the short-term energy estimates (i.e., integrated over one interval in a 3IFC task) of N_0 and N_π, E_0 and E_π, respectively, can deviate substantially from their expected (i.e., long-term) values provided that the product of time and bandwidth is small. Furthermore, the samples taken from the two noise sources can be partially correlated. Fluctuation of the short-term estimate of the noise energy leads to a variability in the interaural correlation for a finite-length noise interval [see the Appendix, Eq. (A2), and Gabriel and Colburn, 1981; Richards, 1987]. The interaural correlation of a finite sample will be referred to as *effective correlation*, ρ_{eff}, while the mean interaural correlation (i.e., the expected value of ρ_{eff}) will be referred to as *reference correlation*, ρ.

We determined the probability distribution for the interaural correlation for an NρSπ condition as a function of the signal-to-noise ratio, the bandwidth, and the duration of the masker. From the mathematical expressions for the effective correlation probability distribution as given in the Appendix, we found that three important factors affect the distribution for the effective correlation.

Bandwidth and duration of the noise. With increasing duration and bandwidth, the variance of the effective correlation, ρ_{eff}, will decrease as a result of the decreasing variances of E_0 and E_π (see the Appendix).

The reference correlation. For a reference correlation of $+1$ (and -1), there is no correlation uncertainty, and the effective correlation will always be $+1$ (-1). On the other hand, for reference correlations between -1 and $+1$, the effective correlation will follow a distribution rather than have a fixed value. For a reference correlation close to zero, the width of the effective correlation distribution will be widest (i.e., the correlation uncertainty is maximum).

The presence or absence of the signal. The addition of an Sπ signal results in a shift of the mean interaural correlation towards -1.

To demonstrate the effect of these properties upon the correlation uncertainty, probability density function (PDFs) for a 300-ms noise and three different combinations of reference correlation and noise bandwidth are shown in Fig. 1. Each panel shows two distributions, one for the noise alone (solid line) and one for noise plus signal (dashed line). The signal had a duration of 200 ms, was temporally centered in the noise, and had a level of 10 dB below the masker level (i.e., S/N$=-10$ dB). It is clear that the *width* of the PDF

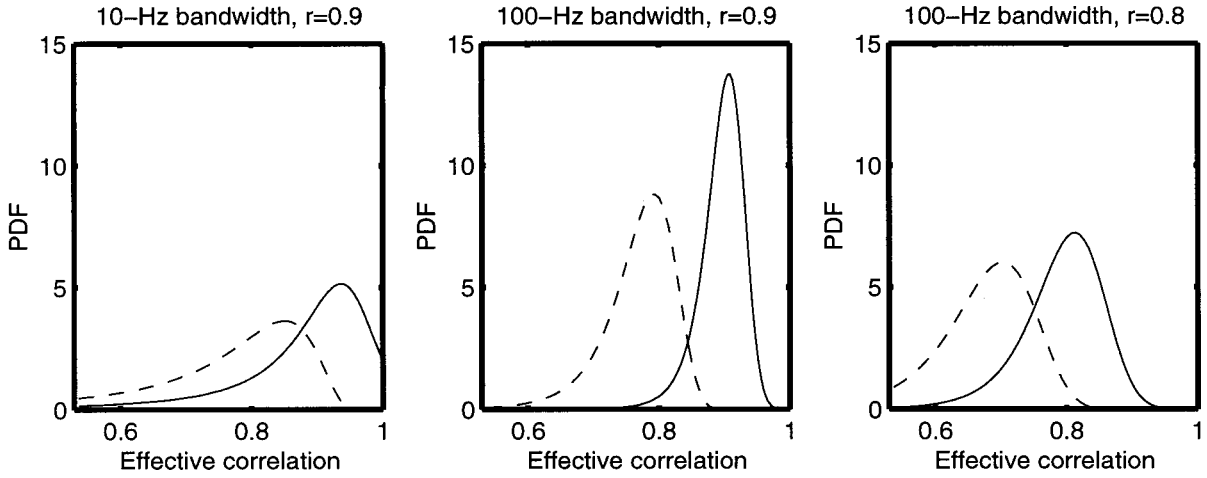

FIG. 1. Probability density functions for the effective interaural correlation. The left panel corresponds to a bandwidth of 10 Hz and a reference correlation of 0.9. For the middle and right panels, these parameters were 100 Hz, 0.9 and 100 Hz, 0.8, respectively. The solid lines represent a 300-ms masker alone; the dashed lines represent a 300-ms masker with a 200-ms Sπ signal. The distributions were calculated for the complete 300-ms interval. The signal-to-masker ratio was -10 dB.

decreases with increasing bandwidth (10 Hz in the left panel, 100 Hz in the middle and right panels) and with increasing correlation (0.9 in the left and middle panels, 0.8 in the right panel), indicating less correlation uncertainty. Furthermore, the peak of the curve migrates towards higher correlation values for decreasing bandwidth. This results from the fact that the interaural correlation is a nonlinear function of the noise energies E_π and E_o. If the mean interaural correlation is set to 0.9 and the bandwidth is 10 Hz, it can be observed in the left panel of Fig. 1 that there is a finite probability for correlation values to be smaller than 0.8, a value that differs by more than 0.1 from the mean interaural correlation. This property is highly asymmetric; correlations higher than $+1$ cannot occur. If, despite this asymmetric property, the mean correlation is 0.9, the peak of the curve must occur at a correlation greater than 0.9.

The addition of the Sπ signal results in a shift of the curves towards lower correlation values. Furthermore, the distributions show a small increase in their widths. In the left panel (10-Hz bandwidth), the shift of the curve is small compared to the width of the distributions. Thus, from a signal-detection point of view, it is likely that at this signal-to-masker ratio, interaural correlation uncertainty can influence the detection performance. For a bandwidth of 100 Hz and a reference correlation of 0.9 (middle panel), the curves for masker alone and masker plus signal show a smaller overlap. If the reference correlation is reduced to 0.8 (right panel), the amount of overlap is increased. We can conclude that both the bandwidth and the reference correlation of the noise have a strong effect on the detectability of the signal in terms of interaural correlation. If human observers indeed use the interaural correlation as a decision variable, thresholds should depend on the stimulus parameters that determine the amount of correlation uncertainty.

In fact, experimental data from Gabriel Colburn (1981) and van der Heijden and Trahiotis (1998) confirm this hypothesis. Gabriel and Colburn (1981) found that if the bandwidth of a noise stimulus is increased from 5 to 1000 Hz, the interaural correlation just noticeable difference (jnd) for a reference correlation of 0 decreases by a factor of 2. Moreover, the change in the correlation jnd was largest for bandwidth below the critical bandwidth. For masker bandwidths beyond the critical bandwidth, the correlation jnd did not change by a large amount. This might indicate that the internal interaural correlation is evaluated after filtering in the periphery of the auditory system. For a reference correlation of $+1$, the correlation jnd remained approximately constant for bandwidths up to the critical bandwidth. However, an increase in the correlation jnd was observed when the bandwidth was increased well beyond this value. Although critical band filtering seems to play a role under these conditions, this increase in thresholds is not yet understood. The data obtained by van der Heijden and Trahiotis (1998) showed that the correlation dependence of thresholds is much stronger at narrow bandwidth (3 Hz) than at large bandwidth (900 Hz). This corresponds to the notion that correlation uncertainty influences detection, because the probability density function for the correlation is wider at narrow bandwidths (see Fig. 1).

The consequences of the use of frozen noise upon correlation uncertainty are very simple. If exactly the same noise waveform is used in each trial and each token of a multiple-interval, forced-choice procedure, there is no uncertainty in the masker interval; the interaural correlation always has the same value. The addition of the Sπ signal results in a deviation from this fixed value. The actual value depends on the signal-to-noise ratio: a higher signal level results in a lower correlation.

B. The EC theory

Durlach's EC theory (Durlach, 1963) is another well-known theory to account for BMLDs. According to this theory the waveforms which arrive at both ears are modified by an interaural time delay and an interaural level adjustment in such a way that the masker waveforms are equalized (the E process). This process is performed imperfectly as a result of internal errors. Subsequently, the stimulus in one ear is

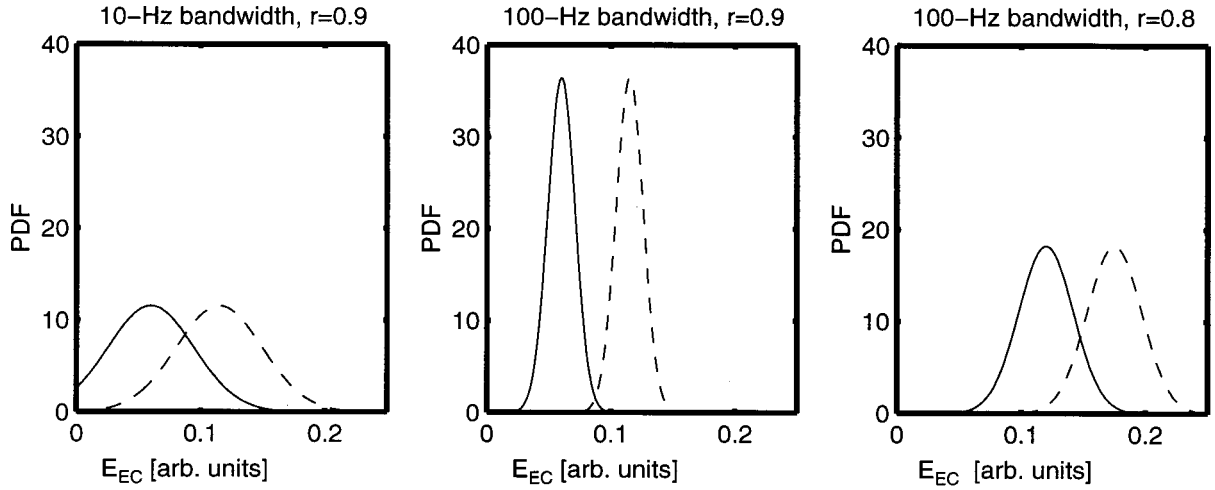

FIG. 2. Probability density functions for E_{EC} (see the text) for different values of the reference correlation and the masker bandwidth. The format is the same as in Fig. 1.

subtracted from the stimulus in the other ear (cancellation, or C process). In binaural conditions, this process often leads to an improvement of the signal-to-masker ratio and hence to the prediction of a positive BMLD. In an NρSπ condition no equalization is available that yields a signal-to-masker ratio improvement. Hence, the improvement in signal-to-masker ratio is obtained by calculating the amount of masker energy that is removed by the cancellation process. From Eq. (1), it can be seen that the common part of the masking noise will be removed and that the N_π masker portion remains. Thus, the amount of stimulus energy that remains after the EC-process equals[1]

$$E_{EC}=2(1-\rho)E_\pi+4E_S, \tag{2}$$

where E_S denotes the signal energy and E_{EC} is the energy of the difference signal between the left and right ears. The use of the difference energy as a decision variable was also suggested by Breebaart *et al.* (1999) for NoSπ stimuli with non-Gaussian maskers. If no signal is present, E_S is simply zero. Assuming that E_{EC} is used as a decision variable, stimulus uncertainty will influence the detection task because E_π is a random variable with a certain mean and standard deviation. Equations (A4) to (A6) in the Appendix give a description of the variability of E_π. A graphical representation of this description is shown in Fig. 2. The format is the same as in Fig. 1; the left and middle panels correspond to a reference correlation of 0.9, the right panel to 0.8. The bandwidth of the noise is 10 Hz in the left panel and 100 Hz in the other panels. Each panel contains two curves; the solid lines represent the PDF for E_{EC} for a masker alone, the dashed lines for masker plus signal. For simplicity it is assumed that the rms value of the noise sources equals 1 (arbitrary units) and the signal-to-masker ratio is -10 dB. The masker had a duration of 300 ms. The curves in Fig. 2 show a similar behavior as in Fig. 1; a wider bandwidth or a higher reference correlation results in a narrower distribution of E_{EC}, and hence a better detectability of the signal.

If a frozen-noise sample is used, E_π has a fixed value. Hence, no uncertainty in terms of the EC theory is present in the stimulus (the power of the difference signal is frozen) and the only limitation for detection is internal noise.

C. Interaural differences in time and intensity

The interaural differences (IIDs and ITDs) present in an interaurally partially correlated noise fluctuate as a function of time. In a running-noise condition, the random fluctuations can be described in terms of a probability distribution. We determined these probability distributions by computing a partially correlated noise in the digital domain of sufficient duration (3 s at a sample rate of 32 kHz). After a Hilbert transform of the left and right signals, the interaural intensity differences and interaural time differences were obtained. From these differences, histograms were computed which are (close) approximations of the PDFs of the IIDs and ITDs. This procedure was repeated for masker plus signal for a signal-to-masker ratio of -10 dB. The results are shown in Fig. 3. The format is the same as Figs. 1 and 2. The left panels correspond to a masker bandwidth of 10 Hz and a reference correlation of 0.9; the middle panels to a bandwidth of 100 Hz and a correlation of 0.9, and the right panels to a bandwidth of 100 Hz and a reference correlation of 0.8. The solid lines represent distributions for the maskers alone, the dashed lines for masker plus signal. The upper panels represent the interaural phase differences (IPD); the lower panels represent the IIDs.

The following facts can be observed in Fig. 3. First, if we compare the middle panels to the left panels (i.e., the effect of bandwidth), no difference is observed. Thus, the width of the PDF for the interaural differences does not depend on the bandwidth and the range of variation of the IIDs and ITDs does not change systematically with bandwidth. The *rate* of variation does, however, increase if the bandwidth is increased. This property is important for our hypothesis about stimulus uncertainty. It if often assumed that the binaural auditory system is sluggish in processing binaural cues (cf. Grantham and Wightman, 1978, 1979; Grantham, 1984; Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998). Thus, because the amount of uncertainty is *not*

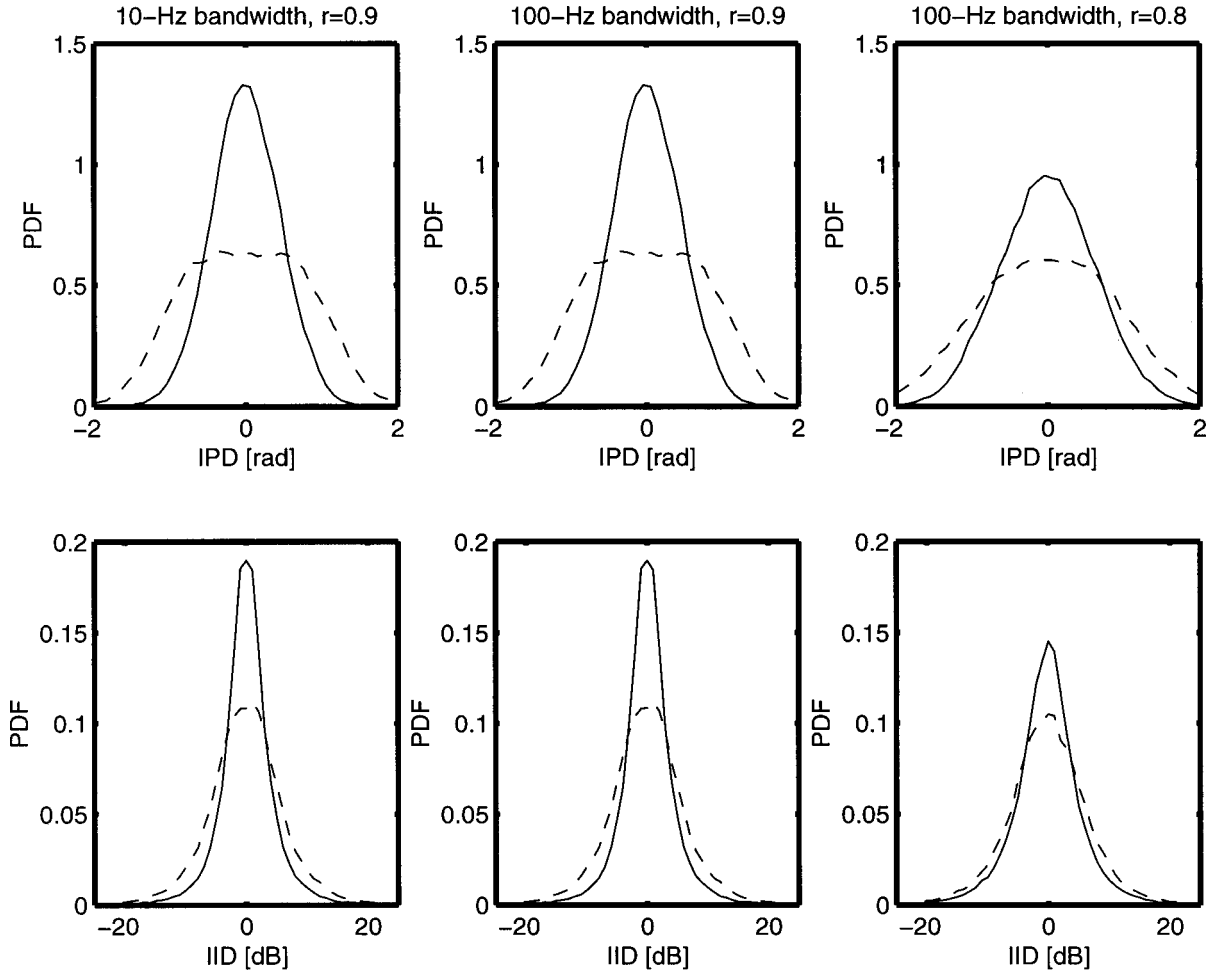

FIG. 3. Probability density functions for the IPD (upper panels) and IID (lower panels) in the same format as Figs. 1 and 2.

changed by the masker bandwidth, it is expected that thresholds will *increase* with increasing bandwidth as a result of the increase in the rate of fluctuation of the IIDs and ITDs. This is in contrast to the expectations based on the EC theory or models based on interaural correlation; these models predict a *decrease* with increasing bandwidth.

The solid line in the right panel of Fig. 3 demonstrates that a decreasing interaural correlation results in an increase in the width of the PDF, a similar effect as observed in the curves for the correlation and the EC theory. The addition of the signal has a different effect on the PDFs compared to the two other models discussed in this paper. Instead of a shift of the mean, an increase in the width of the distribution is observed. This property makes it more difficult to analyze these PDFs in terms of detectability. Nevertheless, the observation that a change in the bandwidth results in different expectations for the three theories makes it valuable also to discuss our data in terms of IIDs and ITDs.

III. EXPERIMENT I

A. Procedure and stimuli

A three-interval, forced-choice procedure with adaptive signal-level adjustment was used to determine masked thresholds. Three masker intervals of 300-ms duration were separated by pauses of 300 ms. A signal of 200-ms duration was added to the temporal center of one of the masker intervals. Feedback was provided after each response of the subject.

The signal level was adjusted according to a two-down, one-up rule (Levitt, 1971), tracking the 70.7% correct score within a 3IFC paradigm. This corresponds to $d' = 1.26$. The initial step size for adjusting the level was 8 dB. The step size was halved after every second track reversal until it reached 1 dB. The run was then continued for another eight reversals. The median level at these last eight reversals was used as the threshold value. At least three threshold values were obtained for each parameter value and subject. All stimuli were generated digitally and converted to analog signals with a two-channel, 16-bit D/A converter at a sampling rate of 32 kHz. The stimuli were presented over Beyer Dynamic DT990 headphones.

The 300-ms masker samples were obtained by adding interaurally in-phase noise and interaurally out-of-phase noise with the appropriate weighting factors [Eq. (1)]. For running-noise conditions, the noise samples for each interval were obtained by randomly selecting 300-ms segments from a two-channel, 2000-ms bandpass-noise buffer. The 2000-ms noise buffer was created in the frequency domain by selecting the desired frequency range from the Fourier transforms of two independent 2000-ms broadband Gaussian noises. After an inverse Fourier transform, and combination of the two

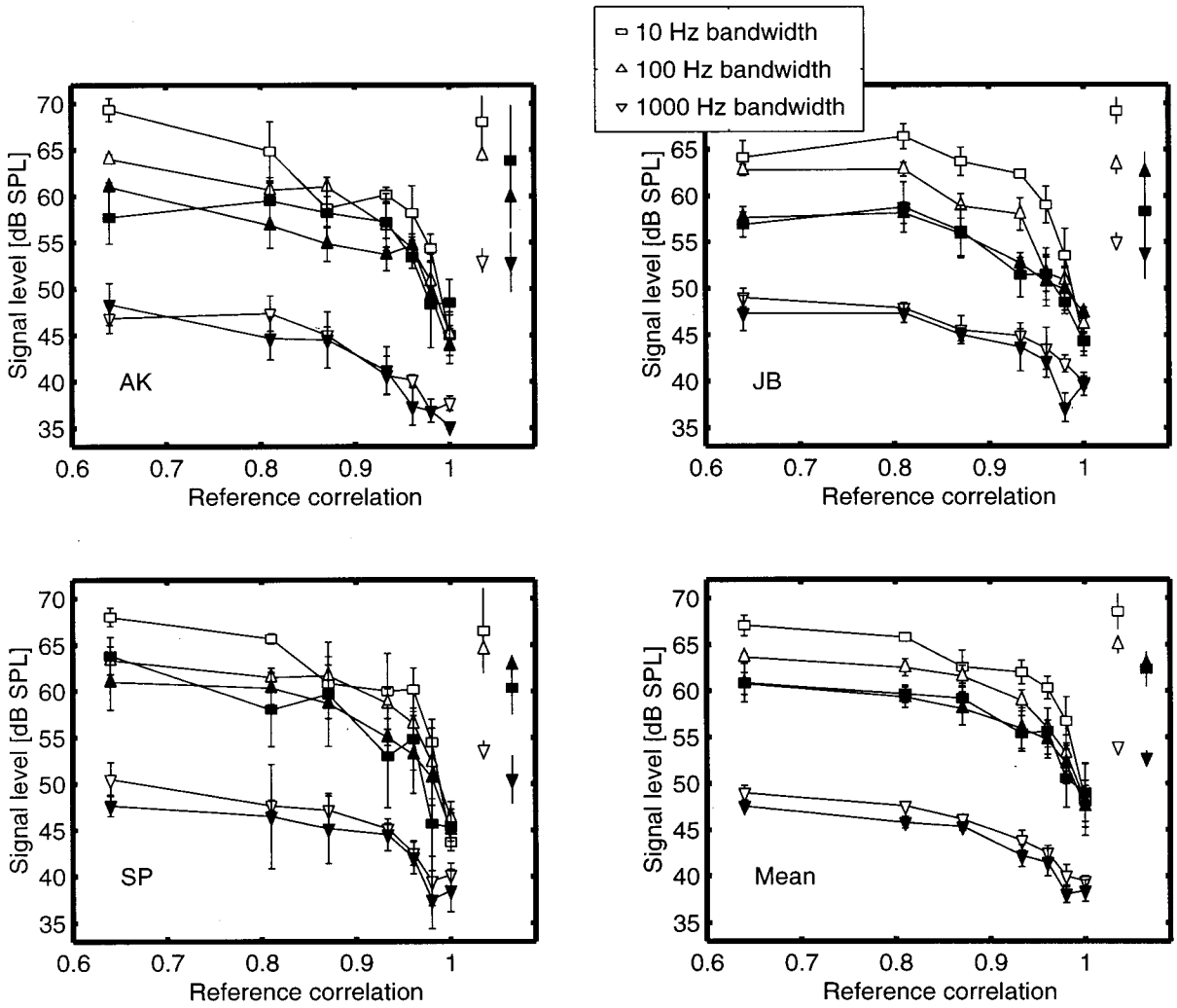

FIG. 4. Binaural masked thresholds as a function of the reference correlation. The bottom-right panel shows the mean thresholds; the other panels show individual thresholds for three subjects. The squares correspond to a masker bandwidth of 10 Hz, the upward triangles to 100 Hz, and the downward triangles to 1000 Hz. The open symbols represent running-noise conditions; the filled symbols represent frozen-noise conditions. The isolated symbols represent NoSo reference data. Error bars denote the standard error of the mean. The masker level for all three bandwidths was 65 dB.

noise signals according to Eq. (1), the two-channel (for the left and right ears) bandlimited noise buffer with the specified reference correlation was obtained. It is important to note that the specified reference correlation is the correlation of the 2000-ms noise buffer. The correlations of shorter segments (like the 300-ms noise segments used in the running-noise experiments) will in general deviate from this exact value (see Fig. 1).

For frozen-noise conditions, only one fixed 300-ms noise sample was used for which the interaural correlation was equal to the reference correlation.[2] This noise sample was generated by adding two independent bandlimited noise samples of 300 ms with a *fixed* rms value. These bandlimited noise samples were generated in the same way as the noise buffers for random-noise conditions followed by a normalization of their rms values. The partially correlated noise was then generated by combining the noises according to Eq. (1). The same noise sample was used during one run. To exclude the possibility that the frozen-noise thresholds would depend on the specific waveform of the token, a different frozen-noise sample was used for each run, and the mean threshold from these runs was used as threshold value. All noise maskers were presented at an overall level of 65 dB SPL.

The 200-ms signals were interaurally out-of-phase sinusoids with a frequency of 500 Hz. In order to avoid spectral splatter, the signals and the maskers were gated with 50-ms raised-cosine ramps. Thresholds are expressed as the means of at least three repetitions per condition and subject. Binaural masked thresholds were measured for $N\rho S\pi$ conditions, where the bandwidth of the noise was either 10, 100, or 1000 Hz. The center frequency of the noise masker was always 500 Hz. Reference correlations of $\rho = 1$, 0.98, 0.96, 0.93, 0.87, 0.81, and 0.64 were used. In addition, NoSo thresholds were also obtained. Three well-trained subjects with normal hearing participated in the experiments.

B. Results

The experimental data are shown in Fig. 4 as a function of the reference correlation. The bottom-right panel shows the mean thresholds, while the other panels show individual thresholds for the three subjects. The squares correspond to a

masker bandwidth of 10 Hz, the upward triangles correspond to 100 Hz, and the downward triangles to 1000 Hz. The open symbols represent running-noise conditions; the filled symbols represent frozen-noise conditions. The NoSo thresholds are plotted in the upper-right corners of each panel. The error bars denote the standard error of the mean.

For both running- and frozen-noise conditions, the NρSπ thresholds increase with decreasing reference correlation. This increase is strongest for the 10-Hz running-noise masker, which increases by 18.8 dB if the correlation is decreased from +1 to 0.64. For the 100-Hz-wide and 1000-Hz-wide running-noise conditions, the increase amounts to 15.5 and 10 dB, respectively. These values are in good agreement with data from van der Heijden and Trahiotis (1998). For frozen-noise maskers, the increase amounts to 12, 13.3, and 9.2 dB, for the 10-, 100-, and 1000-Hz-wide conditions, respectively.

The thresholds for frozen and running-noise maskers are approximately equal for a reference correlation of +1, while for decreasing reference correlations, the difference between frozen and running-noise maskers increases, especially for the narrow-band conditions. The reference correlations at which frozen and running-noise thresholds become different are 0.98 for a bandwidth of 10 Hz and 0.93 for a bandwidth of 100 Hz. As interaural correlation decreases the differences between running- and frozen-noise conditions reach 7 dB for the 10-Hz-wide masker, and 4 dB for the 100-Hz-wide condition. For the 1000-Hz-wide maskers, the thresholds for running and frozen noise are very similar.

The differences between the 100-Hz and 1000-Hz conditions vary considerably across reference correlations, both for running and frozen-noise conditions. For running noise, the difference in thresholds amounts to 9 dB for a reference correlation of +1 and increases up to a value of 14 dB for a reference correlation of 0.64. For frozen noise, these values are 10.6 and 13.3 dB, respectively. Because the overall masker level was kept constant, a difference of about 10 dB between 100-Hz and 1000-Hz thresholds would correspond to a constant signal-to-noise ratio at the output of an auditory filter with a bandwidth of 78 Hz [1 equivalent rectangular bandwidth (ERB) at 500 Hz, Glasberg and Moore, 1990].

The NoSo thresholds for *running* noise show a decrease with increasing masker bandwidth. The signal-to-noise ratio decreases from +4 dB at 10 Hz to 0 dB at 100 Hz and finally to −11 dB at 1000 Hz, very similar to experimental data from van de Par and Kohlrausch (1997, 1999). In contrast, for *frozen* noise, the NoSo thresholds are very similar for the 10- and 100-Hz bandwidth (S/N$=-2$ dB), while for the 1000-Hz bandwidth, the threshold is 10 dB lower. In general, the relation between running- and frozen-noise thresholds in the NoSo condition equals that for NρSπ with $\rho<0.95$. For the smallest reference correlation (i.e., $\rho=0.64$), the running-noise BMLD for a bandwidth of 10 and 100 Hz is almost zero (except for subject JB, who has a BMLD of 5 dB for this condition). For the 1000-Hz-wide condition, the BMLD is 6 dB for $\rho=0.64$, consistent with data from Robinson and Jeffress (1963).

C. Discussion

Following our hypothesis that stimulus uncertainty influences NρSπ thresholds, the difference between frozen and running-noise conditions should be larger at lower reference correlations, as a result of the fact that stimulus uncertainty increases with decreasing correlation. In addition, frozen- and running-noise thresholds should be equal for a reference correlation of +1, because no uncertainty in terms of binaural cues is present in the masker intervals. These effects are clearly visible in our data (Fig. 4) for the bandwidths of 10 and 100 Hz. For a bandwidth of 1000 Hz, there is almost no difference between the running and frozen-noise conditions for all values of the reference correlation. This suggests that for this value of the masker bandwidth, stimulus uncertainty does not influence the detection of the signal and the thresholds are limited by internal noise (similar to the frozen-noise data). Interestingly, the data that are limited by internal errors also show a dependence on the masker correlation. This implies that the net effect of the internal noise must be larger for smaller interaural correlations. One possibility to implement this property in a quantitative binaural model is given in Breebaart *et al.* (2001).

For the quantitative analysis of our data in terms of the binaural models described in Sec. II, we will concentrate on those conditions where, presumably, external variability is dominant over internal noise. In terms of the terminology used by Lutfi (1990), we are interested in conditions with a substantial amount of informational masking. As a measure for this we take the difference between running- and frozen-noise thresholds. Substantial differences are observed for a 10-Hz-wide masker and reference correlations at or below 0.98, and for a 100-Hz-wide masker at or below 0.93. For the 1000-Hz data, the difference is small at all correlation values and those data are therefore not included in the analysis.

The influence of external variability in the binaural data is stronger at 10 Hz than at 100-Hz bandwidth. This supports the expectations based on the EC theory and on correlation uncertainty, because both models predict a stronger difference between frozen and running noise at narrower bandwidths. The data are, however, not in line with the expectations based on the evaluation of IID and ITD cues. The distributions of these cues do not depend on the bandwidth, and hence no effect is expected if uncertainty is considered. Including the effect of binaural sluggishness, this should lead to an increase in threshold with an increase in masker bandwidth. This, however, is not found in the data, which show a decrease of the running-noise thresholds with increasing bandwidth.

In order to verify the hypotheses based on the EC theory and the interaural correlation quantitatively, we computed the detectability of the running-noise thresholds shown in Fig. 4 based on the two models for those conditions in which detection is apparently limited by stimulus uncertainty (i.e., $\rho\leqslant 0.98$ at 10-Hz bandwidth and $\rho\leqslant 0.93$ at 100-Hz bandwidth). For the interaural correlation, the detectability was calculated using the distribution of the interaural correlation expressed in terms of Fisher's Z (see the Appendix). The rationale for the transformation from correlation to Z lies in the fact that the correlation probability does not have a

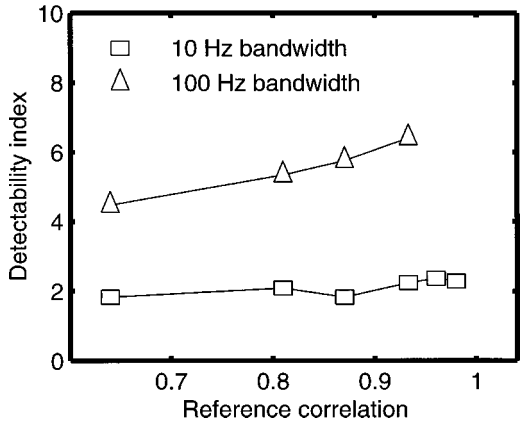

FIG. 5. Detectability index d' in terms of interaural correlation for the running-noise NρSπ conditions as a function of the reference correlation. The squares denote the 10-Hz-wide condition and the upward triangles the 100-Hz-wide condition.

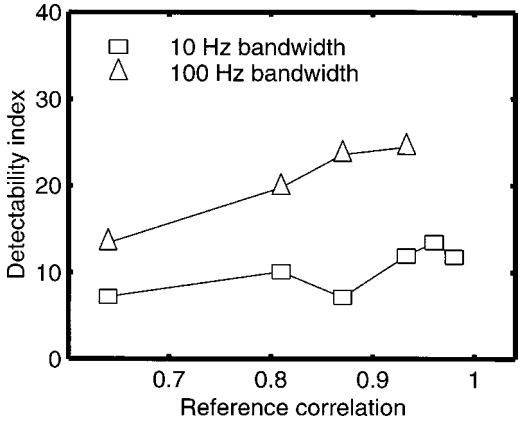

FIG. 6. Detectability index d' in terms of E_{EC} for the running-noise NρSπ conditions as a function of the reference correlation. The squares denote the 10-Hz-wide condition and the upward triangles the 100-Hz-wide condition.

Gaussian distribution, while Fisher's Z does have an approximately Gaussian distribution. For both the masker alone and the masker plus test signal at threshold, the probability distributions for Z were computed and from these distributions, the sensitivity index d' was calculated. For the sinusoidal signals, a length of 200 ms including 50-ms ramps was used to calculate the change in correlation. For the masker-alone correlation interval, a duration of 200 ms was assumed, because this corresponds to the signal length and hence the duration from which the binaural system can extract useful information concerning interaural correlation changes. Peripheral preprocessing was simulated by first filtering the signals with a fourth-order gammatone filter with a center frequency of 500 Hz and a bandwidth of 78 Hz (cf. Glasberg and Moore, 1990). The values for d' are shown in Fig. 5.

The squares denote the 10-Hz masker condition and the upward triangles the 100-Hz condition. Clearly, most values of d' are higher than the theoretical value of 1.26 that results from the applied procedure. The values of d' across reference correlations are relatively constant for 10-Hz bandwidth and increase systematically towards high reference correlations for 100-Hz bandwidth. Only the 10-Hz condition shows a fair agreement with d' in terms of the correlation uncertainty. From this simulation it appears that the correlation uncertainty is a valid statistic only for the 10-Hz-wide condition.

The large values of d' for the 100-Hz-wide conditions may indicate that, in the processing of these stimuli in the auditory system, information is lost. An optimal detector, basing its decision on the correlation change within the 200 ms of signal duration, would perform much better than the subjects, given the high values of d' for correlation discrimination. Such a loss of information might be caused by the fact that the subjects are not able to process the whole stimulus but extract a decision variable based on a shorter part of the sample.

Another possibility is that the correlation hypothesis is not correct and that detection behavior can be better described by another statistic, for example based on the EC theory. Equation (2) gives the relation between the decision variable E_{EC} and the source of stimulus uncertainty, E_π. With the help of Eqs. (A4) to (A6) and (A18) in the Appendix, the detectability index in terms of E_{EC} can be determined for the conditions limited by stimulus uncertainty. These indices are shown in Fig. 6, in the same format as in Fig. 5.

All values for d' in terms of E_{EC} are much higher than the theoretical value of 1.26. This indicates that E_{EC} is not a valid descriptor for the influence of stimulus variability.

An important remark can be made if the values for d' are compared for the correlation (Fig. 5) and the EC theory (Fig. 6). The values are completely different for these theories, the latter being much higher. This observation is particularly of interest given the analysis of Green (1992). He stated that a correlation model leads to identical predictions as an EC model in an NoSo vs NoSπ discrimination paradigm. Figures 5 and 6 show that this conclusion is not valid for conditions that are dominated by stimulus uncertainty.

In summary, both the 10-Hz-wide and the 100-Hz-wide conditions show large differences between the running and frozen-noise conditions, indicating that stimulus uncertainty dominates the detection process. This effect is smaller at 100 Hz than at 10-Hz bandwidth. The thresholds decrease with an increase in masker bandwidth in the running-noise conditions, which is not in line with expectations based on the processing of IID and ITD cues. An uncertainty analysis in terms of the EC theory revealed that an EC process fails to account for the thresholds found in the running-noise conditions.

The only close match between experimental data and predictions was found for the 10-Hz-wide conditions based on the interaural cross correlation. If one assumes that stimulus uncertainty limits the detection and the correlation is a valid statistic for describing thresholds, the psychometric function for an NρSπ condition as a function of the signal level can be predicted. To study to what extent this is true, a second experiment was performed, where predicted and measured psychometric functions were compared.

IV. EXPERIMENT II

A. Procedure and stimuli

In order to further examine the role of stimulus uncertainty in an NρSπ condition, we determined the psychomet-

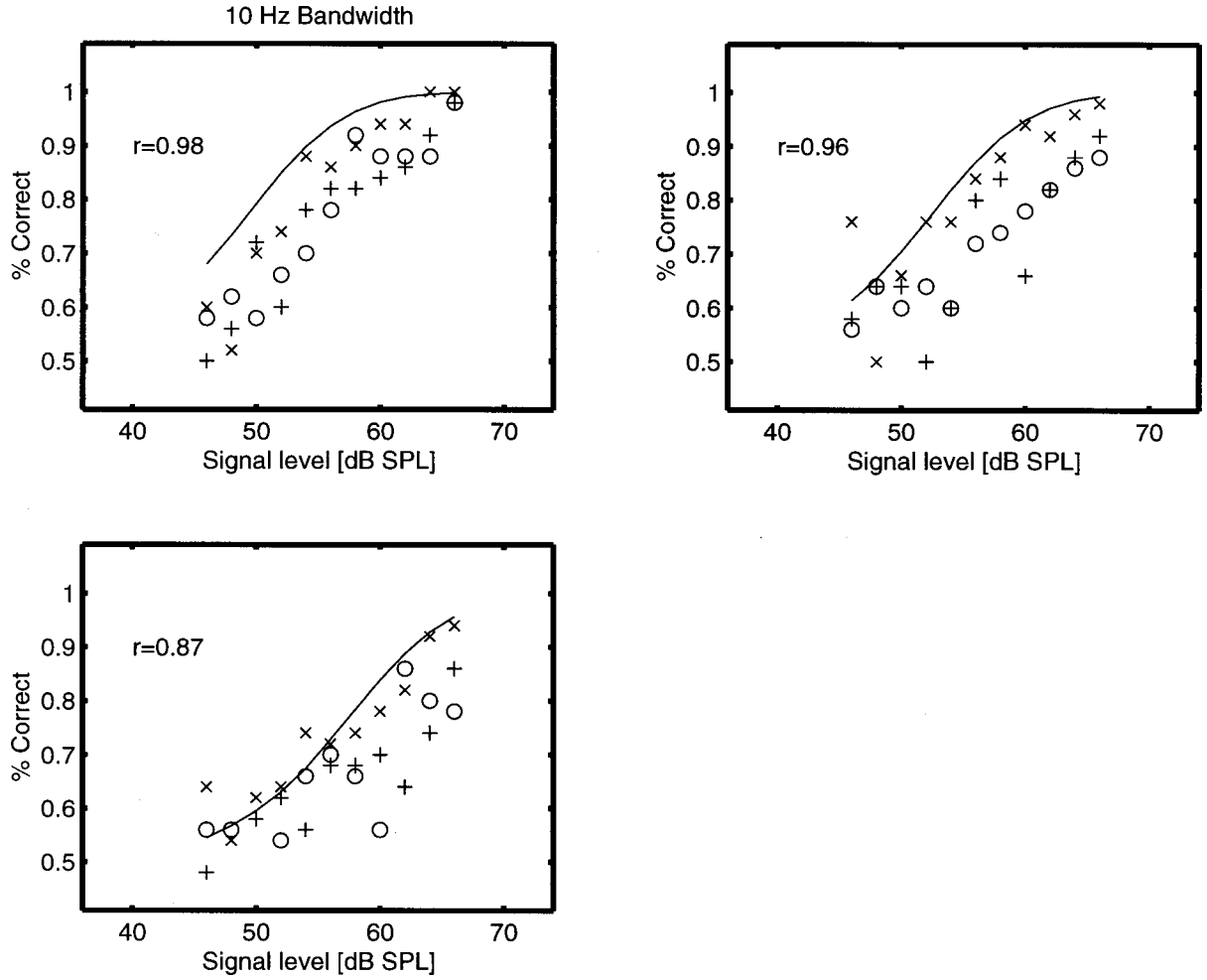

FIG. 7. Proportions correct as a function of the signal level for a running-noise NρSπ condition for $\rho=0.98$ (upper-left panel), $\rho=0.96$ (upper-right panel), and $\rho=0.87$ (lower-left panel). The different symbols represent different subjects. The bandwidth of the masker was 10 Hz. The solid line represents the predictions according to a correlation-uncertainty model (see the text).

ric functions for running-noise conditions at bandwidths of 10 and 100 Hz for reference correlations of 0.98, 0.96, and 0.87. Proportions correct were determined in a 2IFC procedure with 50 trials per condition by 3 subjects. The generation of the stimuli and the method of presentation to the subjects were similar to the method described in experiment I. The signal levels used to determine the subjects' performance were 46 to 66 dB SPL at 10-Hz bandwidth and 42 to 62 dB SPL at 100-Hz bandwidth with a step size of 2 dB.

B. Results

The proportions correct for the NρSπ condition as a function of the signal level are shown in Fig. 7 for a masker bandwidth of 10 Hz and in Fig. 8 for a bandwidth of 100 Hz. The different symbols denote different subjects. The upper-left panel represents data for $\rho=0.98$, the upper-right panel for $\rho=0.96$, and the lower panel for $\rho=0.87$. The data show an increase in the proportion of correct responses from 0.5 to 1 if the signal level is increased from 45 to 65 dB SPL at 10-Hz bandwidth and from 40 to 60 dB SPL at 100-Hz bandwidth. The solid lines represent the proportions correct based on the correlation probability density functions.

We calculated the predicted proportions correct as a function of the signal level based on the sensitivity index (d') determined from the correlation uncertainty. The values for d' were converted to proportions correct (p) by computing the area under the normal curve up to $d'/\sqrt{2}$ (see Green and Swets, 1966):

$$p=\int_{-\infty}^{d'/\sqrt{2}}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx. \tag{3}$$

The predicted proportions correct are shown by the solid lines in Figs. 7 and 8. For a bandwidth of 10 Hz, the curves lie on top, close to the subjects' responses, indicating that the data can quite accurately be described (especially for the subject denoted by ''x'') by the stimulus uncertainty in the interaural correlation. However, at 100-Hz bandwidth, the subjects perform worse than the predictions based on the correlation uncertainty. This indicates that correlation uncertainty is not a valid descriptor for the 100-Hz data.

C. Discussion

Because of the close correspondence between the predicted and observed psychometric functions for the 10-Hz-

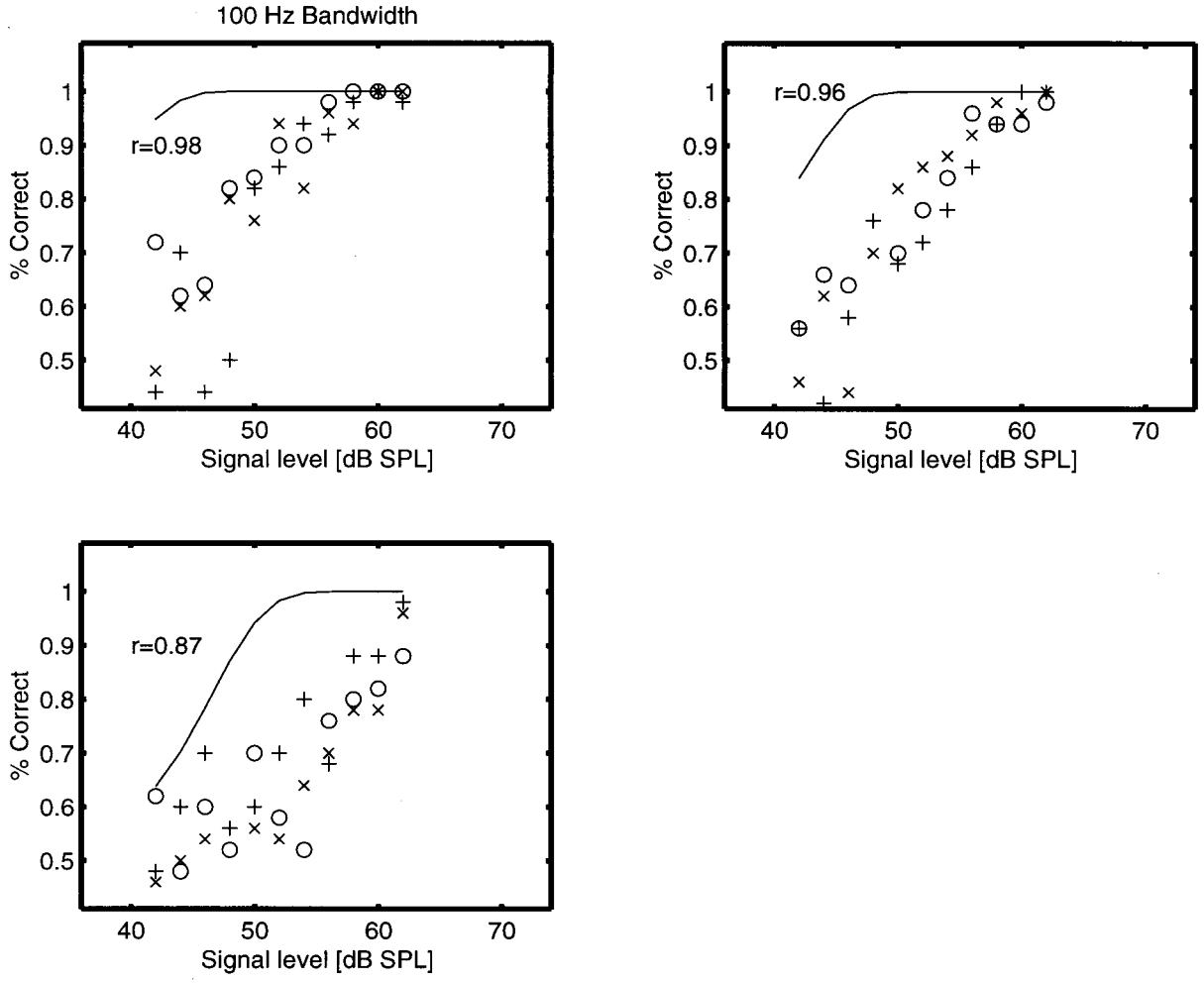

FIG. 8. Same as Fig. 7 for a masker bandwidth of 100 Hz.

wide maskers, it is likely that stimulus uncertainty limits the detection and that this uncertainty expressed in terms of the interaural correlation is a valid way to predict thresholds. For the 100-Hz condition, however, such an analysis overestimates the performance of the subjects.

We want to emphasize that although the interaural correlation is a valid detection statistic in *describing* the 10-Hz-wide running-noise conditions, this does not prove that observes indeed use this particular measure. We have shown, however, that stimulus uncertainty can play an important role in binaural detection paradigms. To further investigate the role of the interaural correlation as a detection statistic and the role of stimulus uncertainty, we performed a third experiment. This experiment is a compromise between the running-noise condition (i.e., with stimulus uncertainty) and the frozen-noise condition (i.e., absolutely no stimulus uncertainty). Two conditions were tested, which we refer to as ''fixed-ρ'' and ''interleaved.''

V. EXPERIMENT III

A. Procedure and stimuli

The following experimental paradigms were used:

(1) Fixed-ρ. In this condition, each trial consisted of three intervals which contained exactly the same noise sample. To one of these noise samples, the signal was added. For each trial, a different noise sample was calculated according to the frozen-noise algorithm described in Sec. III A. This implies that both the interaural correlation and the power of the interaural difference signal was fixed across all intervals of a run, but each noise sample was a different realization under the above constraints. Thus, across trials, the waveforms arriving at both ears were totally different, but the interaural correlation and the power of the interaural difference signal of the masker was fixed.

(2) Interleaved. Similar to the fixed-ρ condition, each trial consisted of three identical masker intervals, and again one interval contained the signal. However, the number of masker realizations was reduced to two. Thus, two frozen-noise samples were calculated as described in Sec. III A. For each trial, one of these realizations was chosen at random and used as the masker in all three intervals of this trial.

The measurement procedure, the signal durations and levels, and the method of presentation to the subjects were the same as described in Sec. III A. We measured thresholds for two masker bandwidths (10 and 100 Hz) and two masker corre-

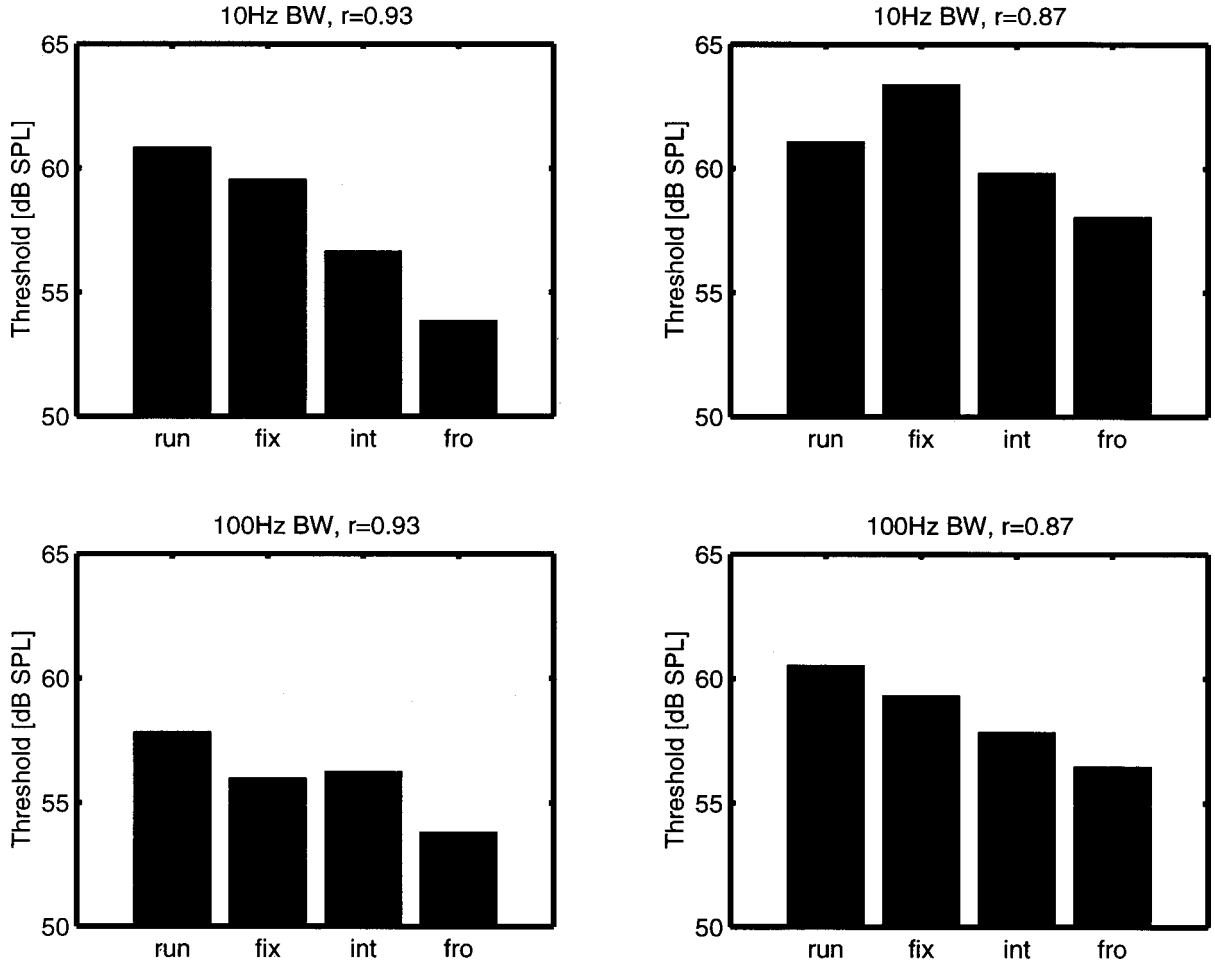

FIG. 9. Mean thresholds across three subjects for the running-noise (run), fixed-ρ (fix), interleaved (int), and frozen-noise (fro) conditions. The upper panels show thresholds for a bandwidth of 10 Hz, the lower panels for 100 Hz. The left and right panels correspond to an interaural correlation of 0.93 and 0.87, respectively.

lations (0.93 and 0.87). Three subjects participated in this experiment.

B. Results

The mean thresholds across subjects are shown in Fig. 9. The upper panels show thresholds for a masker bandwidth of 10 Hz; the lower panels correspond to 100-Hz bandwidth. The left and right panels correspond to an interaural correlation of 0.93 and 0.87, respectively. In each panel, four threshold values are shown. From left to right, these are thresholds for the running-noise condition of experiment I (labeled ''run''), the thresholds for the fixed-ρ condition (''fix''), the interleaved condition (''int''), and the frozen-noise condition of experiment I (''fro'').

As described above, these conditions reflect different levels of stimulus uncertainty. The first level corresponds to absolutely no stimulus uncertainty (frozen-noise conditions) and this condition results in uniformly lower thresholds than all other conditions. If the level of stimulus uncertainty is increased by a small step (the interleaved condition), thresholds increase by 1 to 3 dB for all tested conditions. A third level of stimulus uncertainty was to apply only one restriction to the masker stimuli: the overall interaural correlation and hence the power of the difference signal had to be constant. For three out of four conditions, this also resulted in an increase in thresholds. The differences between frozen-noise and fixed-ρ are about 6 dB at 10-Hz bandwidth and 4 dB at 100-Hz bandwidth. Finally, the highest level of stimulus uncertainty in the present experiments (i.e., running noise) resulted on average in very similar thresholds to those in the fixed-ρ conditions.[3]

C. Discussion

Some striking remarks can be made with respect to the thresholds for stimuli with a fixed interaural correlation (i.e., the fixed-ρ, the interleaved, and the frozen-noise conditions). If the binaural auditory system uses the interaural correlation of each token as a decision variable, the processing of the masker alone would result in an internal estimate of the externally presented interaural correlation. This internal value is fixed and only limited by internal noise. The addition of the signal results in a decrease of the interaural correlation and can thus be detected. Based on such an interaural-correlation processing, all the thresholds for the conditions with a fixed interaural correlation should give the same thresholds. This was not found in our data.

One reason for the differences across these conditions may be peripheral filtering. The externally presented stimu-

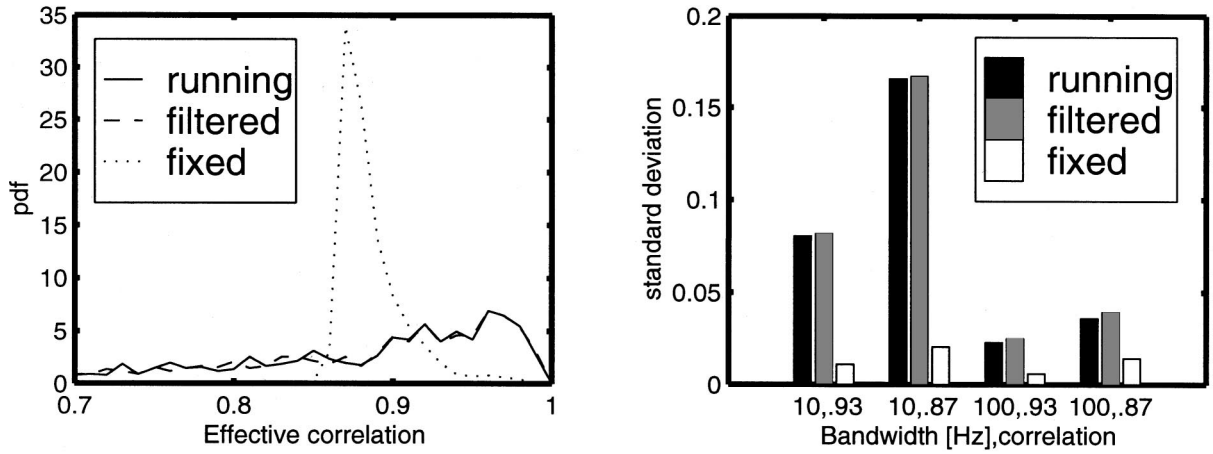

FIG. 10. Left panel: simulated correlation distribution for a 10-Hz-wide, Nρ stimulus with a reference correlation of 0.87. These distributions were obtained for running noise (running), running noise after peripheral filtering (filtered), and for fixed-ρ conditions after peripheral filtering (fixed). Right panel: standard deviations for the same stimuli as in the left panel for different values for the masker bandwidth and correlation.

lus has a fixed interaural correlation. Peripheral filtering in the cochlea results in frequency-dependent phase shifts in the presented waveforms at both ears. These phase shifts result in a change in the interaural correlation. Therefore, for the fixed-ρ condition, the interaural correlation of different tokens after peripheral filtering follows a distribution rather than having a fixed value. To evaluate this hypothesis quantitatively, we computed 1000 partially correlated noises following three different procedures:

(1) Running-noise samples which are generated in the same way as described in Sec. III A.
(2) The same running-noise samples after filtering by a fourth-order gammatone filter with a center frequency of 500 Hz. This filter simulates the effect of peripheral filtering in the inner ear; and
(3) Fixed-ρ samples generated as described in Sec. III A, also after filtering with the gammatone filter.

An example of the correlation distributions that were found with this procedure is given in the left panel of Fig. 10. In this example, the bandwidth of the noise was 10 Hz and the reference correlation was 0.87. The solid line is the distribution for running noise without filtering (i.e., procedure 1). The dashed line corresponds to procedure 2 (i.e., running noise after peripheral filtering). Clearly, these distributions are very similar, indicating that peripheral filtering does not change the statistics of the interaural correlation for running noise. The distribution for fixed-ρ after peripheral filtering is shown by the dotted line. In line with our hypothesis, the distribution has a substantially reduced standard deviation. The values for the standard deviation of the correlation distribution are shown in the right panel of Fig. 10. The black bars correspond to the running-noise procedure without filtering (number 1 in the above description); the gray bars denote running noise after peripheral filtering (2), and the white bars denote fixed-ρ samples after peripheral filtering (3). The x axis indicates the combinations of bandwidth and reference correlation of the noise for each condition. All fixed-ρ conditions have a nonzero amount of correlation uncertainty. This supports the hypothesis that peripheral filtering produces uncertainty in the interaural correlation. The magnitude of this correlation uncertainty is, however, not sufficient to explain the thresholds of the fixed-ρ condition. We computed the values for the detectability index in terms of interaural correlation for the fixed-ρ condition after peripheral filtering as described in Sec. III B. All d' values for the filtered fixed-ρ condition were above 27, indicating that subjects performed worse than expected from correlation uncertainty introduced by peripheral filtering.

An important conclusion that can be drawn from these results is that it is unlikely that the auditory system uses the interaural correlation of the complete token as a decision variable. These results also show that the overall power of the difference signal of the complete token as a decision variable is not a valid descriptor of how the auditory system processes Nρ stimuli.

A possibility to explain the results qualitatively is based on the idea of internal templates (Dau *et al.*, 1996a, b; Breebaart and Kohlrausch, 1999). Assume that listeners develop an internal representation of the interaural differences that occur as a function of time if a masker alone is presented. Such a template can be obtained from the masker-alone intervals in the 3IFC task. One possible realization would be the running average of the difference power based on a time constant that is smaller than the duration of the tokens. If such a template exists, then the task of the listener is to match the template to the internal representation of the actually presented stimuli.

For example, in an NoSπ condition, the masker alone contains no interaural differences. Hence, the template consists of a sequence of zero interaural differences. The addition of the signal results in changes in the interaural differences which can be detected. In this case, there is no uncertainty in the masker-alone representation. The same holds for the frozen-noise conditions. All masker-alone intervals are identical, resulting in the same template. The only task that a listener has to perform is to detect which interval produces an internal representation that differs from the template. This process is in principle limited by internal noise only.

If two different Nρ tokens are used in random order (i.e., the interleaved condition), detection becomes somewhat more complicated. For perfect detection, the listener has to store two templates (one for each token) and must be able to recognize which masker token is used before the templates are compared with the actual stimulus. If the wrong template is matched with the stimulus, all intervals from the trial result in an imperfect match of the template. This increases the probability of choosing the wrong interval, and hence detection performance decreases.

For the fixed-ρ condition, finally, it is only possible to derive an averaged template based on many different noise realizations. This explains the increase in thresholds with respect to the interleaved and frozen conditions, in which the template has a close relation to the actual stimulus. In such a view, fixed-ρ and running-noise conditions are equivalent with respect to the detection strategy. The fixed-ρ condition, however, does allow comparison of internal representations across the three intervals within a trial. Given our experimental data, which show no statistically significant difference between fixed-ρ and running-noise conditions, we can conclude that such an across-interval comparison does not give a significant detection advantage in our conditions.

VI. GENERAL CONCLUSIONS

The results suggest that for binaural signal detection with partially correlated noises, two factors play an important role:

(1) The reference correlation. With decreasing masker correlation, the NρSπ thresholds increase; and
(2) Stimulus uncertainty. Our results show that uncertainty in binaural cues reduces detection performance, especially in narrow-band conditions.

An unresolved issue concerning the data presented in this paper is the method of internal binaural processing. We have shown that the three theories tested (the interaural correlation, the EC theory, or the processing of IIDs and ITDs) cannot account for the results found in this study. The data suggest that the auditory system is able to use internal templates in the process of binaural signal detection. Quantitative tests to support this notion are, however, beyond the scope of this paper and have yet to be performed.

ACKNOWLEDGMENTS

We thank A. Houtsma, S. van de Par, the Associate Editor D. W. Grantham, and the reviewers for their valuable comments on earlier drafts of this paper. The investigations were supported by the Research Council for Earth and Lifesciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO).

APPENDIX

For the generation of an NρSπ stimulus, two independent noise sources $N_0(t)$ and $N_\pi(t)$ with the same rms value are used, which are combined as follows to yield $L(t)$ and $R(t)$ for the left and right ears, respectively:

$$\begin{aligned} L(t) &= \tfrac{1}{2}\sqrt{2}\sqrt{1+\rho}N_0(t) + \tfrac{1}{2}\sqrt{2}\sqrt{1-\rho}N_\pi(t) + S(t), \\ R(t) &= \tfrac{1}{2}\sqrt{2}\sqrt{1+\rho}N_0(t) - \tfrac{1}{2}\sqrt{2}\sqrt{1-\rho}N_\pi(t) - S(t). \end{aligned} \tag{A1}$$

As a result of fluctuations in the energy of a finite-length interval of the Gaussian-noise samples $N_0(t)$ and $N_\pi(t)$, the effective correlation (ρ_{eff}) of the masker sample may deviate from the desired reference correlation (ρ). Because the noise sources $N_0(t)$ and $N_\pi(t)$ are independent, the effective (i.e., physically occurring) interaural correlation (ρ_{eff}) for the NρSπ stimulus can be written as (neglecting the correlation between the two independent noise sources N_0 and N_π; see also footnote 2)

$$\rho_{\text{eff}} = \frac{(0.5+0.5\rho)E_0 - (0.5-0.5\rho)E_\pi - E_s}{(0.5+0.5\rho)E_0 + (0.5-0.5\rho)E_\pi + E_s}, \tag{A2}$$

where E_x is defined as the energy of the stimulus of duration T, according to

$$E_x = \int_{-T/2}^{T/2} N_x^2(t)\,dt. \tag{A3}$$

From Rice (1959), it is known that for a Gaussian-noise sample, E is distributed normally according to:

$$p(E) = \frac{1}{\sigma_E\sqrt{2\pi}} e^{[-(E-m_E)^2]/2\sigma_E^2}, \tag{A4}$$

with

$$m_E = T\int_0^\infty \omega(f)\,df, \tag{A5}$$

and

$$\sigma_E^2 = T\int_0^\infty \omega^2(f)\,df. \tag{A6}$$

Here, $\omega(f)$ refers to the spectral power density of the noise source. The relation between the energies E_0, E_π, and E_s for a certain ρ_{eff} according to Eq. (A2) is given by

$$E_0 = \alpha E_\pi + \beta E_s, \tag{A7}$$

with

$$\alpha = \left(\frac{0.5-0.5\rho}{0.5+0.5\rho}\right)\left(\frac{1+\rho_{\text{eff}}}{1-\rho_{\text{eff}}}\right), \tag{A8}$$

and

$$\beta = \frac{1+\rho_{\text{eff}}}{(1-\rho_{\text{eff}})(0.5+0.5\rho)}. \tag{A9}$$

One way to realize a correlation of ρ_{eff} is to fix E_π at a certain value and compute the necessary value of E_0 according to Eq. (A7). The probability for that realization of ρ_{eff} is then equal to the product of the probabilities $p_E(E_0)$ and $p_E(E_\pi)$. Because there are many possible ways to realize a correlation of ρ_{eff}, we have to sum all possibilities of these realizations

$$p(\rho_{\text{eff}})\Delta\rho_{\text{eff}} = \sum_{E_\pi} p_E(E_0)\Delta E_0 p_E(E_\pi)\Delta E_\pi, \tag{A10}$$

which results in

$$p(\rho_{\text{eff}}) = \int_{E_\pi} p_E(\alpha E_\pi + \beta E_s) p_E(E_\pi) \frac{\partial E_0}{\partial \rho_{\text{eff}}} dE_\pi, \quad \text{(A11)}$$

and hence

$$p(\rho_{\text{eff}}) = \frac{1}{\sigma_E^2 2\pi} \times \int_{E_\pi} \exp\left(\frac{-(\alpha E_\pi + \beta E_s - m_E)^2 - (E_\pi - m_E)^2}{2\sigma_E^2} \right) \times \frac{\partial E_0}{\partial \rho_{\text{eff}}} dE_\pi, \quad \text{(A12)}$$

with

$$\frac{\partial E_0}{\partial \rho_{\text{eff}}} = \frac{0.5 - 0.5\rho}{0.5 + 0.5\rho} \frac{2}{(1-\rho_{\text{eff}})^2} E_\pi + \frac{2}{(1-\rho_{\text{eff}})^2 (0.5 + 0.5\rho)} E_s. \quad \text{(A13)}$$

In summary, if the spectral shape of the noise source and the sample duration are known, Eqs. (A5) and (A6) supply values for m_E and σ_E^2. For a given signal energy E_s and a given reference correlation ρ, Eq. (12) gives the probability density for the occurrence of a certain interaural correlation.

A difficulty arising from the probability density function given by the above equations is that for correlations close to 1, the function becomes skewed. If the distribution for the interaural correlation were Gaussian, it would be easier to calculate parameters like the detectability index d' for two different distributions. Therefore, the Fisher ρ-to-Z transformation is used. This transformation results in a probability density function that behaves approximately normal, and is given by

$$Z = 0.5 \ \ln \frac{1 + \rho_{\text{eff}}}{1 - \rho_{\text{eff}}}. \quad \text{(A14)}$$

Thus, for a given Z, the corresponding interaural correlation becomes

$$\rho_{\text{eff}} = \frac{e^{2Z} - 1}{e^{2Z} + 1}, \quad \text{(A15)}$$

and hence

$$\frac{d\rho_{\text{eff}}}{dZ} = \frac{4e^{2Z}}{(e^{2Z}+1)^2}. \quad \text{(A16)}$$

The probability density function for Z is then given by

$$p(Z) = p(\rho_{\text{eff}}) \frac{d\rho_{\text{eff}}}{dZ}. \quad \text{(A17)}$$

The detectability index for the NρSπ condition is determined by the means and the standard deviations in terms of Z as follows. The mean (μ) and standard deviations (σ) of the distributions of $p(Z)$ are determined for both masker alone and masker plus signal (μ_{N_ρ}, σ_{N_ρ}, $\mu_{N\rho S\pi}$, $\sigma_{N\rho S\pi}$, respectively). The detectability index is then obtained as

$$d' = \frac{\mu_{\text{N}\rho\text{S}\pi} - \mu_{N_\rho}}{\sqrt{\sigma^2_{\text{N}\rho\text{S}\pi} + \sigma^2_{N_\rho}}}. \quad \text{(A18)}$$

[1]We assume that the correlation between the signal S and the noise N is zero. Although this is mathematically not correct for a finite-length interval, a computational analysis revealed that the effect of these correlations is negligible in our analysis.

[2]The correlation between two finite-length samples from independent noise sources is almost never *exactly* zero. In our analysis and generation of stimuli, however, we assume that this correlation is zero. To justify this assumption, we generated 1000 intervals of interaurally partially correlated noise and determined the width of the probability density functions for the interaural correlation after (1) combination of the signals according to Eq. (1), and (2) combination of the signals according to Eq. (1) *after normalizing* the energies of the noise samples to a fixed value (i.e., there was no energy fluctuation for this case). The width of the probability density function for a reference correlation of 0 after normalization of the noise intervals was approximately 10^5 times narrower than without normalization. This indicates that for the way we generated the Nρ stimuli (i.e., with *two* independent noise sources), energy fluctuation is the main cause for correlation fluctuations. Another reason why this assumption is reasonable is the fact that the processing of the cochlea results in phase shifts in the presented stimuli. It is possible to generate a waveform with an exact interaural correlation. But, this is only possible for the waveforms arriving at the eardrums. After the processing in the peripheral hearing system, phase shifts result in changes in the correlation. It is therefore not so valuable to take the correlation between waveforms into account, because this property changes by the processing of the cochlea.

[3]A MANOVA analysis of the data shown in Fig. 9 was performed with the following independent parameters: amount of stimulus uncertainty, stimulus bandwidth, interaural correlation, and subject. The analysis revealed that bandwidth, correlation, and amount of stimulus uncertainty were statistically significant effects at a 95% confidence interval. A multiple comparison procedure (Fisher's least significant difference method) on the means for the different values of stimulus uncertainty revealed that the contrast for running noise vs fixed-ρ was not statistically significant. On the other hand, the contrasts between fixed-ρ, interleaved, and frozen conditions were all statistically significant at a 95% confidence interval.

Bernstein, L. R., and Trahiotis, C. (**1996**). ''The normalized correlation: Accounting for binaural detection across center frequency,'' J. Acoust. Soc. Am. **100**, 3774–3787.

Breebaart, D. J., van de Par, S., and Kohlrausch, A. (**2001**). ''A model for the effective signal processing in the binaural auditory system based on contralateral inhibition. I. Model setup,'' J. Acoust. Soc. Am. (submitted).

Breebaart, J., and Kohlrausch, A. (**1999**). ''A new binaural detection model based on contralateral inhibition,'' in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific, Singapore), pp. 195–206.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**1998**). ''Binaural signal detection with phase-shifted and time-delayed noise maskers,'' J. Acoust. Soc. Am. **103**, 2079–2083.

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., and Summerfield, Q. (**1998**). ''Measurements of the binaural temporal window using a detection task,'' J. Acoust. Soc. Am. **103**, 3540–3553.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996a**). ''A quantitative model of the 'effective' signal processing in the auditory system: I. Model structure,'' J. Acoust. Soc. Am. **99**, 3615–3622.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996b**). ''A quantitative model of the 'effective' signal processing in the auditory system: II. Simulations and measurements,'' J. Acoust. Soc. Am. **99**, 3623–3631.

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.

Durlach, N. I. (**1963**). ''Equalization and cancellation theory of binaural masking-level differences,'' J. Acoust. Soc. Am. **35**, 1206–1218.

Durlach, N. I., Gabriel, K. J., Colburn, H. S., and Trahiotis, C. (**1986**).

''Interaural correlation discrimination: II. Relation to binaural unmasking,'' J. Acoust. Soc. Am. **79**, 1548–1557.
Gabriel, K. J., and Colburn, H. S. (**1981**). ''Interaural correlation discrimination: I. Bandwidth and level dependence,'' J. Acoust. Soc. Am. **69**, 1394–1401.
Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.
Grantham, D. W. (**1984**). ''Discrimination of dynamic interaural intensity differences,'' J. Acoust. Soc. Am. **76**, 71–76.
Grantham, D. W., and Wightman, F. L. (**1978**). ''Detectability of varying interaural temporal differences,'' J. Acoust. Soc. Am. **63**, 511–523.
Grantham, D. W., and Wightman, F. L. (**1979**). ''Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation,'' J. Acoust. Soc. Am. **65**, 1509–1517.
Green, D. (**1992**). ''On the similarity of two theories of comodulation masking release,'' J. Acoust. Soc. Am. **91**, 1769.
Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York).
Hafter, E. R., and Carrier, S. C. (**1969**). ''Masking-level differences obtained with a pulsed tonal masker,'' J. Acoust. Soc. Am. **47**, 1041–1047.
Hirsh, I. (**1948**). ''The influence of interaural phase on interaural summation and inhibition,'' J. Acoust. Soc. Am. **20**, 536–544.
Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (**1952**). ''The masking of tones by white noise as a function of the interaural phases of both components. I. 500 cycles,'' J. Acoust. Soc. Am. **24**, 523–527.
Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (**1962**). ''Masking and interaural phase. II. 167 cycles,'' J. Acoust. Soc. Am. **34**, 1124–1126.
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.
Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Lutfi, R. A. (**1990**). ''How much masking is informational masking?'' J. Acoust. Soc. Am. **88**, 2607–2610.
Rice, S. O. (**1959**). *Selected Papers on Noise and Stochastic Processes* (Dover, New York), Chap. ''Mathematical analysis of random noise.''
Richards, V. M. (**1987**). ''Monaural envelope correlation perception,'' J. Acoust. Soc. Am. **82**, 1621–1630.
Robinson, D., and Jeffress, L. (**1963**). ''Effect of varying the interaural noise correlation on the detectability of tonal signals,'' J. Acoust. Soc. Am. **35**, 1947–1952.
van de Par, S., and Kohlrausch, A. (**1995**). ''Analytical expressions for the envelope correlation of certain narrow-band stimuli,'' J. Acoust. Soc. Am. **98**, 3157–3169.
van de Par, S., and Kohlrausch, A. (**1997**). ''A new approach to comparing binaural masking level differences at low and high frequencies,'' J. Acoust. Soc. Am. **101**, 1671–1680.
van de Par, S., and Kohlrausch, A. (**1998**). ''Analytical expressions for the envelope correlation of narrow-band stimuli used in CMR and BMLD research,'' J. Acoust. Soc. Am. **103**, 3605–3620.
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.
van der Heijden, M., and Trahiotis, C. (**1998**). ''Binaural detection as a function of interaural correlation and bandwidth of masking noise: Implications for estimates of spectral resolution,'' J. Acoust. Soc. Am. **103**, 1609–1614.
Zurek, P. M., and Durlach, N. I. (**1987**). ''Masker-bandwidth dependence in homophasic and antiphasic tone detection,'' J. Acoust. Soc. Am. **81**, 459–464.