Binaural signal detection with phase-shifted and time-delayed noise maskers

Jeroen Breebaart, Steven van de Par, and Armin Kohlrausch
IPO, Center for Research on User-System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands

(Received 25 June 1997; revised 2 October 1997; accepted 5 November 1997)

Detection thresholds were measured for interaurally in-phase sinusoids added to a narrow-band dichotic noise masker which was either interaurally phase shifted (NπSo condition) or time delayed (NτSo condition). The signals were spectrally centered in the noise bands and the delay τ equaled half the signal period. Both conditions were tested at 125 and 500 Hz for noise bandwidths of 10, 25, 50, and 100 Hz. In addition, NoSπ and NoSo thresholds were obtained. In contrast to expectations based on the EC theory, no differences in detection thresholds were observed between thresholds in phase-shifted and time-shifted maskers. The results also cannot be explained on the basis of more recent binaural models. The stimuli presented here might therefore serve as a useful validation tool in the development of new binaural theories and models. © *1998 Acoustical Society of America.* [S0001-4966(98)03904-6]

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

INTRODUCTION

When broadband noise is presented in phase to both ears, and pure tones are presented out of phase to each ear simultaneously (NoSπ condition), the masked threshold is generally lower than for the case when both the noise and the tone are presented in phase (NoSo condition) (Hirsh, 1948; Hafter and Carrier, 1969; Zurek and Durlach, 1987). The difference in detection threshold between the NoSo and the NoSπ condition is referred to as binaural masking level difference (BMLD). The increased sensitivity for out-of-phase signals in a diotic noise is due to the generation of interaural differences by adding the signal to the masker (cf. Zurek, 1991).

If the masker is interaurally phase shifted or time delayed and combined with an in-phase sinusoidal signal (NπSo and NτSo, respectively), the detection threshold is also lower than for the NoSo condition (Webster, 1951; Jeffress *et al.*, 1952). However, the NπSo and NτSo BMLDs are generally smaller than for the NoSπ condition (Hirsh, 1948; Jeffress *et al.*, 1952, 1962; Langford and Jeffress, 1964; Kohlrausch, 1986; Van de Par and Kohlrausch, 1997). The BMLD difference between the NoSπ and the NπSo condition can amount up to about 7 dB for a signal frequency of 167 Hz, and decreases with increasing signal frequency (Jeffress *et al.*, 1962).

One of the theories for interpreting BMLDs is the Equalization and Cancellation (EC) theory, developed by Durlach (1963). The basic idea of this theory is that the auditory system attempts to eliminate the masking components by first transforming the stimuli presented to the two ears in order to equalize the two masking components (E-process). One possible transformation is applying an internal time delay. It is assumed that the E-process is performed imperfectly due to internal errors. Subsequently, the stimulus in one ear is subtracted from the stimulus in the other ear (C-process). For NτSo, where τ equals half the period of the center frequency of the noise and the signal, the cancellation process, after a compensating internal delay, results in an improvement of the signal-to-noise ratio which is equal to the improvement for the NoSπ condition.

For a phase-shifted masker (NπSo), the interaural phase shift has to be compensated by an internal time delay. For a band-limited noise, the consecutive periods of the temporal waveform are not exactly equal. Thus an external phase shift cannot perfectly be compensated by an internal delay. This mismatch of phase shift and time delay will increase with increasing masker bandwidth, due to the decreasing similarity of subsequent periods of the masker. Therefore, the interaural correlation of the Nπ masker after the E-process is smaller than one. This decrease of correlation will be referred to as *decorrelation* of the masker. It is well known that binaural thresholds for an Sπ signal increase with decreasing interaural correlation of a masker (Robinson and Jeffress, 1963). In this respect, and under the assumption that binaural detection for the NπSo condition is degraded by both internal errors and decorrelation (Siegel and Colburn, 1989), NτSo and NoSπ BMLDs should be larger than NπSo BMLDs.

It has been shown previously that for wide-band maskers, the NτSo and NπSo BMLDs are smaller than the NoSπ BMLDs (Jeffress *et al.*, 1962; Langford and Jeffress, 1964), especially at low frequencies. However, a comparison between the NπSo and the NτSo BMLDs for narrow-band maskers as a function of frequency and bandwidth of the noise has never been made. Since the bandwidth and frequency of the masker determine the amount of decorrelation within the framework of the EC theory (and hence the detection threshold), the experiments form a critical test for the EC theory.

To evaluate the hypothesis that the binaural system benefits from the absence of decorrelation in an interaurally time-shifted condition in contrast to an interaurally phase-shifted condition, as suggested by the EC theory, binaural masked thresholds were measured for NτSo and NπSo as a

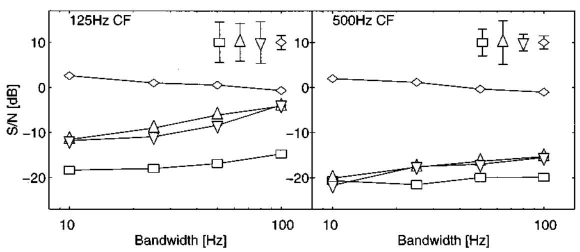

FIG. 1. NoSo (diamonds), NπSo (triangles down), NτSo (triangles up), and NoSπ (blocks) thresholds at 125-Hz (left panel) and 500-Hz (right panel) center frequency as a function of masker bandwidth. Error bars denote the mean standard deviation between subjects for NoSo, NπSo, NτSo, and NoSπ.

function of masker bandwidth. As a reference, thresholds for NoSo and NoSπ were also measured.

I. PROCEDURE AND STIMULI

A 3-interval forced-choice procedure with adaptive signal-level adjustment was used to determine masked thresholds. Three masker intervals of 400-ms duration were separated by pauses of 300 ms. A signal of 300-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). The initial step size for adjusting the level was 8 dB. The stepsize was halved after every second reversal of the level track 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 four 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 masker signals were presented to the subjects over Beyer Dynamic DT990 headphones at a sound pressure level of 65 dB.

The 400-ms masker samples were obtained by randomly selecting a segment from a 2000-ms bandpass-noise buffer. The bandpass-noise buffer was created in the frequency domain by selecting the frequency range from the Fourier transform of a 2000-ms broadband Gaussian noise. After an inverse Fourier transform, the band-limited noise buffer of 2000 ms was obtained.

The 300-ms signals were sinusoids with a frequency equal to the center frequency of the noise masker. In order to avoid spectral splatter, the signal and the maskers were gated with 50-ms raised-cosine ramps. Thresholds are expressed as signal-to-overall-noise-power ratio and are the means of four repetitions per condition and subject. Masked thresholds were measured for NτSo, NπSo, NoSo, and NoSπ conditions where τ equals half the period of the center frequency.

II. RESULTS

Thresholds were obtained from three well-trained subjects with normal hearing. Mean NoSo, NπSo, NτSo, and NoSπ thresholds are shown in Fig. 1 at center frequencies of 125 (left panel) and 500 Hz (right panel). The thresholds for the NoSo condition (diamonds) show a slight decrease with increasing bandwidth for both center frequencies. The slope of 1.3 dB/oct is in line with other data obtained for noise maskers of subcritical bandwidth (de Boer, 1962; Kidd *et al.*, 1989; van de Par and Kohlrausch, 1997). The thresholds for the NπSo (triangles down) and NτSo condition (triangles up) are very similar, while the NoSπ condition (squares) shows lower thresholds. The average difference in detection thresholds between the NπSo and NoSπ conditions is 8 dB at 125 Hz and 4 dB at 500 Hz, where the differences obtained with the 100-Hz masker agree with the results from Kohlrausch (1986) for broadband maskers. The difference between NoSπ and NτSo thresholds for 100-Hz bandwidth at 500 Hz amounts to about 5 dB, a value similar to the 2.9-dB difference measured by Langford and Jeffress (1964) for a broadband masker. Furthermore, there is an increase of the NπSo and NτSo thresholds with increasing masker bandwidth which is stronger at 125 Hz (e.g., 2.1 dB/oct of masker bandwidth) than at 500 Hz (1.2 dB/oct). The mean difference between NπSo and NτSo thresholds amounts to -0.5 dB, implying that NτSo thresholds are slightly higher than NπSo thresholds. According to a one-tailed Student's t test on the pooled differences of all subjects and conditions, this difference was significantly different from zero at a 2% significance level.

III. DISCUSSION

According to the EC theory, an Nπ or Nτ masker is equalized by applying an internal time delay. For the Nτ stimulus, the waveforms at the right and left side after the equalization stage are identical, yielding a perfect interaural correlation. For the Nπ stimulus, the phase shift is also compensated by an internal delay, yielding only a partially correlated masker. Therefore, if only this single time delay is used in the detection process, NτSo thresholds would be expected to be lower than NπSo thresholds. This was not found to be true in our experiments.

At 125-Hz center frequency, the increase of thresholds with increasing masker bandwidth is larger than at 500 Hz. This may be related to the fact that the internal delay necessary to compensate for the external phase shift decreases with increasing center frequency, since the optimal internal delay equals half the period of the center frequency of the noise. Note that the temporal fluctuations of the envelope are independent of center frequency, i.e., the interaural correlation decreases only with increasing delay and bandwidth of the masker. Thus for a certain masker bandwidth, the amount of decorrelation for the Nπ masker at a delay of half the period of the center frequency is larger at 125 Hz than at 500 Hz. Furthermore, an increase in the bandwidth of a noise signal results in a stronger damping of the cross-correlation function. Thus at larger internal delays, the decorrelation grows faster with increasing masker bandwidth than at smaller internal delays. Consequently, a smaller effect of masker bandwidth is expected at the higher frequency.

This qualitative analysis of the effects of masker bandwidth and interaural phase relations of the masker made us wonder how large these effects are quantitatively. We therefore evaluated the NπSo and NτSo BMLDs predicted by the EC model *as a function of the internal delay*. The rationale for this extension lies in the fact that the NτSo condition shows smaller BMLDs than the NoSπ condition. According to the EC theory, both conditions should have equal BMLDs. Furthermore, the optimal internal delay of 4 ms for the NτSo condition at 125-Hz center frequency is rather large in comparison with plausible delays that occur in daily listening conditions and in comparison with delays found in the neuronal system (Palmer *et al.*, 1990; Caird *et al.*, 1991; McAlpine *et al.*, 1996). In this respect, we hypothesized that the internal delay assessed for detection at the 125-Hz condition might be smaller than the optimal internal delay of 4 ms (cf. also the discussion in van der Heijden *et al.*, 1997).

According to the EC theory, the NoSπ BMLD $f(0,\pi)$, as a function of the internal delay Δ is given by (see Durlach, 1972)

$$f(0,\pi)=\frac{k+\cos(\omega\Delta)}{k-\rho(\Delta)}. \tag{1}$$

Here, k is a factor that represents internal errors of the stimulus representation, ω is the center frequency of the stimulus and $\rho(\Delta)$ represents the autocorrelation function of the noise after peripheral filtering. The NoSπ thresholds were used to determine the internal error term, resulting in $k=1.0202$ (based on an NoSπ BMLD of 20 dB).[1] The NτSo BMLD $f(\tau,0)$ as a function of the internal delay is Δ given by

$$f(\tau,0)=\frac{k-\cos(\omega\Delta)}{k-\rho(\tau-\Delta)}, \tag{2}$$

while the NπSo BMLD $f(\pi,0)$ is given by

$$f(\pi,0)=\frac{k-\cos(\omega\Delta)}{k+\rho(\Delta)}. \tag{3}$$

We computed the autocorrelation functions $\rho(\Delta)$ by filtering the power spectrum of the noise maskers by a fourth-order gammatone filter with an equivalent rectangular bandwidth of 38.2 Hz at 125-Hz center frequency and 78.7 Hz at 500-Hz center frequency (Glasberg and Moore, 1990). The bandwidth of 78.7 Hz at 500-Hz center frequency is in fair agreement with the estimated filter bandwidth as provided by Langford and Jeffress (1964). Subsequently, an inverse Fourier transform and a normalization resulted in the autocorrelation function for the stimuli after peripheral filtering. With these assumptions we calculated the BMLD as a function of the internal delay. The predicted NπSo and NτSo BMLDs at 125-Hz and 500-Hz center frequencies for a masker bandwidth of 10 Hz are shown as a function of the internal delay in the left panel of Fig. 2. The solid and the dotted lines represent the BMLDs for the NπSo and the NτSo condition at 125-Hz center frequency, while the nearly identical dashed and the dash-dotted lines represent the BMLDs at 500-Hz center frequency. The right panel of Fig. 2 shows the BMLDs at 100-Hz bandwidth in the same format. If we assume, as in the EC theory, an optimal delay is used for detection, some striking results in comparison with the experimental data can be summarized as follows:

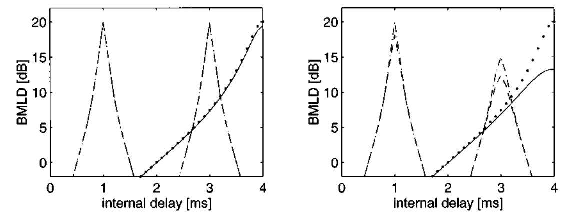

FIG. 2. BMLDs as a function of the internal delay according to the EC theory. NπSo condition at 125-Hz center frequency (solid line); NτSo at 125-Hz center frequency (dotted line); NπSo at 500-Hz center frequency (dashed line); NτSo at 500-Hz center frequency (dash-dotted line). The left panel represents predictions for 10-Hz bandwidth, the right panel those for 100-Hz bandwidth.

(1) The EC theory does not predict any differences in BMLD between the three conditions at 10-Hz bandwidth, assuming that optimal internal delays (1 or 3 ms at 500 Hz, 4 ms at 125 Hz) are used.
(2) At 100-Hz bandwidth, the EC theory predicts a difference of 2 dB between the NπSo and the NτSo condition at 500-Hz center frequency, and a difference of 7 dB at 125-Hz center frequency. These differences are not found in our experimental data.
(3) The predicted NτSo BMLDs remain constant with increasing masker bandwidth, while the experimental NτSo BMLDs decrease from 15 dB at 10-Hz bandwidth to 4 dB at 100-Hz bandwidth for the condition at 125-Hz center frequency.

These observations suggest that, within the framework of the EC theory, binaural detection based on a single, optimal delay cannot account for the experimental data presented here. One could argue that a nonoptimal internal delay is assessed for detection. Such an argument is supported by the findings of Jeffress *et al.* (1962), that the NτSo BMLD for a 167-Hz signal reaches its maximum value already for $\tau=0.5$ ms and does not increase for larger noise delays. From Fig. 2, we see that in order to achieve a BMLD of at least 4 dB for the NπSo condition at 125-Hz center frequency and 100-Hz bandwidth, the internal delay must amount to at least 2.5 ms. Furthermore, assuming that delays within a range of 2.5 ms are available for detection, an NτSo BMLD of 20 dB is expected at 500-Hz center frequency and 100-Hz bandwidth, while the experimental data show a BMLD of only 15 dB.

Also modifications to the EC theory, such as proposed by Green (1966), who suggested that besides subtraction also addition is allowed in the cancellation step, cannot account for our data. Clearly, the addition step would totally cancel the masking noise in an NπSo condition, while the subtraction would totally cancel the masking noise in an NoSπ condition and, after an appropriate delay, in an NτSo condition. Consequently, this would result in equal thresholds for all three conditions independent of the masker bandwidth, which does not correspond to our data.[2]

After having concluded that the EC theory cannot account for the results presented here, we tried to explain the results within the framework of cross-correlation models (see Colburn, 1977; Zwicker and Henning, 1985; Lindemann, 1986; Shackleton *et al.*, 1992; Stern and Shear, 1996; Bernstein and Trahiotis, 1996). A quantitative application of these models to the detection conditions of the present study was beyond our intentions. We think, however, that the existing cross-correlation models cannot account for the dependency of the detection thresholds on the masker bandwidth, due to the fact that the simulated neural activity (or crosscorrelation) in these models is not normalized with the stimulus energy. This leads to the problem that the changes in the neural activity that reflect changes in the interaural correlation are much smaller than the variation in the masker activity. To give an example: for a signal-to-noise ratio of -20 dB, the change in the cross correlation due to the addition of an $S\pi$ signal to a narrow-band diotic masker amounts to 0.02 times the masker power. On the other hand, the standard deviation of the masker energy for a 10-Hz-wide noise of a duration of 400 ms amounts to 0.5 times the masker energy (Bendat and Piersol, 1976). Hence without normalization, the change in the interaural cross-correlation function due to the addition of a signal to a narrow-band masker is much smaller than the uncertainty in the correlation function due to stimulus uncertainty. Thus if a change in the (unnormalized) cross-correlation function is used as a cue for detection, only small BMLDs are expected for narrow-band maskers.[3] Furthermore, since the standard deviation of the masker energy decreases with increasing masker bandwidth, the $NoS\pi$ thresholds should decrease with increasing masker bandwidth. This was not found to be true in our experiments; according to Fig. 1, the $NoS\pi$ thresholds remain approximately constant across bandwidth.

In summary, we think that neither the EC theory nor current crosscorrelation models can predict the experimental data presented here. Therefore, the understanding of the binaural processes involved in the experimental conditions of this study still forms a challenge for psychoacousticians.

ACKNOWLEDGMENTS

We want to thank all our subjects for participating in our experiments which meant spending some of their valuable time in the listening booth. We thank G. B. Henning, Andrew Oxenham, and an anonymous reviewer for their valuable comments on earlier drafts of this paper.

[1]Since we are interested in comparing, at each of the two frequencies 125 and 500 Hz, the relation between the three binaural conditions and the influence of masker bandwidth we used, for simplicity, the same value for k at 125 and 500 Hz.

[2]An attempt to explain our data using the vector theory of binaural interaction (Jeffress, 1972) is also bound to fail. As already noted by Jeffress *et al.* (1962), vector diagrams ''do not explain the change of MLD's with frequency nor the rather large difference possible between the MLD's for $NoS\pi$ and $N\pi So$.'' (Jeffress *et al.*, 1962, p. 1125). We can add that, without further modifications, the vector theory would also be unable to predict a change in binaural thresholds with masker bandwidth.

[3]The model by Stern and Shear (1996) incorporates an automatic gain control stage which adjusts and thus eliminates the influence of overall stimulus level. However, as long as this stage does not follow intensity changes on a short-time basis, it does not solve the bandwidth problem described above.

Bendat, J. S., and Piersol, A. G. (**1976**). *Random Data: Analysis and Measurement Procedures* (Wiley-Interscience, New York).

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

de Boer, E. (**1962**). ''Note on the critical bandwidth,'' J. Acoust. Soc. Am. **34**, 985–986.

Caird, D. M., Palmer, A. R., and Rees, A. (**1991**). ''Binaural masking level difference effects in single units of the guinea pig inferior colliculus,'' Hearing Res. **43**, 1–24.

Colburn, H. S. (**1977**). ''Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise,'' J. Acoust. Soc. Am. **61**, 525–533.

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

Durlach, N. I. (**1972**). ''Binaural signal detection: Equalization and cancellation theory,'' in *Foundations of Modern Auditory Theory*, Vol. II, edited by J. V. Tobias (Academic, New York).

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

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

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. J. (**1948**). ''The influence of interaural phase on interaural summation and inhibition,'' J. Acoust. Soc. Am. **20**, 536–544.

Jeffress, L. A. (**1972**). ''Binaural signal detection: Vector theory,'' in *Foundations of Modern Auditory Theory*, Vol. II, edited by J. V. Tobias (Academic, New York).

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.

Kidd, G., Mason, C. R., Brantley, M. A., and Owen, G. A. (**1989**). ''Roving-level tone-in-noise detection,'' J. Acoust. Soc. Am. **86**, 1310–1317.

Kohlrausch, A. (**1986**). ''The influence of signal duration, signal frequency and masker duration on binaural masking level differences,'' Hearing Res. **23**, 267–273.

Langford, T. L., and Jeffress, L. A. (**1964**). ''Effect of noise crosscorrelation on binaural signal detection,'' J. Acoust. Soc. Am. **36**, 1455–1458.

Levitt, R. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Lindemann, W. (**1986**). ''Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals,'' J. Acoust. Soc. Am. **80**, 1608–1622.

McAlpine, D., Jiang, D., and Palmer, A. R. (**1996**). ''Binaural masking level differences in the inferior colliculus of the guinea pig,'' J. Acoust. Soc. Am. **100**, 490–503.

Palmer, A. R., Rees, A., and Caird, D. (**1990**). ''Interaural delay sensitivity to tones and broad band signals in the guinea-pig inferior colliculus,'' Hearing Res. **50**, 71–86.

Robinson, D. E., and Jeffress, L. A. (**1963**). ''Effect of varying the interaural noise correlation on the detectability of tonal signals,'' J. Acoust. Soc. Am. **35**, 1947–1952.

Shackleton, T. M., Meddis, R., and Hewitt, M. J. (**1992**). ''Across frequency integration in a model of lateralization,'' J. Acoust. Soc. Am. **91**, 2276–2279.

Siegel, R. A., and Colburn, H. S. (**1989**). ''Binaural processing of noisy stimuli: Internal/external noise ratios for diotic and dichotic stimuli,'' J. Acoust. Soc. Am. **86**, 2122–2128.

Stern, R. M., and Shear, G. D. (**1996**). ''Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay,'' J. Acoust. Soc. Am. **100**, 2278–2288.

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 der Heijden, M., Trahiotis, C., Kohlrausch, A., and van de Par, S. (**1997**). ''Binaural detection with spectrally nonoverlapping signals and maskers: evidence for masking by aural distortion products,'' J. Acoust. Soc. Am. **102**, 2966–2972.

Webster, F. A. (**1951**). ''The influence of interaural phase on masked thresholds. I. The role of interaural time derivation,'' J. Acoust. Soc. Am. **23**, 452–462.

Zurek, P. M. (**1991**). ''Probability distribution functions of interaural phase and level differences in binaural detection stimuli,'' J. Acoust. Soc. Am. **90**, 1927–1932.

Zurek, P. M., and Durlach, N. I. (**1987**). ''Masker-bandwidth dependence in homophasic and antiphasic tone detection,'' J. Acoust. Soc. Am. **81**, 459–463.

Zwicker, E., and Henning, G. B. (**1985**). ''The four factors leading to binaural masking-level differences,'' Hearing Res. **19**, 29–47.