A new approach to comparing binaural masking level differences at low and high frequencies

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

A new approach to comparing binaural masking level differences at low and high frequencies[a]

Steven van de Par and Armin Kohlrausch
Institute for Perception Research (IPO), P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands

A new experimental technique for studying binaural processing at high frequencies is introduced. Binaural masking level differences (BMLDs) for the conditions N_0S_π and $N_\pi S_0$ were measured for a tonal signal in narrow-band noise at 125, 250, and 4000 Hz. In addition, ''transposed'' stimuli were generated, which were centered at 4000 Hz, but were designed to preserve within the envelope the temporal ''fine-structure'' information available at the two lower frequencies. The BMLDs measured with the 125-Hz transposed stimuli were essentially the same as BMLDs from the regular 125-Hz condition. The transposed 250-Hz stimuli generally produced smaller BMLDs than the stimuli centered at 250 Hz, but the pattern of results as a function of masker bandwidth was the same. The patterns of results from the transposed stimuli are different from those of the 4000-Hz condition and, consistent with the low-frequency masker data, generally show higher BMLDs. The results indicate that the mechanisms underlying binaural processing at low and high frequencies are similar, and that frequency-dependent differences in BMLDs probably reflect the inability of the auditory system to encode the temporal fine structure of high-frequency stimuli. © *1997 Acoustical Society of America.* [S0001-4966(97)04102-7]

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

INTRODUCTION

It is well known that for broadband noise maskers, binaural masking level differences (BMLDs) resulting from the comparison of an N_0S_π condition (the noise masker, N, and the signal, S, have an interaural phase of 0 and π, respectively) with an N_0S_0 condition are much larger at low frequencies than at high frequencies (Durlach, 1964; Metz *et al.*, 1968). At low frequencies, BMLDs are approximately 15 dB, while at frequencies above about 2 kHz they are only 2–3 dB.

When narrow-band noise maskers instead of broadband maskers are used, the BMLDs are generally much larger, both at low and at high frequencies (Metz *et al.*, 1968; Zurek and Durlach, 1987). In this case, BMLDs at low frequencies can be as high as 25 dB, while at high frequencies they can amount to 15 dB.

There are two mechanisms that could account for the differences between low- and high-frequency BMLDs (e.g., Zurek and Durlach, 1987):

(1) With increasing frequency, the auditory filter bandwidth increases. It is a general rule that the maximum rate of fluctuations within a noise band is proportional to its bandwidth. For a broadband masker, therefore, the rate of changes in interaural time and intensity differences at the outputs of the auditory filter increases with increasing signal frequency. This increased rate is detrimental for binaural unmasking, if one assumes that the auditory system is not able to follow these rapid changes (Perrott and Musicant, 1977; Grantham and Wightman, 1978; Grantham, 1984; Bernstein and Trahiotis, 1992).

(2) With increasing frequency, the responses of the inner hair cells show a decrease in phase locking (Palmer and Russell, 1986). Therefore, at frequencies above about 1.5 kHz, the fine-structure information of the input waveform is gradually lost. As a result, interaural time differences which are present in the fine structure of the waveform are no longer present in the activity of the auditory nerve. Therefore, at high frequencies the binaural system has access only to the interaural intensity differences in the envelope of the stimulus.[1]

Although no interaural time differences in the *fine structure* of high-frequency stimuli can be exploited by the auditory system, several studies have shown that the auditory system is able to process interaural time differences that are available in the *envelope* of high-frequency stimuli. Henning (1974) tested the detectability of an interaural delay in a 300-Hz amplitude modulated high-frequency sinusoid and found that performance was as good as with a 300-Hz pure tone. McFadden and Pasanen (1976) measured the minimal interaural delay needed for lateralization of noise bands of several bandwidths and for two-tone complexes as a function of frequency separation and depth of modulation. They found that: ''In many conditions of listening, sensitivity to interaural time differences at high frequencies compares favorably with the sensitivity at low frequencies.''

The ability of the auditory system to process interaural time delays in the envelope of high-frequency stimuli such as presented in these studies suggests that this ability is also exploited in high-frequency binaural masking experiments. However, it is not possible to relate the results from these experiments directly to the difference in the results for low- and high-frequency BMLDs. Apart from the different experimental approach (lateralization versus binaural detection),

---

the specific stimulus properties do not allow a direct comparison of low- and high-frequency data because the internal representation after transformation in the inner hair cells (half-wave rectification and low-pass filtering) is different. Therefore, the question remains whether the difference in the size of BMLDs with frequency is a result of the loss of information contained in the fine structure of a stimulus in the auditory periphery prior to the binaural processing or whether it is the result of different binaural processing capabilities at low and high frequencies.

A first attempt to study binaural unmasking at high frequencies with specific emphasis on envelope structure was performed by Bernstein and Trahiotis (1992). They added a sinusoid to the envelopes of high-frequency narrow bands of noise. While the noise was in phase, the sinusoid was either homophasic or antiphasic. Subjects had to distinguish between intervals containing the homophasic and the antiphasic sinusoid. With this approach the rate of fluctuation of interaural intensity differences (IIDs) and the inherent rate of fluctuations of the envelope could be adjusted independently by changing the frequency of the sinusoid and the bandwidth of the noise, respectively. The results indicated that there is a rate limitation for the processing of dynamically changing IIDs such as proposed by Grantham (1984).

The experiments of this paper are intended to link the ideas on envelope processing at high frequencies with signal properties in a typical binaural masking experiment in a more direct way. We report the results of experiments with a special type of high-frequency stimulus which contains ''fine structure'' also after the first stages of peripheral transduction (basilar membrane filtering and hair-cell transduction). This property is achieved by encoding in the envelope of a 4-kHz carrier the information that is available after the transformation of a low-frequency stimulus through a simple hair-cell model. The temporal information, now presented in the high-frequency channel, is in principle identical to the low-frequency information. Using this technique, the role of fine structure for binaural processing at high and low frequencies can be compared directly.

The results for transposed stimuli are compared with BMLDs obtained with conventional high- and low-frequency stimuli, where the high-frequency stimuli are in the same spectral range as the transposed stimuli. A comparison of the BMLDs for transposed stimuli with those for high-frequency stimuli will indicate whether the additional envelope information affects binaural interaction at high frequencies. If the properties of binaural processing are the same at low and high frequencies, we would expect that the transposed and low-frequency stimuli give very similar BMLDs. In the next section we will explain the calculation of the transposed stimuli and discuss the properties of these stimuli.

I. TRANSPOSED STIMULI

In the following example, the procedure is described for the generation of a transposed stimulus in an $N_0 S_\pi$ condition. The first step is to generate a conventional low-frequency stimulus. Portions of low-frequency stimuli are shown in panel A of Fig. 1. The interval ranging from 0.0–0.1 s shows the time function of a diotic reference stimulus
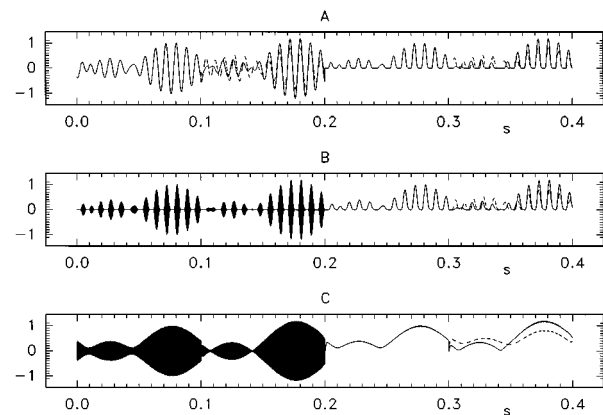

FIG. 1. An example of three different $N_0 S_\pi$ stimuli before and after peripheral processing. Panel A shows a 125-Hz stimulus, panel B shows a transposed stimulus, and panel C shows a 4-kHz stimulus. The intervals 0.0–0.1 s show the N_0 masker alone, the intervals 0.1–0.2 s show the N_0 masker plus the S_π signal at a signal-to-noise ratio of -10 dB, the intervals 0.2–0.3 s show the masker after peripheral processing, and the intervals 0.3–0.4 s show the combined masker and signal after peripheral processing.

(N_0) which is a noise masker with 25-Hz bandwidth centered at 125 Hz. In the interval ranging from 0.1–0.2 s, a dichotic test stimulus $(N_0 S_\pi)$ is shown, with an S_π signal added to the N_0 masker with a signal-to-noise ratio of -10 dB. The two curves in this interval represent the signals at the right and left ear. Comparing the two curves we find interaural time delays in the form of different timings of the zero crossings and we find interaural intensity differences in the form of differences between the envelopes, e.g., at t equals 0.17 s.

The interval ranging from 0.2–0.4 s shows the signals from the first half of panel A, after being processed by a stage that simulates properties of the auditory periphery. These are modeled by half-wave rectifying the input signal and, subsequently, low-pass filtering at 500 Hz. We assume that the signals in the interval 0.2–0.4 s are a reasonable description for the low-frequency stimuli at the level of the inner hair cell.

Multiplying the processed waveforms by a high-frequency carrier (4 kHz in the present experiments), we obtain a ''transposed'' stimulus as shown in panel B. The reference stimulus (noise alone) is plotted in the 0.0–0.1-s interval, the dichotic test stimulus in the interval from 0.1–0.2 s. In the interval 0.2–0.4 s, this transposed stimulus is shown after being processed by the first stages of the auditory periphery. We can now see that with our description of the auditory periphery, essentially the same temporal information is available for the transposed stimulus as for the initial low-frequency stimulus in panel A. The two conditions differ, however, by the center frequency of the auditory channel, through which this information is provided to the binaural processor.

The signals for a standard $N_0 S_\pi$ condition at 4 kHz are shown in panel C. Here, the masker is a 25-Hz-wide noise centered at 4 kHz and the signal is a 4-kHz sinusoid. We see that no information about the stimulus fine structure is present after the peripheral transduction. However, there are interaural differences present in the envelopes of the waveform. By comparing panels B and C, one sees that, with a
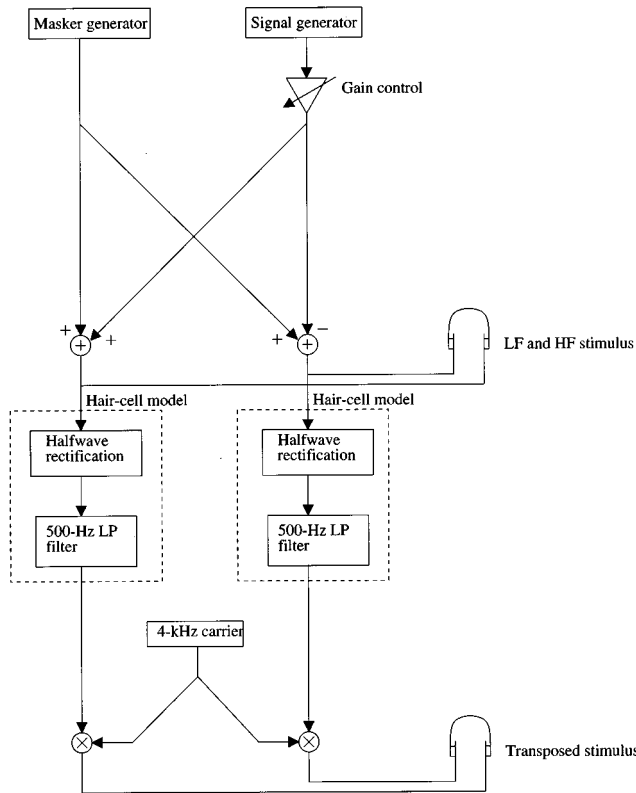
FIG. 2. A schematic overview of the transposed N_0S_π stimulus generation. The scheme starts at the top with the masker and signal generation, followed by a gain control of the signal and the summation and subtraction of the masker and signal. At this point the stimulus for a conventional N_0S_π condition is obtained. This stimulus is then transformed by a simple hair-cell model and multiplied by a 4 kHz carrier resulting in a transposed stimulus.

conventional high-frequency stimulus less information about temporal details is available for any central processing stage following the peripheral transduction.

In Fig. 2 the generation of a transposed N_0S_π stimulus is shown schematically. Since these stimuli will be used in a forced-choice procedure, the transposed stimuli will be either noise alone or noise plus signal. In the upper part of this figure a conventional low-frequency stimulus is generated. The signal-to-noise ratio of the stimulus is adjusted with the gain control. The low-frequency stimulus is then used as an input to generate a transposed stimulus, as is shown in the lower part of Fig. 2. When we discuss the signal-to-noise ratio of a transposed stimulus we will be referring to the signal-to-noise ratio of the underlying low-frequency stimulus.[2]

The hair-cell model that is used in the generation of the transposed stimuli consists of a half-wave rectifier and a second order low-pass filter at 500 Hz. For our purposes these are the important signal-processing characteristics of the inner hair cells. The adaptive and compressive properties of the inner hair cells are not included since we can expect these properties to affect the low-frequency and transposed stimulus similarly once they are transformed by the inner hair cells in the cochlea. The 500-Hz low-pass filter serves as a means to limit the bandwidth of the half-wave rectified signal such that after the multiplication with the 4-kHz carrier, only high-frequency (>1.5 kHz) auditory filters are excited.
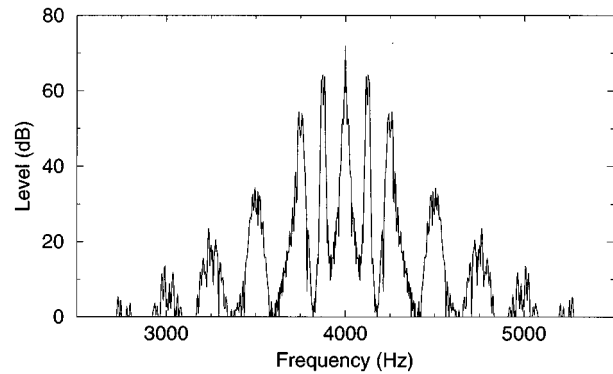

FIG. 3. The spectrum of a 25-Hz-wide 125-Hz transposed stimulus. Most of the stimulus energy is around 4 kHz. The spectrum of the original low-frequency stimulus now occurs at 4125 Hz and the mirror image of that spectrum is at 3875 Hz. Additional peaks in the spectrum occur with spacings of 250 Hz.

The spectrum of a transposed noise band is shown in Fig. 3. The average spectral level is highest around the carrier frequency (4 kHz) and decreases at both sides of the maximum. The spectrum of the original noise band is represented in the two side bands that have a spectral distance from the carrier frequency equal to the center frequency of the noise band. Additional peaks are found at regular intervals of multiples of twice the center frequency of the noise band. More details about the spectrum of the transposed stimulus are given in the Appendix. In our experiments the stimulus energy that was present below frequencies of 1.5 kHz was at least 70 dB lower than the total amount of energy in the stimulus spectrum for all conditions that were measured. Since we presented stimuli at a sound pressure level of about 70 dB, the low-frequency energy was below absolute threshold and could not lead to any binaural cues. This implies that subjects could only use binaural cues at high frequencies where, generally, binaural detection is observed to be worse than at low frequencies.

The introduction of a signal in a transposed stimulus not only leads to changes in the stimulus spectrum near the center frequency of the transposed stimulus, but also to changes in the more remote residual spectral parts. In this respect, the transposed stimulus differs from a conventional high-frequency stimulus. However, an analysis of the transposed stimuli shows that the off-frequency spectral parts are not likely to lead to a better binaural detection than the central part of the spectrum and that therefore the extra spectral components in the transposed stimulus do not affect binaural detection (cf. Appendix). To test this, the spectrum of 125-Hz transposed stimuli was bandpass filtered such that only the central three or five peaks remained.[3] Thresholds for an N_0S_π condition with and without bandpass filtering were measured for subject SP at narrow and broadband conditions. Differences between the conditions were no larger than 1.7 dB, suggesting that the additional spectral components have very little effect on the detection thresholds.

II. EXPERIMENT I

A. Procedure

A three-interval forced-choice procedure with adaptive signal-level adjustment was used to determine masked thresholds. The three masker intervals of 400-ms duration were separated by pauses of 200 ms. A signal of 300-ms duration was added in the temporal center to one of these intervals. The subject's task was to indicate which of the three intervals contained the signal. Feedback was provided to the subject after each trial.

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. After each second reversal of the level track, the step size was halved until a step size of 1 dB was reached. The run was then continued for another eight reversals. From the level of these last eight reversals the median was calculated and used as a threshold value. At least four threshold values were obtained and averaged for each parameter value and subject.

Seven subjects participated in this experiment. They were all laboratory colleagues, except for one subject, and had experience in (monaural) masking experiments.

B. Stimuli

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 signals were presented to the subjects over Telephonics TDH-49P headphones at a sound pressure level of 70 dB.

The 400-ms masker samples for the low- and high-frequency stimuli were obtained by randomly selecting a segment from a 2000-ms bandpass-noise buffer. The band-limited noise buffer was created in the frequency domain by generating a flat spectrum within the passband and randomizing the phases. After an inverse Fourier transform, the 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.[4] Thresholds are expressed as signal-to-overall-noise level.

In this first experiment, BMLDs were obtained by measuring and comparing N_0S_0 and N_0S_π thresholds. Thresholds were measured for maskers with a center frequency of 125 Hz, 250 Hz, and 4 kHz, and bandwidths of 5, 10, 25, 50, 100, and 250 Hz. In addition, a bandwidth of 500 Hz was used for center frequencies of 250 Hz and 4 kHz. Transposed stimuli were obtained from the low-frequency conditions with 125- and 250-Hz center frequencies.

C. Results

Results for the N_0S_0 conditions were very similar across subjects, and therefore only the mean N_0S_0 thresholds are shown in Fig. 4. The symbols in the left panel indicate average results of five subjects for 125 Hz (○), 125-Hz transposed (□), and 4 kHz (◇). Up to a bandwidth of 100 Hz, the three curves are rather similar. Between 100- and 250-Hz
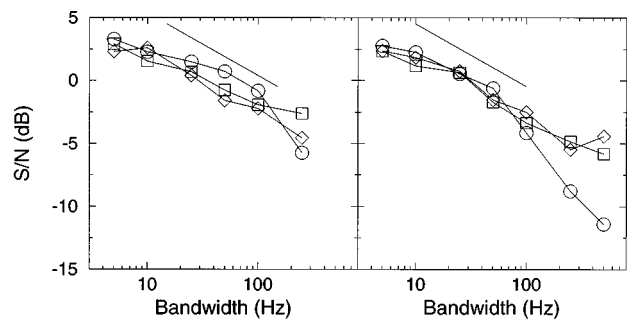

FIG. 4. The N_0S_0 thresholds as a function of masker bandwidth. The left panel shows thresholds for 125 Hz (○), 125-Hz transposed (□), and 4 kHz (◇). The right panel shows thresholds for 250 Hz (○), 250-Hz transposed (□), and 4 kHz (◇). Thresholds are the average of five subjects and are indicated as signal-to-overall-noise level ratio.

bandwidth, the thresholds in the 125-Hz condition seem to decrease more rapidly than in the other two other conditions. This result can be attributed to the fact that the critical bandwidth is narrowest for the 125-Hz condition and that a part of the energy of the 250-Hz-wide masker is filtered out.

The right panel shows average data for the center frequencies 250 Hz (○), 250-Hz transposed (□), and 4 kHz (◇). The data are again averages for five subjects. Three of these subjects also contributed to the data in the left panel. For these conditions the curves are parallel up to a bandwidth of 100 Hz. At wider bandwidths, the 250-Hz curve lies significantly below the two other curves. The 250-Hz transposed and the 4-kHz curves are nearly identical, just as the two corresponding curves in the left panel. These data suggest that monaural processing is similar at low and high frequencies, as long as the auditory filter bandwidth does not affect the stimuli.

With increasing masker bandwidth, thresholds decrease for all conditions. Except for the low-frequency data at large masker bandwidths this is not due to the auditory filter bandwidth but to the variability in the overall stimulus level caused by the fluctuations in the masker envelope (Bos and de Boer, 1966). With a masker of finite length this variability is largest at the narrowest bandwidths. On the basis of the noise statistics it can be shown that the variability in stimulus energy decreases with 1.5 dB/oct (Green and Swets, 1974). If signal detection depends on an energy cue, thresholds can be expected to decrease with 1.5 dB per doubling of the masker bandwidth. The auxiliary lines in the left and right panels of Fig. 4 decrease with this slope and correspond well with the slope in the measured data. In a similar way, a decrease with increasing bandwidth of the signal-to-overall-noise ratio has been reported previously for subcritical bandwidths by several authors (e.g., de Boer, 1962; Weber, 1978; Kidd et al., 1989).

The BMLDs for the N_0S_π condition for center frequencies of 125 Hz, 125-Hz transposed and 4 kHz are shown in Fig. 5. The panels show the individual results of the five observers and at the bottom right, the average of all subjects. The symbol with the error bars indicates the average of the standard deviations over all bandwidths. It is calculated separately for each individual observer and for each stimulus type. In the panel with the average of all observers the stan-
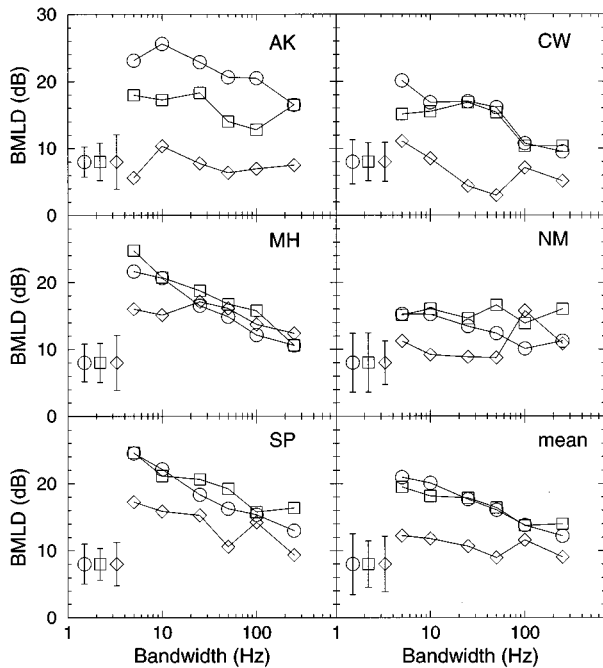
FIG. 5. The $N_0 S_\pi$ BMLDs as a function of masker bandwidth for 125 Hz (○), 125-Hz transposed (□), and 4 kHz (◇). Five panels show data for individual subjects; the panel at the bottom right shows the average of the five subjects. In the bottom right panel, the three symbols with error bars indicate the averaged standard deviation of the mean BMLD for the individual subjects for the three conditions. In all other panels the three symbols with error bars indicate the averaged standard deviation of repeated measurements.

FIG. 6. The $N_0 S_\pi$ BMLDs as a function of masker bandwidth for 250 Hz (○), 250-Hz transposed (□), and 4 kHz (◇). Five panels show data for individual subjects; the panel at the bottom right shows the average of the five subjects.

dard deviation between observers is displayed by the error bars.

In general, we can see that the BMLDs for the 125-Hz and the 125-Hz transposed conditions are larger than BMLDs for the 4-kHz condition. The results also indicate that there is a substantial intersubject variance. However, the BMLDs for the transposed condition show a remarkable similarity with the 125-Hz BMLDs for most subjects. Only for subject AK is there a clear overall difference and for subject NM there is a difference at larger bandwidths. For this subject transposed BMLDs are even larger than the 125-Hz BMLDs. The averaged BMLDs of the five subjects (bottom right panel) again show that transposed BMLDs are very similar to the 125-Hz BMLDs. Furthermore, they are always larger than the 4-kHz BMLDs.

The results for the $N_0 S_\pi$ condition for center frequencies of 250 Hz, 250-Hz transposed, and 4 kHz are shown in Fig. 6. Again, the data from individual subjects are plotted, and the bottom right panel shows the average of all subjects.

Also for the 250-Hz and the 250-Hz transposed BMLDs there is a large variance among subjects. In general, we can see that the BMLDs for the 250-Hz condition are larger than those for the 4-kHz condition. The results are consistent with similar data of Zurek and Durlach (1987).

The BMLDs for the transposed condition show less similarity with the 250-Hz BMLDs than was the case in the comparable situation at 125 Hz. Comparing the 250-Hz transposed BMLDs with the 4-kHz BMLDs at small bandwidths up to 25 Hz, we see that the transposed BMLDs are
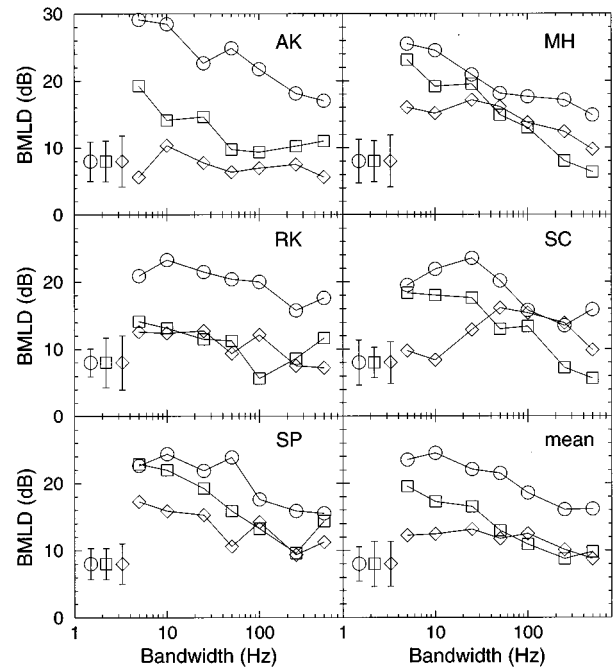
larger for most subjects. For larger bandwidths the transposed BMLDs are generally not higher than the 4-kHz BMLDs except for subject AK. The averaged BMLDs of the five subjects in Fig. 6 (bottom right panel) show that transposed BMLDs are larger than 4-kHz BMLDs at bandwidths below 50 Hz.

III. EXPERIMENT II

A. Stimuli and method

This second experiment is very similar to the first, except that the binaural condition is $N_\pi S_0$ instead of $N_0 S_\pi$. Measurements were performed for 125 Hz, 125-Hz transposed, and 4 kHz. Four of the five previous subjects participated in this experiment.

The rationale for this experiment is that at low frequencies, BMLDs for $N_\pi S_0$ are smaller than for $N_0 S_\pi$ (Metz et al., 1968; Kohlrausch, 1986). If the processing of the transposed stimuli is similar to that of the underlying low-frequency stimuli, a similar difference in size of the BMLDs should exist between transposed $N_0 S_\pi$ and $N_\pi S_0$ conditions.

In Fig. 7, a plot, comparable to that in Fig. 1, is shown for the $N_\pi S_0$ condition. For the low-frequency and the transposed conditions (panels A and B), the fine structure of the waveform is essentially out of phase as a result of the antiphasic masker. This leads to a fundamentally different situation for the two reference stimuli (masker alone), N_0 and N_π. For the N_0 stimulus, the waveforms after peripheral transduction are identical, yielding an interaural correlation of 1. For the N_π stimulus, the correlation is smaller even after an internal delay to the stimulus in one of the ears. Thus, in the internal representation, the N_π stimulus never
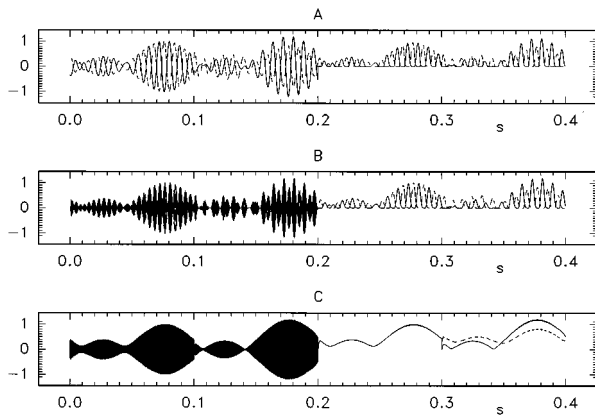
FIG. 7. An example of three different $N_\pi S_0$ stimuli before and after peripheral processing. Panel A shows a 125-Hz stimulus, panel B shows a transposed stimulus, and panel C shows a 4-kHz stimulus. The intervals 0.0–0.1 s show the N_π masker alone, the intervals 0.1–0.2 s show the N_π masker plus the S_0 signal at a signal-to-noise ratio of -10 dB, the intervals 0.2–0.3 s show the masker after peripheral processing, and the intervals 0.3–0.4 s show the combined masker and signal after peripheral processing.

reaches an interaural correlation of 1. This may be one of the reasons for the smaller BMLDs for the $N_\pi S_0$ condition.

B. Results

The BMLDs for the $N_\pi S_0$ condition for center frequencies of 125 Hz, 125-Hz transposed, and 4 kHz are shown in Fig. 8. The first four panels show the data for the individual subjects. The bottom left panel is the average of all subjects
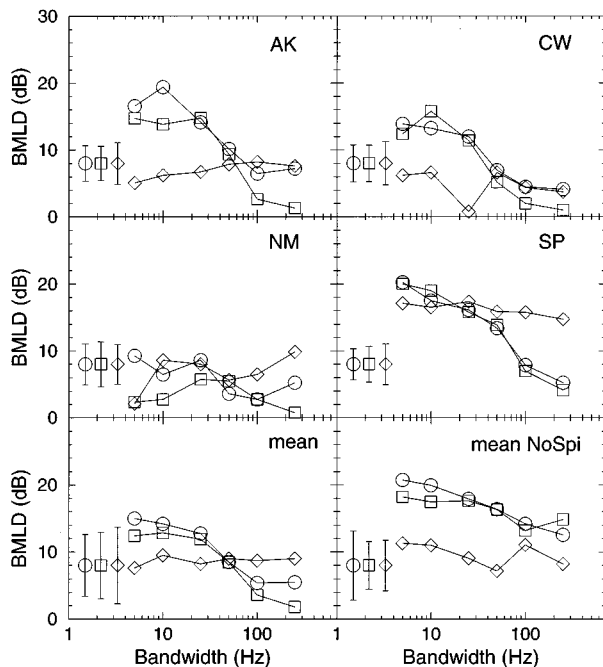

FIG. 8. The $N_\pi S_0$ BMLDs as a function of masker bandwidth for 125 Hz (○), 125-Hz transposed (□), and 4 kHz (◇). The top four panels show data for individual subjects; the bottom left panel shows the average of the four subjects. In addition, the bottom right panel shows BMLDs for the same four subjects for the $N_0 S_\pi$ condition.

for the $N_\pi S_0$ condition and the bottom right panel is the average for the same group of subjects for the $N_0 S_\pi$ condition.

The 125-Hz and the 125-Hz transposed $N_\pi S_0$ BMLDs again show a large variance among subjects. However, the two types of BMLDs are very similar for each individual subject. We can see that the BMLDs for the 125-Hz conditions are not always larger than those for the 4-kHz condition. Especially at the larger bandwidths the reverse is found for three of the four subjects. The averaged BMLDs of the four subjects in Fig. 8 also show that transposed BMLDs are larger than 4-kHz BMLDs up to 50-Hz masker bandwidths and that the opposite is true at larger bandwidths.

Comparing the averaged results in Fig. 5 for the 125-Hz and the 125-Hz transposed conditions we see that, in general, the BMLDs are larger for the $N_0 S_\pi$ condition than the corresponding BMLDs for the $N_\pi S_0$ condition. For the 4-kHz condition, on the other hand, $N_0 S_\pi$ and $N_\pi S_0$ BMLDs are rather similar. This is expected because at 4 kHz binaural processing has to rely on information present in the envelopes, which are indentical for the $N_0 S_\pi$ and the $N_\pi S_0$ conditions.

IV. DISCUSSION

The central question of this study is whether the difference between low- and high-frequency BMLDs is caused by the loss of fine structure information in the auditory periphery, prior to the binaural processing, or is due to poorer binaural processing at high frequencies. With the transposed stimuli we are able to provide the binaural processor with dichotic stimuli that contain similar temporal information in a high-frequency channel as is usually available in a low-frequency channel. If binaural processing capabilities are comparable at low and high frequencies, we expect very similar BMLDs for both types of stimuli.

We found nearly identical BMLDs for the 125-Hz low-frequency and the 125-Hz transposed conditions for both $N_0 S_\pi$ and $N_\pi S_0$ (cf. Figs. 5 and 8). The dependence on bandwidth is very similar and the tendency for low-frequency $N_\pi S_0$ BMLDs to be considerably smaller than $N_0 S_\pi$ BMLDs is also observed for the corresponding transposed conditions. A striking result is that low-frequency and transposed $N_\pi S_0$ BMLDs at bandwidths above 50 Hz are smaller than the high-frequency BMLDs at 4 kHz (Fig. 8). Thus, the notion that low-frequency BMLDs are larger than high-frequency BMLDs is not always true and, apparently, the additional fine-structure information in the transposed stimulus can affect binaural processing negatively in the case of $N_\pi S_0$.

Returning to the central question of this study, the similarity between the 125-Hz and the transposed 125-Hz BMLDs suggests that the information that was coded in the envelope of the 4-kHz carrier and the information present in the 125-Hz waveform is indeed processed similarly. This is in line with the assumption of Colburn and Esquissaud (1976).

For the 250-Hz data shown in Fig. 6 there is not such a clear correspondence between the transposed and low-frequency stimuli. On average the transposed BMLDs are
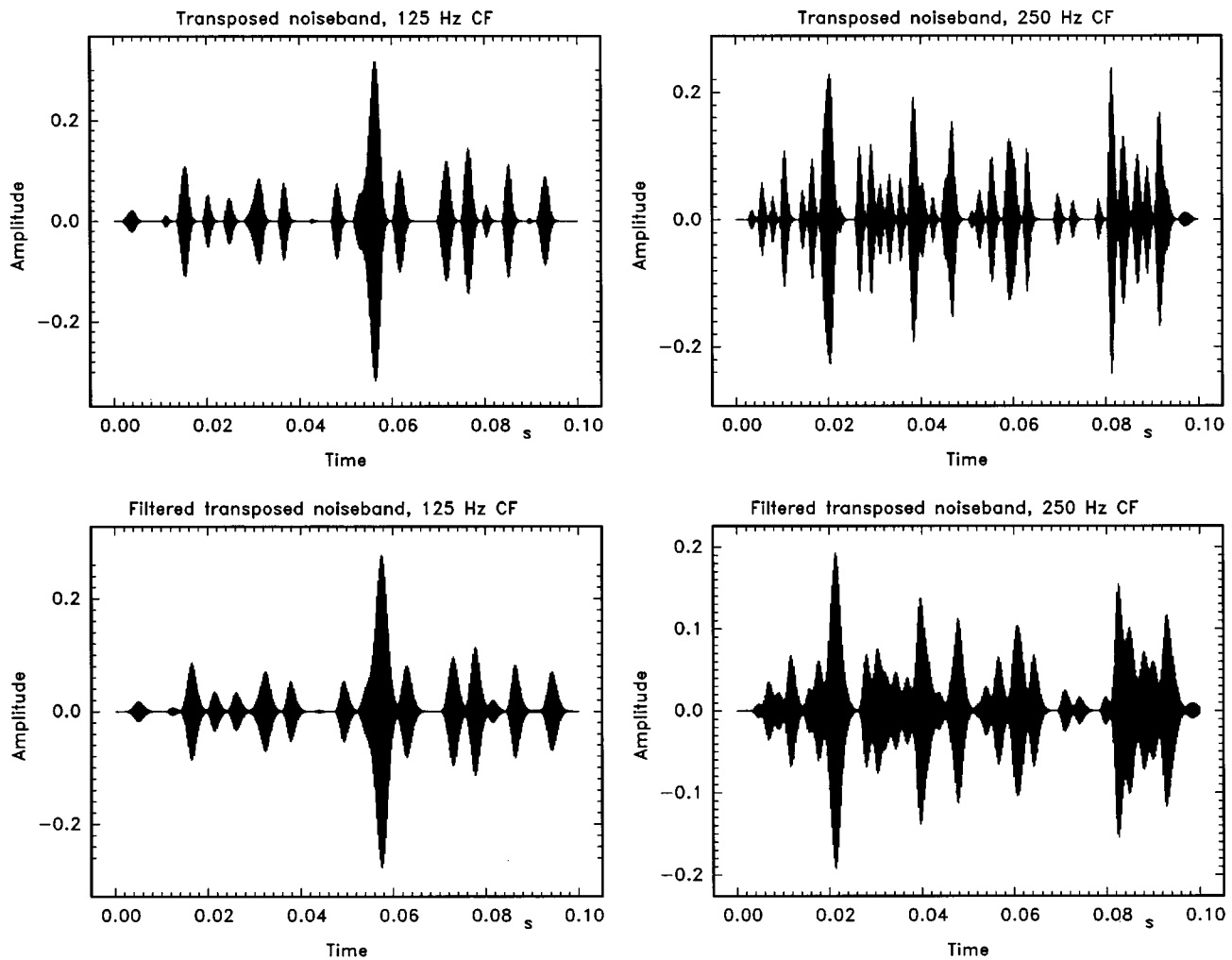
FIG. 9. The waveforms of a 50-Hz-wide band of noise at 125 Hz (left panels) and 250 Hz (right panels) after transposition to 4 kHz. These two conditions are shown before (top panels) and after filtering with a gammatone filter at 4 kHz with an ERB of 456 Hz.

6.5 dB smaller than the low-frequency BMLDs. Nevertheless, we can see that for the 250-Hz case there is a clear increase of transposed BMLDs with respect to 4-kHz BMLDs at the narrowest bandwidths. This shows that the addition of the extra envelope information can still improve binaural processing.

One could argue that for the 250-Hz transposed data, binaural processing at the wider bandwidths is hampered by the rate of interaural time and intensity fluctuations which increases with masker bandwidth. For the corresponding low-frequency condition the auditory filter limits this rate to about 50 Hz which could account for the more efficient processing of these low-frequency stimuli. However, from the study by Bernstein and Trahiotis (1992), it appears that such a rate limitation does only become effective at rates above 160 Hz, which implies that this can not have played an important role in our stimuli. In addition, for the comparable situation at 125 Hz, there is no difference between the low-frequency and the transposed stimuli, even at the large bandwidths.

Another reason could be that the auditory filter at 4 kHz modifies the envelope of the 250-Hz transposed stimulus. The central three bands of the stimulus span a total band-width that is comparable to the auditory filter bandwidth at 4 kHz. For a 125-Hz transposed condition this bandwidth is approximately half that of the 250-Hz transposed stimulus. The effect of filtering a 125-Hz and a 250-Hz transposed stimulus with a gammatone filter centered at 4 kHz and an ERB of 456 Hz is shown in Fig. 9. It is clear that as a result of this filter, the minima in the envelope of the waveform are less wide and have a less steep flank. For the 125-Hz transposed stimulus this effect is less prominent. These effects may explain the lack of correspondence between transposed and low-frequency BMLDs at 250 Hz.

As mentioned before, for 125-Hz conditions and for 125-Hz transposed conditions, $N_\pi S_0$ BMLDs are smaller than $N_0 S_\pi$ BMLDs. An explanation for this difference is related to the fact that for the $N_\pi S_0$ condition the waveforms at both ears are essentially out of phase. The time lag associated with this phase difference is 4 ms for a low-frequency stimulus centered at 125 Hz. This is large with respect to the time lags that occur in daily life as a result of the spatial separation of the ears. In models of binaural processing it has therefore been assumed that the auditory system cannot process these large time differences very efficiently (Langford and Jeffress, 1964; Colburn, 1977).

TABLE I. The decorrelation factor γ which is used in the EC theory is calculated for several bandwidths. This factor appears in the formula for the $N_\pi S_0$ BMLD. The average BMLDs for all subjects are used to derive γ. The second and third columns show γ values for the 125-Hz and the 125-Hz transposed condition, respectively.

Bandwidth (Hz)	Low freq. γ	Transposed γ
5	0.95	0.91
10	0.94	0.93
25	0.92	0.90
50	0.76	0.76
100	0.48	0.18
250	0.52	−0.30

Another explanation for the difference between low-frequency $N_0 S_\pi$ and $N_\pi S_0$ BMLDs can be given on the basis of the equalization and cancellation (EC) model by Durlach (1972). In the EC model it is assumed that in the equalization stage an internal delay is selected such that optimal noise reduction is obtained after the cancellation stage, with the limitation that the internal delay may not be longer than the length of half a signal cycle.

For an $N_\pi S_0$ stimulus, the equalization step will delay the waveforms in one of the two ears by half a period of the central component of the masker spectrum. When the bandwidth of the N_π masker is very small, the improvement of the signal-to-noise ratio after the cancellation step will be very large. However, when the bandwidth increases, the autocorrelation function of the N_π masker will be more damped and the cancellation step will not result in such a large improvement of the signal-to-noise ratio. Therefore, the difference between the $N_\pi S_0$ and $N_0 S_\pi$ BMLDs is predicted to increase with increasing masker bandwidth.

Using the EC theory, we can directly calculate the amount of decorrelation through the internal delay applied in the equalization step. According to this theory, the difference Δ in BMLD between $N_0 S_\pi$ and $N_\pi S_0$ is given by

$$\Delta = 10 \log \left(\frac{k - \gamma}{k - 1} \right). \tag{1}$$

Here, k is a factor that represents internal errors of the signal representation and γ represents the masker decorrelation through internal delay. If we adjust k such that the $N_0 S_\pi$ BMLDs are predicted correctly, we can use the above formula to derive the value for γ from the $N_\pi S_0$ BMLDs. The result of such a calculation is shown in Table I for the 125-Hz and the 125-Hz transposed conditions.

While for bandwidths up to 25 Hz the γ values are above 0.9 for both conditions, they decrease for the larger bandwidths. This provides support for the idea that an increase in bandwidth leads to a decrease in the correlation of the *internally delayed N_π* masker. In this respect it is interesting to note that the γ values for the 125-Hz condition remain constant for the largest masker bandwidths, which probably reflects the bandwidth of the 125-Hz auditory filter. On the other hand, the γ values for the 125-Hz transposed condition decrease further even for larger bandwidths. Since the transposed stimuli are centered at 4 kHz, the auditory filter bandwidth is no limiting factor for bandwidths up to 250 Hz.

This argument about differences between $N_0 S_\pi$ and $N_\pi S_0$ is not applicable to standard high-frequency conditions. Here, only the envelope is available for binaural processing, which has an interaural envelope correlation of 1 for the N_π reference interval. Therefore, in contrast to the 125-Hz $N_\pi S_0$ condition, binaural processing of an N_π masker at 4 kHz does not have to rely on an internal delay. This may explain the larger BMLDs for the 4-kHz $N_\pi S_0$ condition as compared to the 125-Hz and 125-Hz transposed conditions at larger bandwidths.

V. SUMMARY

Our results show that for high carrier frequencies, introducing fine-structure information, normally available at low frequencies, in the envelope can improve as well as hamper binaural processing with respect to a situation where such fine structure is not available. Both for $N_0 S_\pi$ and $N_\pi S_0$, the results for the 125-Hz and 125-Hz transposed conditions are very similar, suggesting very similar binaural processing at high and low frequencies. These data suggest that most of the differences between low-frequency and high-frequency binaural detection can be explained by the frequency-dependent loss of fine-structure information prior to the binaural processor.

ACKNOWLEDGMENTS

APPENDIX: Transposed stimulus properties

In this Appendix, the spectrum of the transposed stimulus is studied in detail and an analysis is made of the binaural spectrum detection performance that is expected if subjects are assumed to listen to one of the off-frequency bands of the transposed stimulus spectrum.

For the generation of a transposed masker stimulus, first a narrow-band noise is generated and half-wave rectified. In Fig. A1 the spectrum of this half-wave rectified narrow-band noise is shown. The structure of this spectrum can be understood when we first consider the spectrum of strongly clipped noise (cf. Lawson and Uhlenbeck, 1950). A strongly clipped noise, $c_g(t)$, is obtained from a narrow-band Gaussian noise, $g(t)$, by defining that $c_g(t) = 1$ for $g(t) \geq 0$, and $c_g(t) = -1$ for $g(t) < 0$. Using this definition for $c_g(t)$, the half-wave-rectified waveform, $h_g(t)$, can be written as

$$h_g(t) = \tfrac{1}{2}(1 + c_g(t)) g(t). \tag{A1}$$

When this equation is transformed to the frequency domain, we obtain

$$H_g(\omega) = \tfrac{1}{2} G(\omega) + \frac{1}{4\pi} C_g(\omega) \otimes G(\omega). \tag{A2}$$
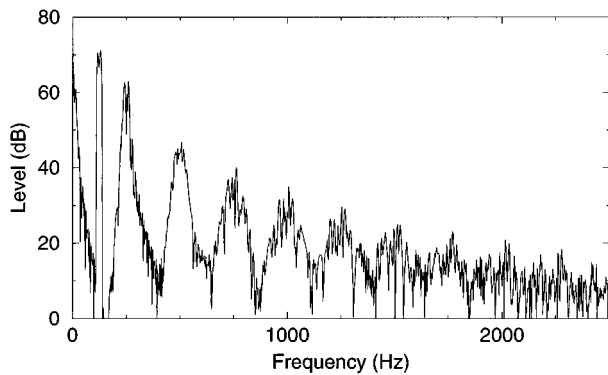
FIG. A1. The spectrum of a 25-Hz-wide 125-Hz half-wave rectified band of noise. Most of the stimulus energy is near 0 Hz. The spectrum of the noise band can be found at it original place. Additional peaks in the spectrum are at the even harmonics of the noise band.

The first term in this equation shows that the original noise spectrum, $G(\omega)$, comes back in the spectrum of the half-wave-rectified noise. In order to understand the second term in Eq. (A2) we have to analyze the spectrum of the strongly clipped noise, $C_g(\omega)$. This spectrum comprises a somewhat smeared version of $G(\omega)$ plus additional spectral bands centered at the odd harmonics of the center frequency of $G(\omega)$. The spectral level of these spectral bands decreases with increasing frequency (cf. Lawson and Uhlenbeck, 1950). Since $G(\omega)$ overlaps with one band from $C_g(\omega)$, the spectrum of $h_g(t)$ contains a smoothed triangular peak around $\omega = 0$ with a bandwidth proportional to the bandwidth of g. And since C_g has peaks at every *odd* harmonic, we will, after convolution, find peaks in the spectrum of h_g at the *even* harmonic frequencies.

In the following step in the transposed stimulus generation, the high-frequency portions of the spectrum are attenuated by a 500-Hz low-pass filter. After multiplication with the 4-kHz carrier, the resulting spectrum is symmetrical around 4 kHz and its energy density decreases with increasing spectral distance from the carrier frequency. The spectrum of this transposed stimulus is shown in Fig. 3.

When the spectra of a transposed masker and a transposed masker plus signal are compared, differences can be found in all frequency bands in the spectrum. In order to find out whether the binaural system might benefit from listening to one of the sidebands of the transposed spectrum, some further analysis was done. It was assumed that at high frequencies, binaural detection can be described by the subject's sensitivity to a decrease in the envelope correlation due to the addition of the signal (Bernstein and Trahiotis, 1996). For this purpose the interaural envelope correlation for each sideband was calculated separately and compared to the envelope correlation of the central three and the central five spectral components. We found that the envelope correlation in none of the sidebands changed more than in the central spectral part and therefore we do not expect that detection can benefit from separately listening to these sidebands.

[1]In order to show how the loss of fine structure may impair binaural detection we assume that at low frequencies, effectively, the full waveform is processed in a correlator, while at high frequencies the envelope of the waveform is processed. Under these assumptions we obtain a 3 dB *smaller* BMLD at high frequencies. Note that when the covariance is used instead of the correlation, the BMLD at high frequencies should be 3.6 dB *larger* compared to low frequencies (cf. van de Par and Kohlrausch, 1995).

[2]Due to the nonlinear interaction between the signal and the masker in a transposed stimulus, it is not possible to separate the signal energy from the masker energy. In order to get an estimate of the signal-to-masker ratio for a transposed stimulus we calculated the quantity $(E_{m+s} - E_m)/E_m$, where E_m and E_{m+s} are the energies of the transposed masker and of the transposed masker plus signal, respectively. We found that the average value of this quantity is very close to the signal-to-masker ratio of the underlying low-frequency stimulus.

[3]Due to the bandpass filtering that was applied for this set of experiments the time that was needed to calculate one trial for our experiments increased by several seconds. Therefore, no bandpass filtering was applied in the rest of the experiments.

[4]Initially, ramps of 20 ms were used. These were found not to be sufficiently long to avoid audible spectral splatter. Some subjects reported hearing the on- and offsets of the signal for the narrow-band N_0S_0 conditions. Using the 50-ms ramps instead increased the thresholds for these subjects to a level comparable to that of other subjects.

Bernstein, L. R., and Trahiotis, C. (**1992**). ''Detection of antiphasic sinusoids added to the envelopes of high-frequency bands of noise,'' Hear. Res. **62**, 157–165.

Bernstein, L. R., and Trahiotis, C. (**1996**). ''On the use of the normalized correlation as an index of interaural envelope correlation ,'' J. Acoust. Soc. Am. **100**, 1754–1763.

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

Bos, C. E., and de Boer, E. (**1966**). ''Masking and discrimination,'' J. Acoust. Soc. Am. **39**, 708–715.

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.

Colburn, H. S., and Esquissaud, P. (**1976**). ''An auditory-nerve model for interaural time discrimination of high-frequency complex stimuli,'' J. Acoust. Soc. Am. **59**, S23.

Durlach, N. I. (**1964**). ''Note on binaural masking level differences at high frequencies,'' J. Acoust. Soc. Am. **36**, 576–581.

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

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.

Green, D. M., and Swets, J. A. (**1974**). *Signal Detection Theory and Psychophysics* (Wiley, New York); reprinted by Krieger, New York.

Henning, G. B. (**1974**). ''Detectability of interaural delay in high-frequency complex waveforms,'' J. Acoust. Soc. Am. **55**, 84–90.

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,'' Hear. 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.

Lawson, J. L., and Uhlenbeck, G. E. (**1950**). *Threshold Signals* (McGraw-Hill, London).

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

McFadden, D., and Pasanen, E. G. (**1976**). ''Lateralization at high frequencies based on interaural time differences,'' J. Acoust. Soc. Am. **59**, 634–639.

Metz, P. J., von Bismarck, G., and Durlach, N. I. (**1968**). ''Further results on binaural unmasking and the EC model. II. Noise bandwidth and interaural phase,'' J. Acoust. Soc. Am. **43**, 1085–1091.

Palmer, A. R., and Russell, I. J. (**1986**). ''Phase-locking in the cochlear nerve of the guinea pig and its relation to the receptor potential of the inner hair cells,'' Hear. Res. **24**, 1–15.

van de Par, S., and Kohlrausch, A. (**1995**). ''Analytical expressions for the envelope correlation of certain narrowband stimuli,'' J. Acoust. Soc. Am. **98,** 3157–3169.

Perrott, D. R., and Musicant, A. D. (**1977**). ''Minimum auditory movement angle: Binaural localization of moving sound sources,'' J. Acoust. Soc. Am. **62,** 1463–1466.

Weber, D. L. (**1978**). ''Suppression and critical bands in band-limiting experiments,'' J. Acoust. Soc. Am. **64,** 141–150.

Zurek, P. M., and Durlach, N. I. (**1987**). ''Masker-bandwidth dependence in homophasic and antiphasic tone detection,'' J. Acoust. Soc. Am. **81,** 459–464.