The contribution of static and dynamic interaural differences to low-frequency BMLDs

A. Kohlrausch, J. Breebaart, S. van de Par

IPO Center for Research on User-System Interaction
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

1. Introduction

Binaural Masking Level Differences (BMLDs) are still a popular means to study binaural interaction in auditory perception. Adding an out-of-phase sinusoid to an in-phase Gaussian-noise masker (a condition labelled NoSπ) causes the interaural correlation of the stimulus to decrease, a cue to which the auditory system is particularly sensitive. On the other hand, the addition of the signal causes both the interaural differences of time or phase (ITD, IPD) and the interaural intensity difference (IID) to fluctuate around a mean of zero. The rate of these fluctuations depends on the effective bandwidth of the noise masker, i.e. the bandwidth after peripheral filtering in the inner ear. In terms of spatial perception, we can state that the internal image of an NoSπ stimulus is centered in the head (no lateralization cue), but that it is less compact than the image of the noise alone. A drawback of the BMLD paradigm lies in the difficulty in linking the results and models to lateralization and localization phenomena.

The present paper introduces an experimental procedure based on a noise type different from Gaussian noise, which allows us to separately investigate the contribution of static (i.e., lateralization) and dynamic (i.e., decorrelation) cues in a BMLD setting. In addition, the procedure allows us to create only IID or only ITD cues. In experiments using Gaussian-noise maskers, both these cues are available and their standard deviations are coupled (see Zurek, 1991). We expect that this experimental paradigm will help in further improving binaural models and finally allowing a unification of models that are able to predict localization and those that are able to predict binaural signal detection thresholds (cf. Stern and Shear, 1996).

2. Method

2.1 Properties of multiplied noise

Multiplied noise is created by multiplying a low-pass noise with a sinusoid. The resulting bandpass noise has a center frequency equal to the frequency of the sinusoid and a bandwidth of twice the lowpass cut-off frequency. Two differences with respect to Gaussian noise should be mentioned: the envelope distribution of multiplied noise is equal to the positive half of a Gaussian distribution with a zero mean, while it is a Rayleigh distribution for Gaussian noise. The zero crossings of multiplied noise occur at regular time instances and have a temporal separation equal to half the period of the sinusoid used in creating the noise (for more details, see van der Heijden and Kohlrausch, 1995).

This latter property can be exploited in masking experiments, if the test-signal frequency is identical to the center frequency of the noise. In this case, the fine-structures of noise and signal are phase coupled and the relative phase ϕ between masker and signal becomes an additional parameter. The following description concentrates on the interaural differences for a dichotic NoSπ condition with a multiplied-noise masker and a sinusoidal test signal. If the relative phase ϕ is zero, the combined stimulus will only show interaural intensity differences, because the fine-structures in both ears remain in phase or have a phase difference of 180 degrees. On the other hand, if the relative phase equals 90 degrees, the combined stimulus only has interaural time differences and no interaural intensity differences, because the envelope values in both ears are always the same.

The situation is comparable to binaural tone-on-tone masking conditions (see Yost, 1972), where the relative phase angle between masker and signal determines which of the binaural cues are introduced on adding the signal. In tone-on-tone masking, the interaural cues are static and lead to a lateralization of the NoSπ stimulus. In our experiments with multiplied-noise maskers, the interaural cues vary dynamically with a rate proportional to the bandwidth of the masker. That is, the interaural differences are distributed around a mean of zero and the NoSπ stimulus will thus be perceived in the center of the head, but will be less compact than the noise masker alone.

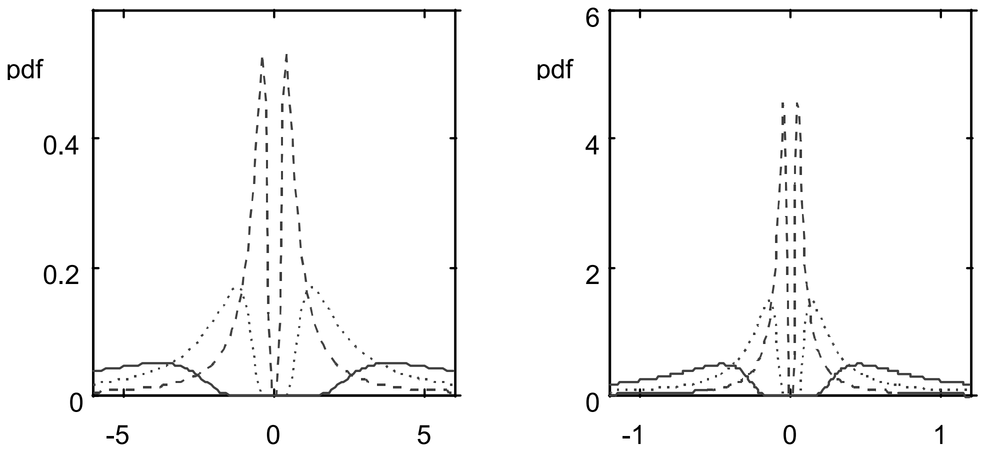

Fig. 1: Distributions of interaural intensity differences (left panel) and interaural phase differences (right panels) for a multiplied-noise masker and a sinusoidal signal in the condition NoSπ. Three different signal-to-noise ratios: -10 dB (continuous curve), -20 dB (dotted curve), -30 dB (dashed curve).

The two panels in Fig. 1 show the distributions of the interaural intensity difference (IID, left panel) and of the interaural phase difference (IPD, right panel). The three curves in each panel are for signal-to-masker ratios of -10, -20, and -30 dB, respectively. As explained before, the distributions are symmetric with mean zero, and they become narrower with decreasing signal level. From these graphs, one can relate an experimentally determined masked threshold to, e.g., the standard deviation of the dynamically varying IID and IPD cues and, by comparing thresholds in IID and IPD conditions, derive a kind of `time-intensity trading ratio' for dynamically varying interaural cues.

A slight variation to the procedure in creating the multiplied noise also allows us to induce static interaural cues in the NoSπ condition. Before the Gaussian lowpass noise is multiplied by the sinusoid, a DC offset is added to the noise waveform. The amount of this offset, relative to the rms value of the original Gaussian noise, is indicated by the parameter m. After multiplication, we again have a bandpass noise, but now with an increased central component. The zero crossings of this noise are still regular, but its envelope is no longer equal to the positive half of a Gaussian distribution. In Fig. 2, we show such distributions for three values of m, namely 0, 1, and 2. With increasing m, the envelope distribution becomes narrower and, asymptotically for large m, it will become equal to the constant value of a sinusoid.

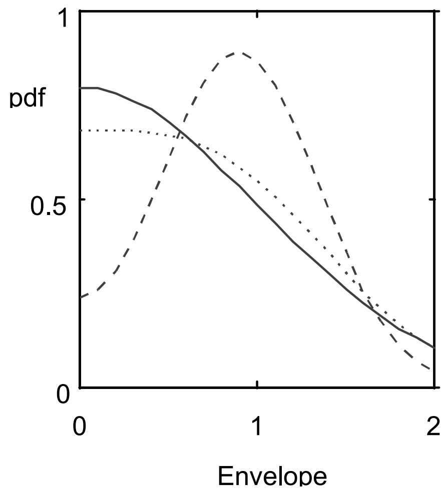

Fig. 2: Envelope distributions for multiplied noise for three values of the DC paramter m: m=0 (continuous curve), m=1 (dotted curve), m=2 (dashed curve).

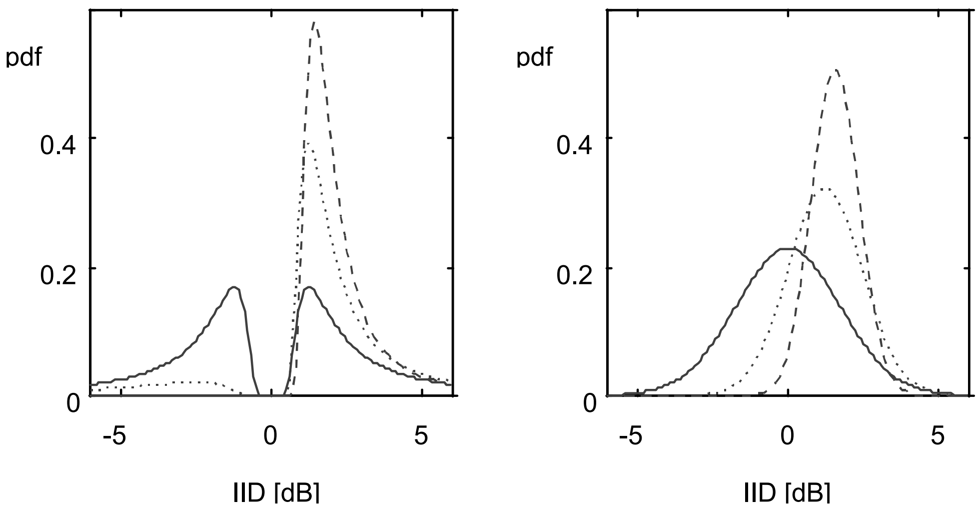

Fig. 3: Distributions of IIDs in the condition NoSπ for an S/N ratio of -20 dB. In the left panel, multiplied noise is the masker and the signal is a sinusoid. In the rigth panel, a sinuosid is the masker and multiplied noise is the signal. The three curves are for three values of m, namely m=0 (continuous curve), m=1 (dotted curve), m=2 (dashed curve).

The left panel in Fig. 3 shows the distribution of the IID in an NoSπ condition as a function of the static component m for a fixed signal-to-noise ratio. With increasing m, the mean of the distribution shifts away from zero, implying that the presence of the signal can be detected on the basis of lateralization. The stronger this shift, the narrower the distribution becomes. The parameter m thus allows us to create a continuum of masking conditions with a different mixture of dynamic and static binaural cues, and so allows us to compare the signal detectability across these conditions.

A different way of introducing static and dynamic interaural differences is by using a sinusoid as masker and multiplied noise as the test signal (i.e., exchange the role of masker and signal compared to the above situation). Also in this case, we can selectively introduce IIDs or IPDs, and we can also vary the relative amount of dynamic vs. static cues. The major difference is that, for the same signal-to-masker ratios, the distributions have a much smaller standard deviation of the IID and the IPD. This is shown in the right panel of Fig. 3 for the distributions of IIDs and the three values 0, 1, and 2 for m. It should be noted that the corresponding graphs for the distributions of IPDs have a similar shape.

2.2 Procedure and stimuli

Masked thresholds were determined using a 3-Interval Forced-Choice procedure with adaptive level adjustment. The three masker intervals of 400 ms duration were separated by gaps of 300 ms. In one interval, a 300-ms signal was temporally centered in the masker interval. The masker was presented interaurally in phase, the signal was phase inverted. The subject's task was to indicate which of the three intervals contained the signal. A slightly different procedure was used in experiment 2 (see below). The signal level was adjusted using a two-down one-up rule (Levitt, 1971). The step size for level changes was 8 dB in the beginning of each run. It was halved after every second reversal of the signal level, until a step size of 1 dB was reached. A run was extended over 14 reversals. The first six reversals were discarded and the median value of the last eight reversals was used as the threshold value. Four threshold values were obtained for each parameter value and subject. Four subjects participated in the experiments and the means across their individual values are shown. Masked thresholds are expressed as signal-to-overall-noise ratio (S/N).

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. Masker and signal were both gated with 50-ms raised-cosine ramps and the masker was presented at an overall level of 65 dB SPL. The center frequency of the multiplied noise was always 500 Hz; sinusoidal signals had a frequency of 500 Hz.

3. Results

3.1 Experiment 1

In the first experiment, the masker was a multiplied noise with a bandwidth of either 10 or 80 Hz. The signal was added in such a way that either IIDs or ITDs were created in the interval containing the $S\pi$ signal. This gives four different experimental conditions, the results of which are indicated by the four different symbols in Fig. 4. This figure shows the mean masked thresholds as a function of the static component m. Data points at the left of the figure are for `pure' multiplied noise (no lateralization), for increasing values of m lateralization becomes an increasingly relevant cue.

For all conditions, the threshold values lie in a narrow range between -26 and -20 dB. A small, but systematic increase in thresholds is seen for increasing values of m.

The thresholds for the two masker bandwidths of 10 and 80 Hz do not differ systematically from each other, indicating that the rate of fluctuation of the interaural parameter does not influence detectability.

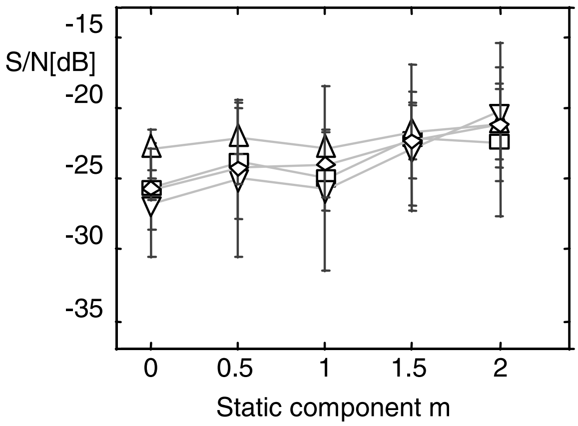

Fig. 4: NoSπ thresholds for multiplied-noise maskers of 10 or 80 Hz bandwidth and a sinusoidal signal. Squares: IIDs only, 10 Hz bandwidth; upright triangles: IIDs only, 80 Hz; inverted triangles: ITDs only, 10 Hz bandwidth; diamonds: ITDs only, 80 Hz bandwidth. Means and standard deviations across the results of four subjects.

3.2 Experiment 2

This experiment differed from the previous one in the exchange of the roles of masker and signal. The masker was now a 500-Hz sinusoid, and the signal was multiplied noise. Since it is well known that a sinusoid does not mask a noiseband very efficiently, one has to be cautious to ensure that thresholds in the dichotic condition are not determined by monaural cues. This goal can be achieved by adding the noisy test signal to the two reference intervals with an interaural phase difference of 0. Thus, the task in this experiment was to discriminate between NoSπ and NoSo.

Again, the same four combinations of noise bandwidth and type of interaural cue were tested and results are indicated by the same symbols as before (Fig. 5). Overall, thresholds are in the same range as in experiment 1 and increase somewhat with increasing m. Only ITD thresholds for a bandwidth of 80 Hz are systematically lower than the values in the other conditions.

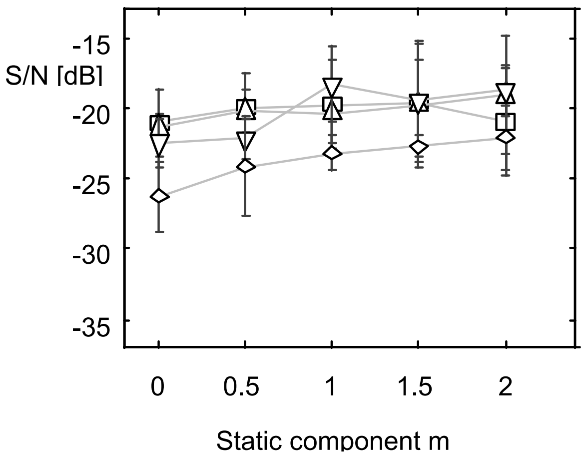

Fig. 5: As in Fig. 4, but the masker was now a sinusoid and the signal was multiplied noise.

3.3 Experiment 3

In this final experiment, we combined ITD and IID cues. This is possible by adding multiplied noise and sinusoid with a relative phase of 45 or 135 degrees. The idea behind this measurement was to test whether signal detectability is different if the two interaural cues point in the same direction (as for 45 degrees) or in opposite directions (as for 135 degrees). For reference, we also included the two conditions with ITDs only and IIDs only. This experiment was performed for multiplied-noise maskers with static component 0 and bandwidths of 10 and 80 Hz. In addition, we combined a tonal masker with the tonal signal, representing the extreme case of only static cues (this case corresponds to m equal to infinity).

The three different symbols in Fig. 6 indicate the different masker types, while the four values on the abscissa indicate the different interaural cues. In this figure we see that there is not only no difference between ITD and IID conditions, but that also the two mixed-cue conditions result in very similar thresholds.

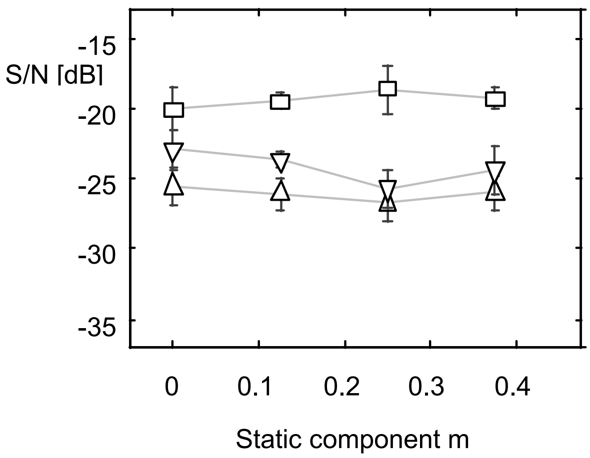

Fig. 6: Thresholds in the condition NoSπ for a sinusoid masked by a sinuoid (squares) or multiplied noise of 10 (triangles) or 80 Hz bandwidth (inverted triangles). The interaural cues vary along the abscissa as follows: Phase 0: IIDs only; 0.125: IIDs and ITDs pointing in the same direction; 0.25: ITDs only; 0.375: IIDs and ITDs pointing in opposite directions.

4. Discussion

As we have seen in the previous section, all experimental conditions lead to very similar thresholds in terms of the S/N ratio. From lateralization experiments it is known that ITDs and IIDs can compensate each other to a certain extent (time-intensity trading). In order to compare time-intensity trading with the ratio of just detectable deviations in time and intensity, we calculated the standard deviations for the distributions of the cues ITD and IID at threshold. For experiment 1 and m = 0, these values are about 4.5 dB and 185 μs, a ratio of about 40 μs/dB. For experiment 2, the ratio is about 25 μs/dB. These appear to be typical trading ratios for low-frequency stimuli (cf. Blauert, 1997).

In order to investigate the effects of both static and dynamic cues, we can analyze the distributions of the interaural cues in a different way: Fig. 7 shows the means and the standard deviations for the various values of m. Open symbols are for data from experiment 1, filled symbols for those from experiment 2. The asterisk shows, in addition, the value for a purely static cue. Both for the IID (left panel) and IPD (right panel), the standard deviations decrease somewhat when lateralization also becomes available as a cue.

Probably more interesting is the comparison of the open and filled symbols. This reveals that, for a similar value of the mean, totally different standard deviations of the interaural cue are found, differing by about a factor 3 to 5. This means that we cannot in a simple way derive a predictor for the binaural masked threshold by reducing the distribution of the interaural cue to one or two numbers. A better measure for the subject's sensitivity is obviously given by the S/N ratio, which has very similar values independent of the exact nature of the binaural cue. Such a result favors models that relate binaural thresholds to differences in the interaural crosscorrelation. Problems with such models so far are, however, that they cannot deal with static interaural level differences, or that they cannot predict correctly the observed constancy of binaural thresholds with masker bandwidth. The results presented in this study hopefully stimulate the development of more comprehensive binaural models for detection as well as localization.

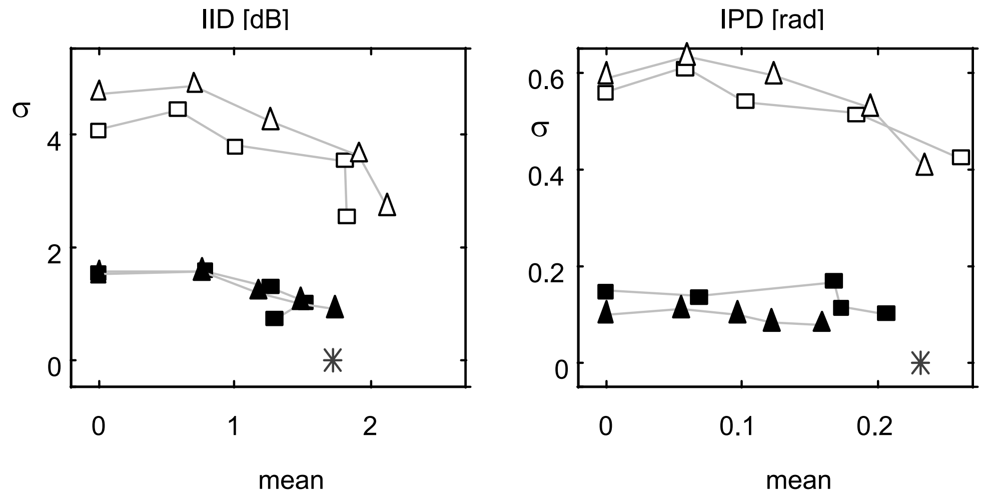

Fig. 7: Means and standard deviations of the interaural cues IID (left panel) and IPD (right panel) at threshold. Open symbols indicate results from experiment 1, filled those from experiment 2. Squares are for multiplied noise with 10 Hz bandwdidth, triangles for 80 Hz bandwidth. The asterisk indicates the static interaural differences from the tone-on-tone masking condition.

5. References

Blauert, J. (1997). Spatial hearing: the psychophysics of human sound localization. MIT Press, Cambridge.

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

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 der Heijden, M. and Kohlrausch, A. (1995). The role of envelope fluctuations in spectral masking. J. Acoust. Soc. Am. 97, 1800-1807.

Yost, W.A. (1972) Tone-on-tone masking for three binaural listening conditions. J. Acoust. Soc. Am. 52, 1234-1237.

Zurek, P. M. (1991). Probability distributions of interaural phase and level differences in binaural detection stimuli. J. Acoust. Soc. Am. 90, 1927-1932.