Proceedings of The Institute of Acoustics

DYNAMIC BINAURAL PROCESSING: MASKED THRESHOLD ADAPTATION WITH A NON-STATIONARY MASKER

Trevor M. Shackleton

Dept. of Physics, University of Surrey, Guildford, Surrey, GU2 5XH

INTRODUCTION

When faced with a natural auditory environment, a human listener has several problems to solve. The first of these is to determine when an event has occurred against a background of changing interfering noise. Secondly the listener has to determine what that event is, and whether it is worthy of further attention. At the same time, the listener determines where in the environment the event took place and in what direction it is moving.

It has become apparent that the use of two ears considerably enhances our ability to perform the detection and localisation tasks just mentioned. This improvement is due to the analysis of differences in the sound at the two ears. This analysis is generally called the binaural processing of interaural difference cues, the most studied natural cues being those of interaural intensity and time differences.

This paper will concentrate upon the detection problem, or more specifically the problem of detecting a tone pulse in a noise background with changing interaural parameters. It is worth pausing, however, briefly to outline how detection is improved in a static noise background.

Signal detection experiments with static maskers.
The first studies of signal detection in a static noise environment were made by Licklider [8] who found that the detectability of speech in noise was improved if either the speech or the noise (but not both) was inverted at one ear and Hirsh [5] who studied the detectability of a pure tone in white noise and found a similar (but larger) effect. We denote the signal by S, the noise by N and the interaural phase relationship by either π or 0 depending on whether the signal/noise was inverted or not, respectively. When the interaural relationships of the noise and signal are the same we refer to the condition as being homophasic (eg. N0S0 or NπSπ). If the relationships are different then the condition is antiphasic (eg. N0Sπ or NπS0). We further define a masking-level difference (MLD) to be the difference between any two detection thresholds (standard terminology refers to the MLD as being between the homophasic and antiphasic thresholds). Using this notation, we may restate the above results as: the N0S0 and NπSπ thresholds are the same as the monaural masked threshold, but the N0Sπ and NπS0 thresholds are between 5-20 dB lower for tones (ie. more detectable).

Jeffress et.al. [6] measured the detectability of a 500 Hz tone as the noise and tone were independently delayed at one ear, thus creating interaural time differences in both the tone and the noise. They found that the detectability was greatest when the difference in the delay of the noise and signal was equal to one-half of the period of the tone (ie. 1 ms). The detectability was least when the relative delay was equal to an integer number of periods.

DYNAMIC BINAURAL PROCESSING: MASKED THRESHOLD ADAPTATION

That is, if the signal is placed at one side of the head, the detectability will be greatest when the noise is on the opposite side, and least when it is on the same side of the head.

If we consider the bandwidth of noise effective in masking the signal to be narrow, then these results may be interpreted as corresponding to the NOSO and NπSπ conditions for integer period relative delays, and to NOSπ and NπSO for half period relative delays.

Robinson & Jeffress [10] measured the detectability of SO and Sπ tones of 500 Hz as a function of the correlation of the noise between the two ears. They found that for both SO and Sπ the threshold decreased slowly, but with increasing slope as the correlation was moved away from the homophasic point (ie. a correlation of +1 for SO and -1 for Sπ), so that at the antiphasic point the slope was very steep. They compared their results with those of Jeffress et.al. [6] using the assumption that the bandwidth of noise effective in masking the signal was 50 Hz. This figure was obtained from studies of monaural masking. They considered that a delay decorrelated the noise band and compared MLDs at equivalent correlations, the agreement was poor. However, Langford & Jeffress [7] calculated the noise auto-correlation function for a band-width of 100 Hz and found good agreement for all correlations except those around zero. There are several experiments which suggest that the wider bandwidth used by Langford & Jeffress is valid.

The above experiments illustrate how interaural time and correlation differences affect the detection of tones in a static environment, and what the relationship between the two parameters is. Until recently, however, the only binaural effects studied have been those involving static interaural cues. The notable exception has been the study of binaural beats. It has been tacitly assumed that once the static results have been found then the dynamic ones can be deduced from them, however recent research has shown that this may not be the case.

<u>Signal detection experiments with dynamic maskers.</u>
The experiments to be described in this section are essentially a frequency-domain measurement of the response to a dynamic binaural cue. In an experiment designed to determine the rate at which the binaural system could track modulated interaural time differences, Grantham & Wightman [3] constructed a noise-like masker which had a time-varying perceptual location between the ears (lateralisation). This stimulus was then used to mask a tone pulse located towards one side of the head (ie. with an interaural delay of 0.5 ms). This experiment, then, is a dynamic analogue of the Jeffress et.al. [6] experiment described above. As expected, the signal threshold was greatest when the signal and noise were on the same side of the head, whereas it was least when they were on different sides. As the rate at which the noise was tracked across the head increased it was found that all the thresholds tended towards the higher (homophasic) threshold. At zero modulation rate the MLD between the condition with colateral signal and noise and that with antilateral signal and noise was around 6 dB, this decreased to 3 dB at a modulation rate of 2.5 Hz and by a modulation rate of 10 Hz the MLD was practically zero.

In a later experiment Grantham and Wightman [4] measured the threshold of an S tone burst masked by noise with a sinusoidally modulated interaural correlation. They varied the modulation rate and the noise correlation at the instant of signal presentation. It was found that the thresholds determined at the instants corresponding to the classic N_0S_π and $N_\pi S_\pi$ conditions decrease from about 16 dB (for 500 Hz) at zero modulation rate to 12 dB at a modulation rate of 0.5 Hz and at a modulation rate of 4 Hz the MLD is almost zero.

The experiments considered above measure the dynamic behaviour of the binaural system by attempting to measure the modulation rate transfer function, in this sense they are a frequency-domain measurement. To the extent that the binaural system is linear, the same information may be derived from time-domain experiments (eg. a measurement of the "impulse response").

Signal detection experiments with variable duration masker "fringe".

The time-domain experiments we now consider have not previously been considered in terms of the dynamic response of the binaural system, but as measures of the static threshold. Their common feature is that either the masker is turned on, or its interaural parameters are changed, a certain length of time before the signal is presented. This "masker fringe" length is the independent variable in the experiments. McFadden [9] measured the MLD between N_0S_π and N_0S_0 conditions as a function of fringe length using a 125 ms long tone burst at 400 Hz. Robinson & Trahiotis [11] performed a similar experiment, but using a 500 Hz tone burst of durations 32 and 256 ms. Both these experiments showed that there was virtually no change in N_0S_0 threshold over the range of fringe lengths used. However, there was a drop of 2-5 dB in the N_0S_π threshold as the fringe length was increased from 0 ms to 500 ms, the shorter signal having the greater drop. Bell [1] performed a similar experiment in which the threshold of a S_π, 500 Hz tone burst of 125 ms duration was measured as a function of the fringe length after a transition from uncorrelated noise (Nu) to correlated noise (N_π). The reverse transition occured concurrently with signal termination. Under these conditions there was a drop of 3 dB in the threshold between fringe lengths of 0 ms and 100 ms, with little reduction after that. The thresholds in this experiment were about 4 dB higher than those in the previous experiments [9,11], although the method and signals were similar.

These results all suggest that the binaural system is slow relative to the monaural system at tracking dynamically varying cues. It has been suggested that this "sluggishness" may be incorporated into any model of binaural procesing by including a low-pass filter (or integrator) into the mechanism doing the binaural analysis [2,4].

EXPERIMENTAL DESIGN

This experiment was designed to test the integrator hypothesis by measuring the effect on the detectability of a tone burst of a brief change in an interaural noise parameter before the signal occurence. By varying the duration of the change and time of signal occurence it should be possible to calculate the characteristics of the filter/integrator. The interaural parameter that was chosen to be varied was the interaural correlation (ie. jumps between noise conditions N_0 and N_π were used), because of the ease of experimental realisation and the availability of reliable static threshold results. Work on a similar theme is being done at Göttingen University by B. Kollmeier et.al. (most of this work is published in German, except [13]).

The noise was turned on with a rise/fall time of about 20 ms in either No or Nπ configuration and after 200 ms the noise in one ear was inverted for a variable duration. At a variable delay after this the signal was presented concurrently with a signal on light. The noise and signal-on light were switched off 200 ms after this. The signal was a 1 ms pulse filtered through a 12% bandpass filter centered at 500 Hz. The signal thus lasted about 10 ms. The noise was wide-band gaussian, filtered only by the headphones. It was presented at a spectrum level of 40 dB. The equipment produced spurious, wide-band, monaural clicks when the noise inverters were switched, but since these were masked by noise 30 dB lower than that used in the experiment they were considered unimportant. The effect on monaural threshold due to the invertor switching was measured in a separate experiment using 3 subjects, the effect was found to be less than 0.5 dB (this is an upper limit set by the power of the t-test, visual inspection would put the difference at less than 0.2 dB).

A 2-IFC design was used, with a 400 ms gap between intervals, the next presentation began 400 ms after the subject's response. The 71% threshold was determined using a 2-down, 1-up tracking technique. A practice period with correct answer feedback consisting of 6 reversals was followed by the main measurement without feedback consisting of 11 reversals, of which the last 10 were used to determine the threshold.

The experimental conditions were arranged in a factorial structure. A repeated measures design was used, with the 4 inversion durations (4, 16, 64 and 256 ms) and the two phasic conditions (antiphasic & homophasic) being the between-subject factors. 16 subjects were used, so each combination was presented to two subjects. The within-subjects factors were the initial noise phase and signal delay (either 0.5, 3.7, 27 or 202 ms). The subjects attended 8 sessions of about 1/2 hour each. Each session had constant noise phase. Every signal delay was presented in each session. A measurement of the static threshold was made before each main measurement. The session and position in session for each noise-phase were arranged in a latin square, with signal delay as the latin letter, to minimise the effect of learning.

RESULTS

The static, homophasic and antiphasic, thresholds were found to vary by about 1 dB across conditions for any individual subject and noise phase. Data are therefore presented in terms of the MLD between the dynamic threshold and the static threshold (dynamic - static). The MLDs are plotted in Fig. 1 and are analysed statistically in Table 1 where a univariate analysis of variance for the repeated measures design is presented. In Table 1 the data has been split according to whether the conditions at the instant of signal presentation were homophasic or antiphasic. Columns 1 & 3 give the statistical significance level of the named effect tested against the error term nearest above it. Columns 2 & 4 give the percentage of overall variance which is accounted for by that effect, that is, a measure of the magnitude of that effect. Figs. 1a and 1b show the same antiphasic data but with axis and parameter interchanged for clarity. The data are shown averaged over noise phase and subject since these effects only cause a vertical shift in the curves (ie. they have non-significant interactions with other effects). In Fig. 1 the data points for different inversion durations correspond to different pairs of subjects.

As inversion duration is increased there is a rise in the dynamic-static MLD for both homophasic and antiphasic conditions. For a signal delay of 0.5 ms the homophasic MLD rises linearly from -2 dB to +1 dB as inversion duration is increased from 4 ms to 256 ms. For the other signal delays, the MLD rises from about -0.3 dB to +1 dB between inversion durations of 4 ms and 64 ms and remains constant up to the longer inversion. In the antiphasic condition however, the graph for a long signal delay falls slightly. After an initial rise of between 4 dB and 10 dB the MLDs at the other signal delays drop by about 4-6 dB between inversion durations of 64 ms and 256 ms. The integrator hypothesis would predict monotonically increasing MLDs for the antiphasic conditions and decreasing MLDs for the homophasic conditions as inversion duration was increased. This inversion duration effect is not statistically significant. If however we study the percentage of variance this effect accounts for (7-10%), we are inclined to believe that it is a real effect, masked in the statistics by being subject to a non-powerful test.

As signal delay is increased in the antiphasic case the MLD for the shortest inversion duration (4 ms) remains constant, but the MLDs for the longer inversion durations all begin to decrease after a signal delay of 3.7 ms. The decay for an inversion duration of 16 ms appears to take longer initially than for the longer inversion durations. In the homophasic case the MLDs for the three shorter inversion durations remain constant until the signal delay reaches 27 ms and then decrease by about 2 dB between 27 ms and 202 ms. The MLD for an inversion duration of 256 ms remains constant. For the homophasic data, a t-test shows that the average MLD is significantly (at 1%) positive (the integrator hypothesis would suggest that the MLDs should be negative for the homophasic condition). The analysis of variance shows that the signal duration effect is very highly significant, for both homophasic and antiphasic conditions, as is the interaction between signal delay and inversion duration (the interaction corresponds to differences in the shape of the graphs).

DISCUSSION

The purpose of the present experiment was to test the validity of the integrator hypothesis [2,4]. A prediction of this hypothesis is, crudely, that all the thresholds for conditions with a period of inverted noise before them will be intermediate between the homophasic and antiphasic thresholds. Consider a period of N0 noise followed by a short burst of Nπ before a reversion back to N0. If we assume that the integration is done over correlation, then the net correlation at the time of signal onset will be somewhere between +1 and -1. From the variation of threshold with correlation [10], we deduce that in the dynamic condition there will be a threshold intermediate between the homophasic and antiphasic thresholds. Any model which averages at the periphery will similarly predict intermediate thresholds. Since the MLDs plotted in Fig. 1 are dynamic threshold minus static threshold, we expect negative MLDs for the homophasic condition and positive MLDs for the antiphasic conditions (Figs. 1a and 1b). We would expect the magnitude of the MLD to depend both on the inversion duration (longer inversions resulting in larger MLDs) and upon the signal delay (longer delays resulting in smaller MLDs).

Table 1. Analysis of variance for the dynamic-static threshold MLD. A univariate analysis of variance is used. Those rows without significance levels stated are the error terms to be used in the following block. Key to terms: Inversion duration, I; Noise Phase, N; Signal Delay, D; Subject nested within inversion duration, SwI.

Source of variance	df	HOMOPHASIC		ANTIPHASIC	
		% level of significance	% of variance accounted for	% level of significance	% of variance accounted for
Within cells error	192		62.4		16.5
SwI (error)	4		12.6		7.8
I	3	25	7.6	13	10.4
N x SwI (error)	4		0		0
N	1	5	1.7	0.01	1.6
N x I	3	25	2.0	10	0.8
D x SwI (error)	12		0		1.6
D	3	0.01	6.7	0.01	42.0
D x I	9	-	0.3	0.1	16.3
N x D x SwI(error)	12		0		0
N x D	3	25	0.8	12	0.9
N x D x I	9	10	5.9	28	1.1

Figure 1. Difference between antiphasic dynamic and static thresholds (dynamic - static) for a 500 Hz tone burst of duration 10 ms masked by wide-band white noise.
The masker in the static condition was either N0 or Nπ gated on from silence (200 + inversion duration + signal delay) ms before the signal (either Sπ or S0 respectively) was presented. In the dynamic noise, the same initial noise was used, but after 200 ms the noise was inverted for the inversion duration, the tone was then presented after the signal delay. In both cases the noise was gated off 200 ms after the signal onset.
Data are shown averaged over both subjects and initial noise phases.
Fig. 1a shows the MLD as a function of inversion duration with signal delay as parameter. A signal delay of 0.5 ms is denoted by O, 3.7 ms by +, 27 ms by ✱, and 202 ms by ◆. Fig. 1b shows the same data but plotted as a function of signal delay with inversion duration as parameter. The inversion duration for 4 ms is denoted by O, 16 ms by +, 64 ms by ✱ and 256 ms by ◆.

Fig. 1a

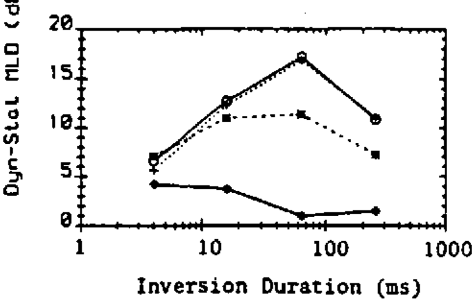
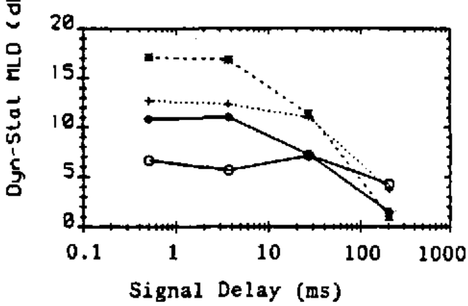


Fig. 1b



The signal delay effect does obey these loose predictions, however the inversion duration effect disagrees with them since the slope is in the wrong direction for the homophasic data and the antiphasic MLD does not increase monotonically. Also, the homophasic MLDs are mostly positive, in direct contradiction of the integrator hypothesis.

Since thresholds greater than the monaural threshold are obtained, especially in the homophasic condition, but also in certain antiphasic conditions, we may conclude that the binaural mechanism cannot simply average the interaural cues. However, we still need to account for Grantham & Wightman's [3,4] findings that the binaural system can only track low frequency modulations. To do this we need to look at the output from the interaural difference detector mechanism.

An influential model in the development of binaural theory is the Webster-Jeffress lateralisation model [eg. 12]. This model depends for its operation upon neural summation of impulses arriving simultaneously from the two ears. The simultaneity is achieved by introducing a set of complementary delays into the left and right channels. The resulting output of the summators has a maximum at a position corresponding to the interaural time difference. In essence this network calculates the cross-correlation function of the noise from the two ears. This array is useful in binaural detection, since in any condition other than the homophasic condition, the positions of the noise alone and noise plus signal stimuli will be different. Thus the detection problem resolves into one of deciding whether the position of the array output is due to noise alone or noise plus signal.

McFadden [9] suggested that the usefulness of the masker fringe in his experiment was that it provided a base noise position from which it was possible to perceive the image movement provoked by the addition of the signal. He also suggested that the masker fringe may instead be used to refresh the subject's memory of the interaural parameters of the noise alone, so that when the signal was presented a smaller intensity would be needed to convince the subject that what had occured was not merely a random fluctuation of the noise. The first of these hypotheses refers primarily to the use of signal onset cues, whereas the second is more concerned with the ongoing cues present throughout the signal presentation.

What happens to the output of these detectors when we vary the interaural parameters of the noise? For simplicity we will assume that the integration time of the system peripheral to the binaural analyser is short relative to the response time of the "higher centres.

The output of the Webster-Jeffress lateralisation array will track the change in noise lateralisation. McFadden's first hypothesis would predict an increase in threshold if addition of the signal caused the image to move in the same direction as the noise. There would probably be a reduction in threshold if the image movement was reversed upon adding the signal. His second hypothesis would probably have the same effect.

In this experiment in the antiphasic conditions both the noise inversion and signal addition cause a "movement" from a centred image to a diffuse image filling the head. We would therefore predict increased thresholds. In the homophasic conditions the noise inversion causes an image movement but the signal addition does not. It is quite likely from McFadden's hypotheses that this condition will also result in an increase in threshold. If we assume that longer inversion durations and signal delays reduce these effects, then we would predict a drop in threshold for long signal delays and inversion durations.

In conclusion we have presented an experiment which is difficult to explain in terms of a simple integrator hypothesis [2]. Extensions of the Webster-Jeffress model [12] following McFadden [9] are suggested and the predictions are found to agree fairly well with the data.

REFERENCES

[1] D.W. Bell, 'Effect of fringe on Masking-Level Difference when gating from uncorrelated to correlated noise', J.A.S.A., Vol. 52, 525-529, (1972).

[2] K.J. Gabriel, H.S. Colburn, P.M. Zurek and N.I. Durlach, 'A general modification for models of binaural interaction', J.A.S.A., Vol. 74(S1), S 85, (1983).

[3] D.W. Grantham and F.L. Wightman, 'Detectablity of varying interaural temporal differences', J.A.S.A., Vol. 63, 511-523, (1978).

[4] D.W. Grantham and F.L. Wightman, 'Detectablity of a pulsed tone in the presence of a masker with time-varying interaural correlation', J.A.S.A., Vol. 65, 1509-1517, (1979).

[5] I.J. Hirsh, 'The influence of interaural phase in interaural summation and inhibition', J.A.S.A., Vol. 20, 536-544, (1948).

[6] L.A. Jeffress, H.C. Blodgett and B.H. Deatherage, 'The masking of tones by white noise as a function of the interaural phases of both components: I. 500 cycles', J.A.S.A., Vol. 24, 523-527, (1952).

[7] T.L. Langford and L.A. Jeffress, 'Effect of noise cross-correlation on binaural signal detection', J.A.S.A., Vol. 36, 1455-1458, (1964).

[8] J.C.R. Licklider, 'The influence of interaural phase relations upon the masking of speech by white noise', J.A.S.A., Vol. 20, 150-159, (1948).

[9] D. McFadden, 'Masking-Level differences with continuous and with burst masking noise', J.A.S.A., Vol. 40, 1414-1419, (1966).

[10] D.E. Robinson and L.A. Jeffress, 'Effect of varying the interaural noise correlation on the detectability of tonal signals', J.A.S.A., Vol. 35, 1947-1952, (1963).

[11] D.E. Robinson and C. Trahiotis, 'Effects of signal duration and masker duration on detectability under diotic and dichotic listening conditions', Percept. Psychophys. Vol. 12, 333-334, (1972).

[12] L.A. Jeffress, 'Binaural signal detection: Vector theory', in J.V. Tobias, Foundations of Modern Auditory Theory, Vol. 2, 349-368, Academic Press, New York, (1972).

[13] B. Kollmeier, 'A comparison of the temporal response of the auditory system under dichotic and diotic conditions', J.A.S.A., Vol. 74(S1), S39, (1983).