THE UNIVERSITY OF MICHIGAN

DEPARTMENT OF PSYCHOLOGY

Sensory Intelligence Laboratory

Annual Progress Report

MECHANISMS FOR IMPROVING HUMAN VISUAL AND AUDITORY SKILLS
IN AIR FORCE TASKS

Wilson P. Tanner, Jr.

(Prepared by Eli Osman)

ORA Project 02795

Recent activities in the Sensory Intelligence Laboratory have involved work by Eli Osman, Sabin Head, and Gary Sylvester. Osman has been working on a model of analysis of binaural inputs by the auditory system, and has been conducting related experiments on binaural detection of signals in noise. Head has been conducting experiments concerning the time utilized by the auditory system in processing information in a psychophysical task. Sylvester has been concerned with psychophysical learning and with theories of signal recognition. He is currently working on the development of theory and the planning of experiments for his doctoral dissertation.

Sabin Head is currently conducting an experiment concerned with psychophysical information processing and decision times. A modified auditory Yes-No detection task is being employed with an eye toward controlling in a precise way the interval of time devoted to processing information in a reaction time task. Just prior to the signal observation period the listener is given a cue that, if processed in time, could enhance his detection performance. By varying the cue's lead time, a function relating "speed" to "accuracy" can be empirically determined.

The signal embedded in noise is the occurrence of one of two possible tones (500 or 2000 cps, individually detectable at approximately a d' of 2.0), and this is to be distinguished from noise alone. Noise alone has a probability of .50 on each trial. The two signals are equally probable on signal plus noise trials. The cue occurs on all trials and its frequency (900 or 1100 cps) in all cases indicates which of the two signals will occur, if any. Performance will be evaluated at a number of lead times of from -50 msec (cue follows signal) to +100 msec (cue leads signal). At least two separate experimental sessions are being run for each such point. The same series of points is being studied for a constant-cue condition where the cues are all 1000 cps (thus providing no frequency information), to assess the extent to which the data reflect the result of temporal information also provided by the cues. In addition to these two main conditions, performance will be measured without cues and also for each of the signals individually. From pilot work we expect to find a decrease in detection performance where the cue follows the signal and an increase in detection performance peaking where the cue leads the signal by 25 msec.

Temporal cues have been found most effective at 250 msec in other types of studies, but we suspect that the figure may be spurious and a result of a listener's particular strategy of the moment. We do not expect temporal cueing to be any longer or shorter than frequency cueing in this particular experiment.

Osman's "correlation model" of binaural analysis is discussed below, using excerpts from a paper which is being prepared for publication. This paper refers to a class of experiments which can be characterized as follows. A waveform $X_1(t)$ is presented to one ear and another waveform $X_2(t)$ is presented to the other ear during the same observation interval [0,T]. Each input $X_i(t)$ may consist of noise alone, $N_i(t)$, or noise plus signal, $N_i(t) + S_i(t)$. The observer is required to decide as to which of two hypotheses is most likely. The hypotheses are

$$H_o: \quad X_i(t) = N_i(t), \quad i = 1,2$$

$$H_1: \quad X_i(t) = N_i(t) + S_i(t), \quad i = 1,2 .$$

We will assume that the noise waveforms $N_i(t)$ are generated by stationary (wide-sense), jointly ergodic, (or quasi-ergodic) Gaussian processes. For our purposes the experimental variables of greatest interest will be the correlation of the input signal waveforms, ρ_s, and the correlation of the noise inputs, ρ_N.

The results of such binaural masking experiments are commonly reported in terms of "binaural masking level differences," BMLD's. Generally the masking level difference expressed in decibels is defined by

$$\text{MLD(dB)} \quad = \quad 10 \log_{10} \left[(\frac{Es}{N_o})_r \ / \ (\frac{Es}{N_o})_T \right] ,$$

where $(Es/N_o)_r$ is the ratio of signal energy (Es) to noise power density (N_o) of a reference condition and $(Es/N_o)_T$ is the ratio of signal energy to noise power density of the specified test masking condition, and where detection performance is equated for the two conditions. For binaural MLD's a commonly used reference condition is input monaural. Often input binaural with $p_S = p_N = +1$, a homophasic condition, is used as reference, where the BMLD for this condition with reference to the monaural condition is negligible. The BMLD has been found to reach a maximum in the range of 10 to 20 dB when $p_S = +1$ and $p_N = -1$ or $p_S = -1$ and $p_N = +1$, both antiphasic conditions.

Two models have been proposed for the purpose of describing quantitatively the data from studies of BMLD's. The "phase-difference" (PD) model (Jeffress, 1965) proposes that signal detection in BMLD conditions is based on analysis of interaural phase shifts (transformed into time differences) by the auditory system, while detection in non-BMLD conditions (such as monaural input or binaural input with $p_S = p_N = \pm 1$) is based on a process of monitoring the input level at one ear in a frequency band containing the signal. The

"equalization-cancellation" (EC) model (Durlach, 1963) proposes that detection in BMLD conditions is based on an attempt to equalize and then cancel (by subtraction) the masking noise to result in an improvement in detection due to an increase in the effective signal to noise ratio, while detection in non-BLMD conditions is again based on a process of monitoring the input level at one ear in a frequency band containing the signal. In non-BMLD conditions such as $p_S = p_N = \pm 1$, the binaural processing proposed by either model could only result in degraded or chance performance, and so monaural processing must be employed. It is implicit in both models that the receiver must "know" the interaural relations of both signal and noise in order first to decide whether binaural or monaural processing is appropriate, and then to execute the binaural process if called for. Green and Henning (1969) have briefly reviewed these two models of binaural analysis.

The notion that one of two mechanisms must be employed depending on whether or not a BMLD condition is the case may be considered an unlikely hypothesis since even with experimental inputs favoring the non-BMLD mechanism, any degradation of the interaural noise correlation would favor use of the BMLD mechanism, although the benefit may be too small to be detected by experiment. The BMLD is a continuous function of changes in the interaural noise correlation from -1 to +1 (Robinson and Jeffress, 1963; see Figure 4 of this report), and the purely homophasic condition favoring the non-BMLD mechanism may never be attainable for any human observer. There is also other evidence suggesting that the nature of the information processing performed is not dependent on the receiver knowing the interaural relations of both signal and noise. The magnitude of the BMLD may be unaffected by observer uncertainty concerning interaural relations for the signal (observations by Osman) or for the noise (McFadden, 1967). The detection of a signal under the homophasic non-BMLD condition is phenomenologically no different, as judged by verbal report, from the detection of a signal under the antiphasic conditions yielding the largest BMLD's (McFadden, 1967). Furthermore, binaural beats resulting from low-frequency sinusoids can be strikingly enhanced by adding in-phase noise binaurally, which effectively results in interspersing homophasic and antiphasic binaural input configurations in continuous fashion (Egan, 1965).

We propose an alternative model of BMLD's, which will be referred to as the "correlation" (ρ) model, which is simpler than the PD or EC models in the sense that the same form of processing of input waveforms over the observation interval [0,T] is assumed, regardless of the nature of the input configurations, that is whether or not a BMLD condition is the case. In addition, the receiver need never know the interaural relations of the signal, and can estimate needed weights based on the interaural relations of the noise after the input waveforms are processed. The functional model we propose here provides a description of a large mass of results, simply organized with regard to interaural correlation. Furthermore, the processing reduces to that of a simple energy detector for monaural detection.

A discussion of the monaural energy detector model for detection of sinusoidal signals in noise and relevant empirical results may be found in Green and Swets (1966). The energy detector is the optimal information processing model for a receiver listening for a signal in a noise background, where the signal itself is a sample of noise. For a two channel receiver the optimal detector listening for noise signals jointly presented over both channels, again in a noise background, with varying interchannel correlations for signal and for noise, can be shown to be equivalent to one which operates by computing a linear combination of three quantities, the energy levels at each channel and the interchannel crosscorrelation. This is the nature of the decision variable we will consider for the ρ-model. It happens also to be the nature of the decision variable of the EC model operating in the binaural mode. Major differences lie in the weighting functions and the presummed character of errors inherent in the receiver.

Specifically, for the ρ-model, the processing consists of the computation of a decision variable or test statistic, D, where

$$D = A \int_0^T x_1^2(t)dt + B \int_0^T x_2^2(t)dt + C \int_0^T x_1(t) \, x_2(t) \, dt \, ,$$

and where each $x_i(t)$ denotes the result of adding internal noise to the input and filtering. Appropriate choice of the coefficients A, B, C, leads to the following formula for BMLD's when the signal is binaural and the signal and noise levels are both equated for the two ears:

$$BMLD \ (dB) = 10 \log [2(1-\rho_s\rho_n) \, / \, (1-\rho_n^2)] + constant,$$

where ρ_n is the effective noise correlation after the addition of internal noise. (Since some sort of internal noise is a necessary assumption for each of the models proposed, basic differences between all three models are in terms of the assumptions made concerning internal noise.) If the signal is monaural and the noise binaural, again equated at the ears, then the result is the following:

$$BMLD \ (dB) = 10 \log [1/(1 - \rho_n^2)] + constant.$$

The expression for D in these cases is

$$D = [\int_0^T x_1^2 (t)dt + \int_0^T x_2^2 (t)dt - 2\rho_n \int_0^T x_1 (t) \, x_2 (t)dt]/(1-\rho_n^2).$$

We will want the probability density function and mean and variance of D conditional on each of H_O and H_1 in accordance with the notion that the receiver functions as a statistical hypotheses tester. Note that we will always have $-1 < \rho_n < +1$, perfect interaural correlations of effective noise being unattainable. The receiver operation of the ρ-model looks very much like that of the binaural mode of the EC model, and indeed the two are identical for the unrealizeable case of an error-free receiver. However, in the EC model the process is of the form

$$\int_0^T [ax_1(t) - bx_2(t)]^2 \, dt = a^2 \int_0^T x_1^2(t)dt + b^2 \int_0^T x_2^2(t) \, dt - 2ab \int_0^T x_1(t)x_2(t)dt$$

and it is not possible to have $a^2 = b^2 = C$ with $ab = C\rho_n$, except for the special cases of $\rho_n = \pm 1$.

The next steps concern the mathematical development of the ρ-model, including a precise specification of the additive internal noise which is assumed, followed by an extensive evaluation of the model. Evaluation involves comparing theoretical results derived from the ρ-model with empirical results from a variety of experiments. A sample of the predictive power of the model is provided in Figures 1, 2, 3, and 4. Explanations are given in the figure captions. The model requires that only one free parameter be evaluated to generate each curve in Figures 3 and 4, and for Figures 1 and 2, the two free parameters of the model were evaluated using the two conditions in Figure 1 and these results were used to predict the empirical results for four other conditions as shown in Figure 2. The results are quite satisfactory, and the agreement between theory and data is at least as good for the ρ-model as it is for the EC model, judging from published reports.

Current experiments for which data is being collected and analyzed with regard to BMLD's concern observer uncertainty with respect to interaural parameters and the effects of varying temporal positions of signal relative to noise. As described above, BMLD's have been studied extensively and two other models, the equalization-cancellation (EC) model and the phase difference (PD) model, have been published which describe the processing of auditory inputs that might explain the data in this field. A crucial point made here is that both theories suppose that the way in which the inputs are processed varies with the interaural conditions. For each model, the detector must know and use the interaural relations of both signal and noise in order to process input information appropriately and achieve maximum detectibility in MLD and non-MLD conditions.

In MLD conditions a binaural mechanism is used; the decision is based on an energy-like quantity at the output of the EC mechanism or a timing signal derived from interaural phase differences. In non-MLD conditions both models regard detection as based on a monaural mechanism; monitoring some amplitude measure at the output of a filter at one ear.

In order to test the notion that detection under MLD conditions is based on a process quite different from that under non-MLD conditions, McFadden (1967) measured detectibility with and without uncertainty in a Yes-No task for a homophasic condition, NOSO, and an antiphasic condition, NπSO. Noise alone or signal plus noise was gated for the duration of a 125 msec observation interval. The signals at the two ears were always in phase; uncertainty was only with respect to the interaural noise phase. Signal levels were adjusted to achieve approximately equal levels of detectability for the two conditions. It was stated that the listener's operated with a "medium criterion" so that the results, reported as percent correct and calculated by averaging obtained hit and correct rejection rates together and across subjects, was a satisfactory measure of performance. For these results uncertainty had a negligible effect on the MLD's. The conclusion drawn was that either the inputs are always routinely processed by both the binaural and monaural mechanisms, or the early portion of each observation interval was used to estimate the noise correlation and then only the appropriate processing system was attended to for the duration of the interval. McFadden admitted to a bias towards the first interpretation while noting that there was no good basis for rejecting the second.

The second interpretation is not unreasonable since the BMLD has been shown to be nondecreasing as signal duration is reduced from 100 to 10 msec with continuous noise (Green, 1966), and if the observer is allowed to listen to noise alone for a short period on every trial under conditions of no uncertainty, the size of the MLD is increased (McFadden, 1966). Monitoring noise correlation for a portion of the observation interval may be a routine operation with or without uncertainty.

If it is the case that some sort of estimation of the interaural relations for both signal and noise is required on every trial, and that only one of two processing systems, binaural or monaural, is subsequently employed, then a decrease in the BMLD should be expected if the observer has uncertainty as to which of the two systems to use and cannot reduce this uncertainty by observation of the inputs. This would be the case if the uncertainty was with respect to the interaural correlation of the signal. Thus one study was intended to examine the effects of uncertainty with regard to the homophasic condition NOSO and the antiphasic condition NOSπ.

Qualitative analysis of results for several experiments on several subjects involves a simple categorization in accordance with whether the differences between performance observed under conditions of no uncertainty and under conditions of uncertainty for NOSO and NOSπ can be reasonably explained by a change in decision criterion or whether a change in detectability is implied. Three effects were observed: no change in performance and changes which could be attributed to modification of decision criterion (both results conforming to expectations based on single process models), and changes which appeared to be the result of a change in detectibility for a certain percentage of subjects (conforming to expectations based on multiple process models). Analysis of results will proceed with reference to ROC (receiver operating characteristic) curves.

The PD model, the EC model, and the ρ-model disagree in their choice of the variable being attended to under MLD and non-MLD conditions. For the EC and ρ models the variables are all energy-like quantities, although different in each case. The vector or PD model receiver listens to interaural phase differences under MLD conditions, but not under non-MLD conditions. The differential role of transients due to the addition of signal to noise under different interaural conditions is not clearly specified by any one of these models. An examination of BMLD's for different temporal relations of signal and noise for both NOSO and NOSπ was carried out to answer and possibly raise questions related to the differences between current models of BMLD's and the inadequacies of each. Data were collected in a two-interval forced choice (2IFC) task and in a Yes-No (Y-N) task using a signal of 400 cps. The results for one observer are shown in Figure 5. They show several significant effects:

(1) A 10 msec signal presented during a 125 msec noise burst is more easily detected as it is placed closer to the termination of the noise burst. This is true for both NOSO and NOSπ. The BMLD is little affected by these temporal changes.

(2) For NOSO a 10 msec signal added to a 10 msec noise burst is as detectable as the same signal added to a 125 msec noise burst provided that signal and noise terminate together. For NOSπ detection is improved in the case where there is more noise. The BMLD is therefore larger when there is more noise (before or after the signal).

(3) The BMLD for a 125 msec signal added coincidentally to 125 msec noise lies somewhere between the two results indicated in (2) above. The Y-N experiment is being conducted in an attempt to clarify the size of the BMLD differences for three such conditions, since the result may be of significance for interpreting the role of the noise fringe (which plays a different theoretical role in each model).

(4) Small BMLD's exist for signals ending before the noise burst begins, and relatively large BMLD's exist for signals beginning after the noise burst ends. These results are most difficult to explain with the PD model.

(5) The shape or slope of the psychometric function does not appear to differ for NOSO as compared to NOSπ under the same condition.

(6) There may be a change in slope of the psychometric function as signal and noise vary in relative positions. The function may be shallower when the signal ends before the noise begins.

The experiment summarized in Figure 5 was carried out with three observers and the points listed above are in general agreement with the results of all three, but with some qualifications. The results will be analyzed further and extended observations may be made before any final interpretation is reached.

Figures 1 and 2. BMLD's expressed in dB referred to the homophasic condition
NOSO (interaural noise phase shift of 0° and interaural signal phase shift
of 0°) as a function of signal frequency. The notation NOSπ, NπSO, NOSm,
NπSm, NπSπ, and NmSm is interpreted as follows: N refers to noise, S re-
fers to signal, the second and fourth symbols give the conditions for N and
S, respectively, with 0,π, and m referring to interaural phase shift of 0°,
interaural phase shift of 180°, and monaural input, respectively. Two
parameters reflecting the magnitude and interaural correlation of internal
noise in the ρ-model were estimated for each signal frequency for NOSπ and
NπSO in Fig. 1, and the predictions derived from these are plotted (o-open
circles) in Fig. 2 along with the corresponding empirical results ($\bullet$-filled
circles) for NOSm, NπSm, NπSπ, and NmSm. (The two open circles for NmSm
at each frequency define the endpoints of a range within which the theoreti-
cal value must fall.) Empirical results from Hirsh and Burgeat (1958).

Figure 3. BMLD's in dB referred to NOSO for fixed interaural time delays for
noise, τ_N, as a function of interaural signal phase shift, φ_S, for a signal
of 500 cps (τ_S denotes interaural time delay for signal). A single param-
eter reflecting the "effective" noise correlation defines all the theoreti-
cal curves simultaneously. Only the curves at τ_N = 0 and -1.0 msec are
drawn, the others would also be in good agreement with the data. Empirical
results from Jeffress, Blodgett, and Deatherage (1952).

Figure 4. BMLD's in dB referred to NOSO for homophasic (SO) and antiphasic
(Sπ) signals in noise of varying interaural statistical correlation, ρ_N,
as a function of $|\rho_N|$. For SO noise correlation is negative (o-open
circles) and for Sπ it is positive (x-crosses) to the right of 0. The
opposite holds to the left of 0. Signal frequency is 500 cps. Empirical
results from Robinson and Jeffress (1963).

Figure 5. Percent correct in a 2IFC task as a function of signal level for
NOSO and NOSπ. Signal frequency is 400 cps. Noise level is fixed.
Curves are for different temporal relations of signal (S) and noise (N)
as indicated on upper right corner of figure, where numbers indicate
duration of 10 msec or 125 msec, and S and N either start together, have
coincident middurations, or end together, or S begins 2 msec after N
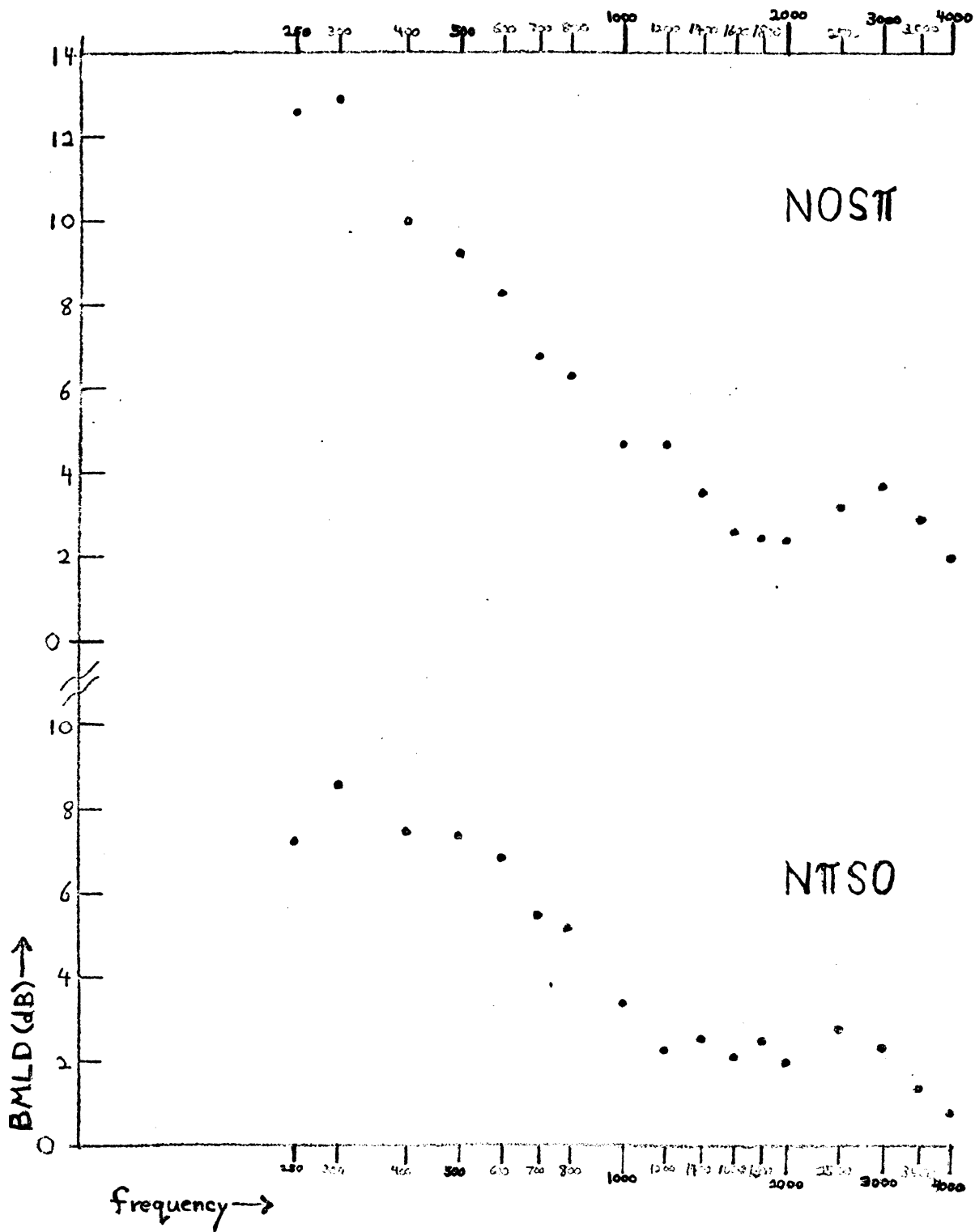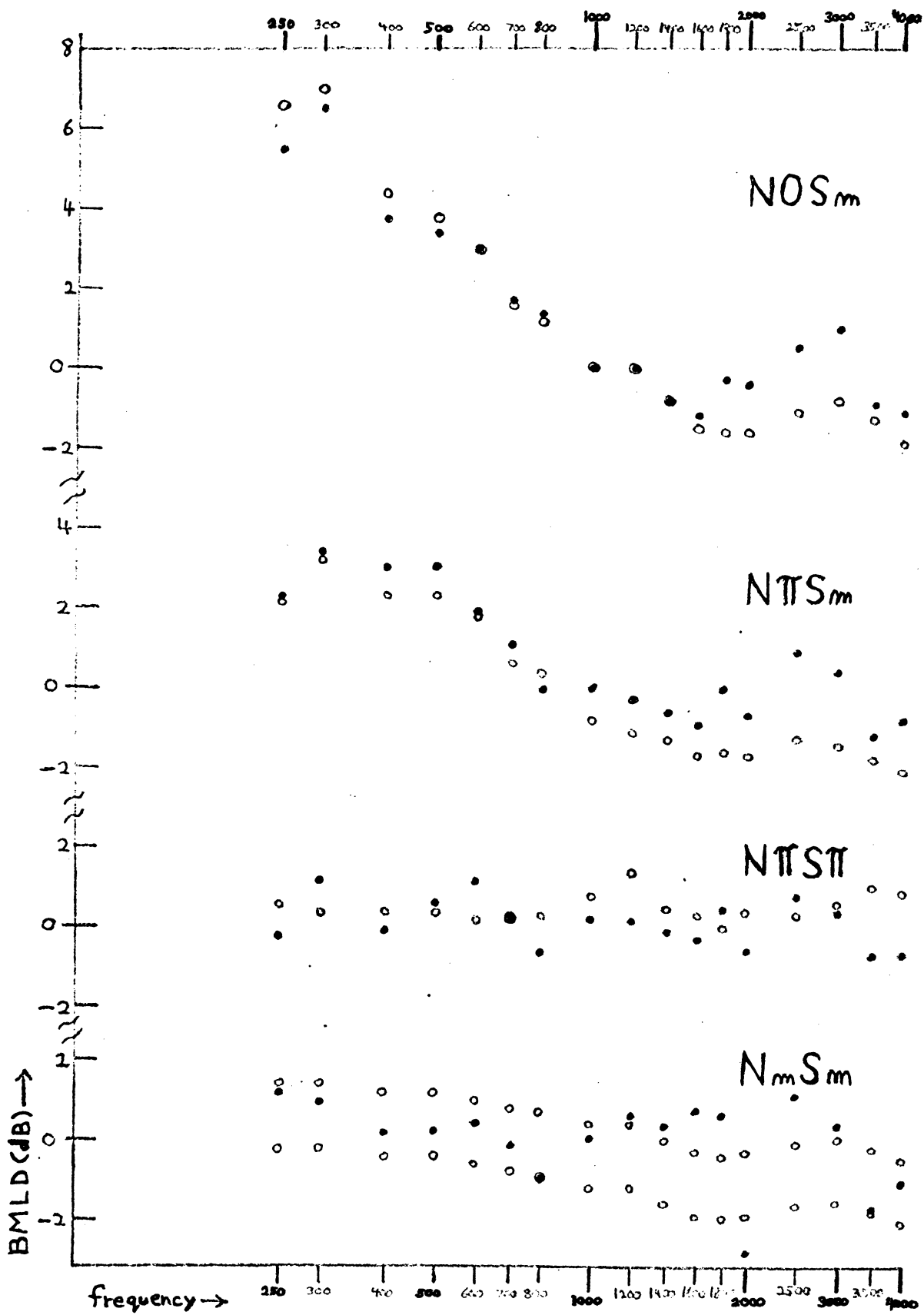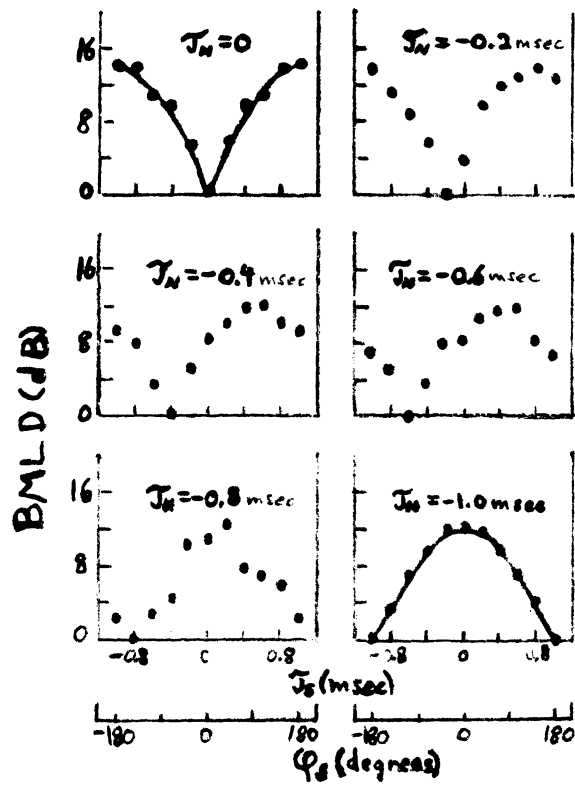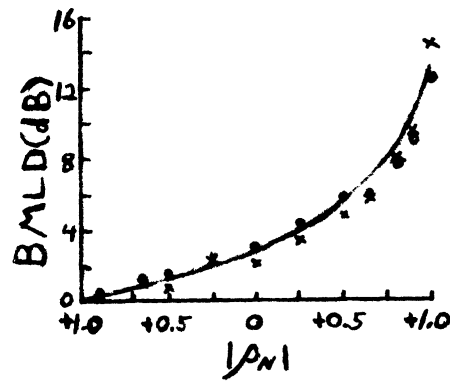ends, or S ends 2 msec before N begins.

Figure 1

9

Figure 2

Figure 3

Figure 4

Tom, units 1970 (plat 4-1-70)

		NOSC. NOST	
S–10, N–125	at tag	● ⊖	
S–10, N–125	mid con	▲ △	
S–10, N–125	end tog	▶ ▷	
S–125, N–125	at tag	◆ ◇	
S–10, N–10	at tog	■ □	
S–10, N–125	2 mass off	✕ ⊗	
S–10, N–125	2 mass bef	＋ ⊕	

[multiple pts. displaced to right on left]
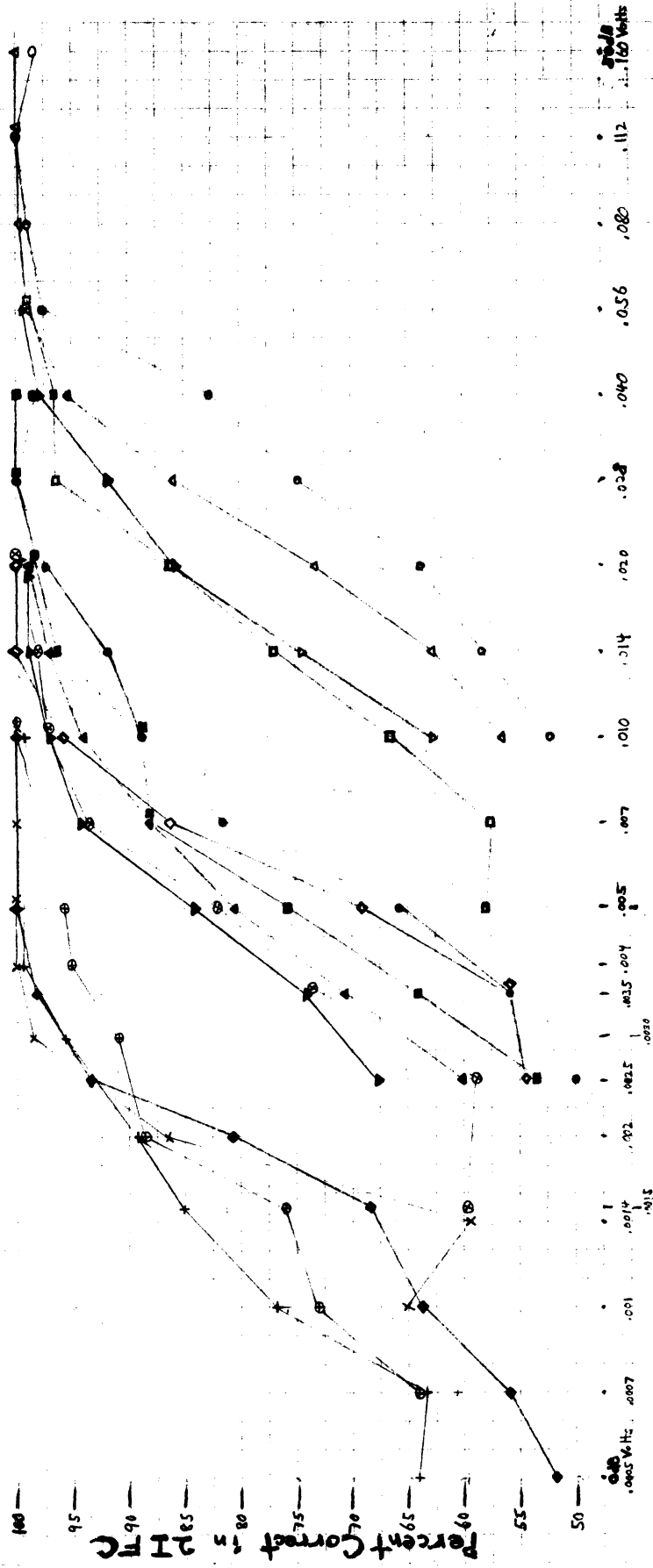
Signal Level
(Scale is linear in dB)

Percent Correct in 2 IFC

Figure 5

REFERENCES

1. Durlach, N. I.   Equalization and cancellation theory of binaural masking level differences,   J. Acoust. Soc. Am. , 35, 1206-1213 (1963).

2. Egan, J. P.   Demonstration of masking level differences by binaural beats, J. Acoust. Soc. Am. , 37, 1143-1144 (1965).

3. Green, D. M.   Interaural phase effects in the masking of signals of different durations, J. Acoust. Soc. Am. , 39, 720-724 (1966).

4. Green, D. M. , and Henning, G. B.   Audition, Ann. Rev. Psychol. , 20, 105-128, (1969).

5. Green, D. M. , and Swets, J. A.   Signal Detection Theory and Psychophysics, John Wiley & Sons, Inc. , New York (1966).

6. Hirsh, I. J. , and Burgeat, M.   Binaural effects in remote masking.   J. Acoust. Soc. Am. , 30, 827-832 (1958).

7. Jeffress, L. A.   Binaural signal detection: vector theory, Defense Res. Lab. Acoust. Rept. No. 245, Univ. Tex. (1965).

8. Jeffress, L. A. , Blodgett, H. C. , and Deatherage, B. H.   Masking of tones by white noise as a function of the interaural phase of both components. I. 500 cycles.   J. Acoust. Soc. Am. , 24, 523-527 (1952).

9. McFadden, D.   Detection of an in-phase signal with and without uncertainty regarding the interaural phase of the masking noise, J. Acoust. Soc. Am. , 41, 778-781 (1967).

10. McFadden, D.   Masking-level differences with continuous and with burst masking noise, J. Acoust. Soc. Am. , 40, 1414-1419 (1966).

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