NONCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A244 392

DTIC ELECTE JAN 15 1992 S D D

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

(2)

1a. REPORT SECURITY CLASSIFICATION

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 91 1022

6a. NAME OF PERFORMING ORGANIZATION
Central Institute for the Deaf

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research

6c. ADDRESS (City, State, and ZIP Code)
818 South Euclid Avenue
Saint Louis, Missouri 63110

7b. ADDRESS (City, State, and ZIP Code)
Bolling Air Force Base
District of Columbia 20332

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL (If applicable)
NL

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-890302

8c. ADDRESS (City, State, and ZIP Code)
Bolling Air Force Base
District of Columbia 20332

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2313	A6	

11. TITLE (Include Security Classification)
Binaural Masking: An analysis of models

12. PERSONAL AUTHOR(S)
Robert H. Gilkey

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final Technical	FROM 4/1/89 TO 1/31/91	91-11-21	8

16. SUPPLEMENTARY NOTATION

92-01105

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The goal of this program of research has been to specify auditory processing in the presence of noisy backgrounds. A variety of experimental and modeling approaches have been employed to examine this processing. Overall the results suggest the importance of spectral and temporal comparisons in signal detection and suggest that similar processing underlies monaural and binaural detection. The introduction of masker energy in temporal intervals that did not overlap with the signal could be shown to either enhance or degrade detection performance, depending on the interaural parameters of the stimuli. Experiments on remote masking and suppression showed excitatory and inhibitory effects that extended across more than an octave. These results are being used to evaluate a nonlinear model of cochlear processing. The responses of subjects to individual stimuli (reproducible noise samples) were highly correlated between monaural and binaural conditions that had seemed dramatically different when the ensemble performance was considered (i.e., data averaged across noise samples). These results

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
John F. Tangney

22b. TELEPHONE (Include Area Code)
(202) 767-5021

22c. OFFICE SYMBOL
NL

DD Form 1473, JUN 86 Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
NONCLASSIFIED

25 NOV 1991

92 1 13 023

Abstract (cont.)

suggests that similar processing occurs in the monaural and the binaural systems, and indicates the need for models that reflect this similarity. Overall, the work examined issues and models of contemporary interest and thus has implications for auditory theory in general and for the study of auditory pattern analysis and auditory masking in specific.

Final Report

AFOSR 89-0302

April 1, 1989 to January 31, 1991

I. SUMMARY

The goal of this project has been to specify the transformations used by the auditory system in order to determine the presence of the signal in an auditory masking task, with particular emphasis on the role of processes that compare information in the frequency domain and in the time domain. Studies of binaural masking show that masking noise that does not overlap with the signal in time can either improve or degrade the detectability of the signal, depending on the interaural phase relations among the masker, the fringe, and the signal. The results from studies of monaural pure tone masking and suppression are being used to evaluate a non-linear model of cochlear processing that may play a significant role in the spectral comparison process. Studies that examine the responses of subjects to each individual noise alone and signal-plus-noise stimulus (Molecular Psychophysics) suggest that similar cues determine performance in monaural and binaural masking tasks, a result not predicted by many models of binaural processing.

II. RESEARCH OBJECTIVES

The overall goal of this program of research is to specify the processes used by the auditory system to detect signals presented in noisy backgrounds. It is assumed that the behavior of the subject can be modeled by a system that on each trial computes a single quantity, the "decision variable" of the model, which in the manner described by the Theory of Signal Detectability provides the basis of the subject's decision about the presence or absence of the signal. Within this framework our task is to determine the decision variable of the subject. For the tone-in-noise detection task we have been investigating, classical models argue that the decision variable is based on processing within the narrow frequency band centered around the signal (i.e., the critical band) and within the brief temporal window that contains the signal. We have used a variety of approaches to demonstrate that these classical models are oversimplifications, to develop models that provide a more accurate description of the responses of the subject, and to delineate the relation between the mechanisms underlying monaural and binaural masking.

III. STATUS OF THE RESEARCH

Additional support for this research has been provided by a grant, NIH (DC-00786) "Monaural masking, binaural masking and their interrelations," period of support May 1, 1990 through April 30, 1994, R.H. Gilkey, PI.

<u>**Molar psychophysical analysis of models of masking**</u>

<u>Binaural temporal masking</u>. Because of the binaural Masking Level Difference (MLD), if the interaural phase of a noise masker is switched during the observation interval from in phase (N0) to 180° out of phase ($N\pi$) or from $N\pi$ to N0, a brief interaurally out-of-phase signal ($S\pi$) will be about 15 dB more

detectable in the N0 portion of the noise than in the Nπ portion. By investigating the change in detectability as a function of the delay (Δt) between the onset of the signal and the phase transition in the noise, the temporal response of the binaural system can be evaluated. The results of this case can be contrasted with a set of conditions in which the interaural phase of the noise is held constant (Nπ), but the level of the noise is reduced or increased by 15 dB halfway through the observation interval. Within a model such as the EC model (N.I. Durlach, in J.V. Tobias (Ed.) Foundations of Modern Auditory Theory II, 371-462, 1972), the first case produces a change of level only in the binaural channel. The second case produces a change in the level in the monaural channel as well. The curves that describe the relation between threshold and Δt can be thought of as temporal masking functions. They show, like traditional temporal masking data, that the decay of backward masking (cases where the N0 segment of the noise precedes an Nπ segment or where the lower intensity segment of the noise precedes the higher intensity segment) is more rapid than for forward masking. Double-sided exponential integration windows have been fit to the forward and backward masking functions. The equivalent rectangular duration of the best-fitting window under monaural conditions ranges from 12-26 ms, somewhat larger than those estimated by Moore et al. [J. Acoust. Soc. Am. 83: 1102-1116, 1988]. The equivalent rectangular duration for the binaural conditions ranges from 41-83 ms, similar to estimates by Grantham and Wightman [J. Acoust. Soc. Am. 65: 1509-1517, 1979]. The observed differences between monaural and binaural conditions were taken as additional evidence that the binaural system responds sluggishly to changing stimulation [Grantham and Wightman, J. Acoust. Soc. Am. 63: 511-523, 1978]. A published paper is included in the appendix [Kollmeier and Gilkey, J. Acoust. Soc. Am. 87, 1709-1719, 1990].

In studying the effects of a forward masker fringe, Yost [J. Acoust. Soc. Am. 78: 901- 907, 1985] found that the threshold for a brief Sπ signal masked by a brief N0 masking noise was not changed when an Nπ forward masker fringe was added. This result was somewhat surprising in light of results such as those of McFadden [J. Acoust. Soc. Am. 40: 1414-1419, 1966] who showed that an N0 forward fringe substantially improved performance in an N0Sπ detection task, and concluded that the system uses the forward fringe as a diotic reference against which to detect the dichotic signal. If an N0 forward fringe provides a useful reference, it might be expected that an Nπ forward fringe would provide a detrimental reference. Yost's results also seemed to conflict with the interpretations of Kollmeier and Gilkey [op. cit.], who thought of the Nπ fringe as a forward masker. One possibility was that the function that relates threshold to Δt for the Nπ forward fringe condition intersects with the function that relates threshold to Δt for the pulsed masker condition at Δt = 0, even though the functions are different elsewhere. To resolve these questions, the detectability of an Sπ tonal signal was investigated as a function of Δt, in the presence of an N0 "masker" that was preceded by quiet or by an Nπ "forward fringe," and followed by quiet or by an N0 or Nπ "backward fringe." The results show that the functions for the Nπ forward fringe condition and the pulsed masker condition are indeed different and that they do not intersect. Overall, the results failed to replicate those of Yost, showing instead that the presence of either an Nπ forward fringe or an Nπ backward fringe reduced detectability for all subjects under a variety of conditions. The results are a further indication that the auditory system uses information that does not overlap with the signal in the temporal domain. Subsequent measurements indicate that the difference between Yost's results and ours cannot be explained based on differences in psychophysical procedure, the amount or type of training received by the subjects, or the duration of the signal. A published paper is included in the appendix [R.H. Gilkey, B.D. Simpson, and J.M. Weisenberger, J. Acoust. Soc. Am. 88, 1323-1332, 1990].

McFadden [J. Acoust. Soc. Am. 83: 1685-1687, 1988] investigated the detectability of a brief tonal signal in the presence of a long duration wideband masking noise. While the "overshoot" effect [E. Zwicker, J. Acoust. Soc. Am. 37: 653-663, 1965] was observed for diotic stimuli (N0S0), no overshoot was observed with dichotic stimuli (N0Sπ). (Overshoot is defined as the difference between the threshold for a signal whose onset is near the beginning of the masker, and the threshold for a signal whose offset is near the end of the masker). Comparable data [D.E. Robinson and C. Trahiotis, Percept. Psychophys. 12: 333-334, 1972; C. Trahiotis, T.R. Dolan, and T.H. Miller, Percept. Psychophys. 12: 335-338, 1972] indicate no overshoot under the monaural condition, but about 6 dB of overshoot under the binaural condition. In a recent experiment we found 4-8 dB of overshoot under both N0S0 and N0Sπ condition when a wideband masker was used, but no overshoot under either condition when a narrowband masker was used. McFadden had used a 750 Hz signal frequency, whereas most of these studies, including ours, used a 500 Hz signal frequency. To eliminate the unlikely possibility that this small difference in frequency could have produced the observed discrepancy, we conducted additional measurements using both 500 Hz and 750 Hz signals and wideband maskers. The results showed no difference in overshoot for the two signal frequencies, but also showed very little overshoot for most subjects under either the N0S0 or the N0Sπ condition. The total pattern of results for our two experiments indicates that there is considerable between subject variability in the observed overshoot.

Psychophysical evaluation of a physiologically based model of auditory processing. In the classical literature both the masking and the suppression of one tone by a second tone of lower frequency have been shown to be nonlinear functions of overall level. We have partially replicated the experiments of Wegel and Lane [Physiol. Rev. 23: 226-285, 1924] on remote masking and of Duifhuis [J. Acoust. Soc. Am. 67: 914-927, 1980] on suppression, using modern adaptive psychophysical techniques and the same subjects in both experiments. The data are comparable to those from earlier studies and agree with the Multiple Band Pass Nonlinearity (MBPNL) model (Goldstein, 1989, 1990, op. cit.). This model is based on current knowledge of auditory physiology and describes the response of the peripheral auditory system at each frequency as the result of a nonlinear interaction between a linear lowpass ("tail") filter and a compressive bandpass ("tip") filter. This view suggests that both excitatory and suppressive mechanisms influence remote masking. The data indicated that the relative influence of these two mechanisms varies from subject to subject. Estimates of the exponent of the compressive nonlinearity of the model obtained from the simultaneous masking experiment agree with those obtained from the suppression and forward masking experiments. Reexamination of the simultaneous masking data that Gagné [J. Acoust. Soc. Am. 83: 2311-2321, 1988] obtained with hearing-impaired subjects, indicates that they are also compatible with the MBPNL model, if it is assumed that the tip filter is damaged (gain set to zero). The results of these experiments were presented to the Acoustical Society of America [Goldstein, Gilkey, and Quiñónez, J. Acoust. Soc. Am., 86, S24(A), 1989].

Molecular psychophysical analyses of models of masking

In most studies of auditory masking, including those described above, both the stimulus and the performance of the subjects are described by their statistical properties (e.g., the average power of the stimulus and the average probability of a correct response). The outputs of models are described by their distributional properties and the average performance of a model is fit to the average performance of a subject. Another approach was described by Green (Psychol. Rev. 71: 392-407, 1964) and referred to as "molecular" psychophysics. In this approach, reproducible noise is used as a masker, such that the stimulus can be specified exactly on every trial. Similarly, the responses of the subject are considered on a trial-by-trial basis. The outputs of models are determined for each stimulus and the fit of the model is evaluated by comparing these outputs to the associated responses of the subjects.

The relation between monaural and binaural masking. The large masking level difference observed between monaural and binaural tone-in-noise masking tasks has been used to suggest that quite different processing is employed under the two conditions (e.g., energy detection vs. interaural time processing). However, when Gilkey et al. [J. Acoust. Soc. Am.78: 1207-1219, 1985] examined the responses of subjects to individual wideband reproducible noise samples, they found that the responses under the N0S0 and N0Sπ conditions were highly correlated. On the other hand, when Isabelle and Colburn [J. Acoust. Soc. Am. 82: 109(A), 1987] examined the responses to individual narrowband reproducible noise samples, they found correlations that were much weaker and often negative. They attributed the differences between their data and those of Gilkey et al. to the differences in the bandwidth of the masker. If so, this would suggest that the correlation observed by Gilkey et al. would more appropriately be attributed to similarities in across critical band processing rather than to similarities in within critical band processing, as Gilkey et al. had implied.

To investigate further the effect of masker bandwidth, the experiment of Gilkey et al. was replicated using both wideband (100-3000 Hz) and narrowband (third octave) maskers. Although the correlation between N0S0 and N0Sπ performance was, in general, somewhat weaker for the narrowband condition all observed correlations were significant ($p << .001$), reaffirming the strong correlation between N0S0 and N0Sπ performance. Again, this result has significant implications for models of both monaural and binaural performance. It is typically assumed that monaural performance is governed by the energy in the stimulus, while binaural performance is related to interaural differences in the stimulus, particularly interaural differences in time. The results of this experiment imply that binaural performance might be based on an energy-like cue (e.g., the E-C model), or that monaural performance might be based on a timing-like cue (e.g., the model of Bilsen and Goldstein [J. Acoust. Soc. Am. 55: 292-296, 1974]).

More recent studies extend this approach to conditions employing interaurally uncorrelated noise. The N0S0 and N0Sπ conditions show very different performance at the molar level, but very similar performance at the molecular level. We have shown that when the interaural correlation of the external noise is 1.0, the output of the E-C mechanism is highly correlated with the waveforms in the monaural channels. Said differently, the effective maskers under the N0S0 and N0Sπ conditions are highly correlated. Therefore, the observed correlation between the N0S0 and the N0Sπ molecular responses of human subjects is expected. On the other hand, it might be expected that even though the N0S0 and NUS0 conditions yield similar performance on the molar level, quite different performance would be seen when the data are analyzed on the molecular level. Under the NUS0 condition the interaural correlation of the masker is zero. The E-C model would argue that the effective masker under these conditions is only partially correlated with either of the maskers in the monaural channels. We collected molecular psychophysical data under the N0S0, N0Sπ, NUS0, and NUSπ conditions. Under the N0S0 condition performance was measured separately for the waveforms reaching each ear under the binaural conditions. The strong correlation between N0S0 and N0Sπ molecular responses was replicated. As expected, the responses under either the NUS0 or the NUSπ condition were not well predicted by the responses under the N0S0 condition to either monaural masker (i.e., the maskers in each ear under the NU conditions). However, to our surprise, when the N0S0 responses to the maskers in the two ears were averaged and used to predict the responses under the NU conditions the correlations were again quite strong. These results will be compared to the detailed predictions of the E-C model and to the predictions of lateralization models.

Improved molecular psychophysical methods. Although the molecular psychophysical approach has proven extremely useful, data collection is slow. The amount of data needed is increased, over that obtained in a molar experiment, by a factor of approximately N, where N is the number of reproducible noise samples employed, because the approach requires us to estimate the value of the subjects' decision variable in response to each noise sample. In the past we have inferred this value based on binary responses in a simple yes/no detection task. The use of binary responses places a practical limit on the amount of information that can be transmitted on each trial (W. R. Garner and H. W. Hake, Psychol. Rev. 58, 446-459, 1951). In addition, the approach requires a number of additional assumptions that have not been tested. Finally, the variability and the expected value of the estimate of the decision variable are not independent.

With these problems in mind, we have been developing a continuous rating procedure based on the procedure of Watson, Rilling and Bourbon [J. Acoust. Soc. Am. 36, 283-288, 1964]. The procedure is straightforward; the subjects' task is to use a mouse to position a cursor along the bottom of a CRT screen to indicate his confidence that a signal was presented on a particular trial. Positioning the cursor to the right of the screen indicates confidence that the signal was present, while positioning the cursor to the left indicates confidence that the signal was not presented. Two simple experiments were conducted to evaluate the procedure. In the first, the subjects' task was to determine whether a particular three digit number presented on the CRT screen was drawn from a population of "noise alone" numbers or from a population of "signal plus noise" numbers. In the second, the subjects' task is to detect the presence of a tonal signal in the presence of a broadband noise masker. Preliminary analyses of the visual data indicate that the subjects can, with training, produce reliable rating judgments. The expected value of these judgments appears to be a simple function of the stimulus magnitude. Ratings obtained for repeated presentations of the same stimulus are consistent. Thus, subjects' ratings can be used to reproduce stimulus distributions (i.e., by deriving receiver operating characteristics or frequency histograms). Molecular psychophysical data collected with this technique will be compared to molecular psychophysical data obtained using binary responses.

IV. PUBLICATION ACTIVITY

Publications

Kollmeier, B., and Gilkey, R.H. (1990). "Binaural temporal masking: Evidence for sluggishness in binaural detection," J. Acoust. Soc. Am. 87, 1709-1719.

Gilkey, R.H., Simpson, B.D., and Weisenberger, J.M. (1990). "Effects of masker fringe on binaural detection," J. Acoust. Soc. Am. 88, 1323-1332.

Papers in preparation

Gilkey, R.H., and Meyer, T.A. "Modeling subject responses in a reproducible noise masking task"

Planned papers

Gilkey, R.H. "The relation between monaural and binaural masking"

Gilkey, R.H. "Effects of manipulating the spectral shape of reproducible noise samples"

Goldstein, J.L., Quiñónez, R.E., and Gilkey, R.H. "A psychophysical evaluation of a model for suppression and excitation in remote masking"

V. PARTICIPATING PROFESSIONALS

Robert H. Gilkey

De Anza College, Cupertino, CA			
University of California, Berkeley, CA	B.A.	1976	Psychology
Indiana University, Bloomington, IN	Ph.D.	1981	Psychology

Dissertation title: "Molecular psychophysics and models of auditory signal detectability."

VI. INTERACTIONS

Invited papers and conference talks

Goldstein, J.L., Gilkey, R.H., and Quiñónez, R.E. (1989). "A psychophysical evaluation of a model for suppression and excitation in remote masking," J. Acoust. Soc. Am. 86, S24(A).

Gilkey, R.H. (1990). "The relation between monaural and binaural tone-in-noise masking," Association for Research in Otolaryngology Midwinter Meeting, St. Petersburg Beach, FL

Binaural forward and backward masking: Evidence for sluggishness in binaural detection

Birger Kollmeier
Drittes Physikalisches Institut, Universität Göttingen, Bürgerstr. 42-44, D-3400 Göttingen, Federal Republic of Germany

Robert H. Gilkey
Central Institute for the Deaf, 818 S. Euclid, St. Louis, Missouri 63110

(Received 17 April 1989; accepted for publication 6 December 1989)

The threshold of a short interaurally phase-inverted probe tone (20 ms, 500 Hz, S_π) was obtained in the presence of a 750-ms noise masker that was switched after 375 ms from interaurally phase-inverted (N_π) to interaurally in-phase (N_0). As the delay between probe-tone offset and noise phase transition is increased, the threshold decays from the $N_\pi S_\pi$ threshold (masking level difference = 0 dB) to the $N_0 S_\pi$ threshold (masking level difference = 15 dB). The decay in this "binaural" situation is substantially slower than in a comparable "monaural" situation, where the interaural phase of the masker is held constant (N_π), but the level of the masker is reduced by 15 dB. The prolonged decay provides evidence for additional binaural sluggishness associated with "binaural forward masking." In a second experiment, "binaural backward masking" is studied by time reversing the maskers described above. Again, the situation where the phase is switched from N_0 to N_π exhibits a slower transition than the situation with constant interaural phase (N_π) and a 15-dB increase in the level of the masker. The data for the binaural situations are compatible with the results of a related experiment, previously reported by Grantham and Wightman [J. Acoust. Soc. Am. 65, 1509–1517 (1979)] and are well fit by a model that incorporates a double-sided exponential temporal integration window.

PACS numbers: 43.66.Pn, 43.66.Nm, 43.66.Dc, 43.66.Mk [WAY]

INTRODUCTION

Since the pioneering work of Hirsh (1948), the phenomena of binaural masking have been extensively investigated (see Durlach and Colburn, 1978, for a review). The vast majority of studies have focused upon stationary conditions, that is, situations where the interaural parameters of the signal and of the masker are fixed within a given trial. This contrasts with most real-world listening situations where the interaural parameters of the stimulus fluctuate as a result of both head and source movement. Obviously, the temporal properties of the binaural system play an important role in these situations, and an analysis of these properties is necessary if we hope to understand human perception in the complex stimulus situations of everyday life.

Some insights into the temporal limits of binaural processing have been obtained by direct extension of techniques previously applied to monaural phenomena. For example, Blodgett *et al.* (1958) and Green (1966) investigated the detectability of a tonal signal in the presence of a continuous noise masker as a function of the duration of the signal and found that the ear can integrate energy linearly up to a maximum integration time of 100–200 ms under both "monaural" and "binaural" conditions.[1]

In contrast, while under monaural conditions the detectability of a brief probe tone is almost independent of the duration of a longer simultaneous masker, under binaural conditions detectability increases dramatically with masker duration until the duration of the portion of the masker that precedes or follows the probe tone exceeds 200–600 ms. Although the results of these experiments are not completely consistent, it has been assumed that the binaural system requires several hundred milliseconds to determine and respond to the parameters of the masker (see, for example, McFadden, 1966; Robinson and Trahiotis, 1972; Trahiotis *et al.*, 1972; Zwicker and Zwicker, 1984; Yost, 1985; Kohlrausch, 1986; Kollmeier, 1986).

When a brief probe tone is presented after the termination of a masker (i.e., forward masking), the decay of masking is found to be more gradual under binaural conditions than under monaural conditions. Similar results are also found for backward masking. That is, under both forward and backward masking conditions, the masking level difference (MLD, i.e., the difference of the masked threshold in monaural versus binaural conditions) decreases as the signal is moved away from the masker in time (see Small *et al.*, 1972; Wightman, 1973; Berg and Yost, 1976). Unfortunately, there are several possible interpretations of these results that have not been resolved in the literature. One possibility is that the change in the level of the effective masker in binaural channels is more gradual than in monaural channels. Another explanation is that information about the interaural phase relation of the masker exists only during the time the masker is present (Kohlrausch and Fassel, 1988).

A third explanation notes that both monaural and binaural thresholds approach the same value (i.e., absolute threshold) as the probe tone is moved away from the masker. Said differently, the magnitude of the MLD is known to decrease with decreases in monaural masking. Thus we would expect the MLD to decrease because monaural of masking decreases as the probe tone is moved away from the masker.

A problem in interpreting the data from all these approaches is that the responses of both the monaural and the binaural systems have to be considered. That is, the stimulus manipulations influence both monaural and binaural "channels" and it is difficult to attribute the observed temporal effects to a particular system. Thus another approach to studying temporal effects in binaural unmasking solves these problems by using noise maskers without monaural changes, but with a temporally varying interaural correlation. Grantham and Wightman (1979) used a short interaurally phase-inverted (S_π) probe tone in the presence of a continuous noise masker whose interaural correlation varied sinusoidally between -1.0 and 1.0. When the rate of modulation is slow, an MLD is obtained when the probe tone is presented at a time when the interaural correlation of the masker is positive and no MLD is found when the masker correlation is near -1.0. As the modulation frequency increases, the difference in MLD between the positive and negative interaural correlation decreases rapidly and levels off for modulation frequencies above 4 Hz. This cutoff frequency is much lower than those obtained in comparable monaural experiments with amplitude-modulated stimuli (Viemeister, 1977). Grantham and Wightman (1979) termed this insensitivity to rapidly varying binaural cues as "binaural sluggishness" and estimated a "binaural minimum integration time" of 44–243 ms.

The evidence of binaural sluggishness, the concept of a binaural minimum integration time, and the order of magnitude of these time constants agree well with temporal properties of localization (Blauert, 1968), lateralization (Blauert, 1972; Grantham and Wightman, 1978; Grantham, 1984), and binaural correlation discrimination (Pollack, 1978; Grantham, 1982). In most of these experiments, however, the average performance of the binaural system is obtained for stimuli with periodically changing interaural parameters, and the transient properties of the binaural system in response to a rapid change in parameters are revealed only indirectly.

In the experiments described here, the change in the detectability of a brief probe tone is observed in response to a single (nonperiodic) transition in the "effective level" of the masker. This transition is introduced by rapidly changing the interaural phase of the masker or by changing the overall level of the masker in both ears. The conditions employed are analogous to forward and backward masking conditions. However, while performance in response to the overall level change is assumed to be governed by the monaural system, the interaural phase change will not alter the level of the effective masker in the monaural channels; thus performance is assumed to be governed by the binaural system. Hence, the transient properties of the binaural system should be revealed.

I. METHOD

A. Subjects

Two female and two male college students, aged between 19 and 25 years, were paid for participation in the experiments. All had clinically normal hearing and received at least 20 h of training before data collection began.

B. Apparatus

Signal and noise stimuli were generated on a Data General NOVA 4x computer. They were produced through separate 12-bit digital-to-analog converters for each binaural channel at a sampling rate of 10 000 samples/s and passed through 5-kHz low-pass elliptical filters. The level of the signal was controlled by separate programmable Charybdis attenuators for each subject. The signal and noise waveforms were then added with an analog mixer and presented to the subjects through TDH-49 headphones mounted in 001A cushions. The four subjects were seated in individual sound-attenuating chambers during the experiment. Timing and response recording were controlled by the computer.

C. Stimuli

In each interval, the masker was randomly sampled from the output of a 33-bit shift register whose repetition period was 5.2 days (Gilkey *et al.*, 1988) and was switched on for 750 ms without shaping the envelope. Transitions in interaural phase or in overall level were generated digitally by the appropriate computation of each noise sample for each binaural channel. The D/A-converted and low-pass-filtered masker was bandpass filtered from 100 to 2000 Hz with a Krohn-Hite 3270 filter and presented to the subjects at a reference spectrum level of 40 dB SPL/Hz. The signal was a 500-Hz sinusoid with a total duration of 20 ms, including 5-ms raised-cosine onset and offset ramps.

D. Conditions

The four conditions with masker transitions are sketched in Fig. 1. In the "binaural" conditions, an interaurally phase-inverted noise masker (N_π) is switched to interaurally in-phase [N_0, Fig. 1(a)] or an N_0 masker is switched to N_π [Fig. 1(c)].[1] Conversely, in the "monaural" conditions, the level of an N_π masker is lowered by 15 dB [Fig. 1(b)] or increased by 15 dB [Fig. 1(d)].

Two of the conditions were used in experiment I: the $N_\pi N_0 S_\pi$ condition [Fig. 1(a)], a 375- ms segment of interaurally phase inverted noise followed by a 375-ms segment of interaurally in-phase noise, and the $N_\pi(-15\ \mathrm{dB})N_\pi S_\pi$ condition [Fig. 1(b)], a 750-ms interaurally phase-inverted noise masker, which was attenuated by 15 dB 375 ms after its onset. The masked threshold of the interaurally phase-inverted probe tone (S_π) was measured as a function of the delay time between the transition of the noise and the signal *offset*. In this experiment, the delay time varied from -180 to $+320$ ms.

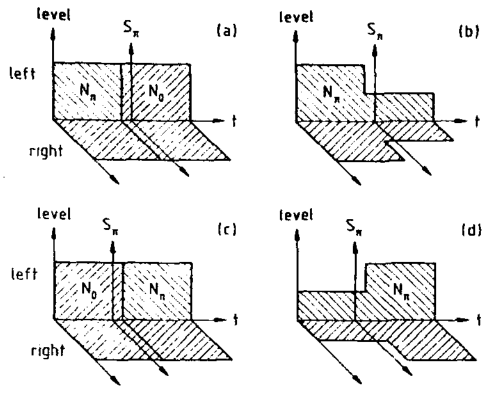

FIG. 1. Schematic diagram of the masker configurations and signals used in the experiments. Panel (a) depicts the $N_\pi N_0 S_\pi$ configuration, where the interaural phase relation of the masking noise is switched from 180° to 0° without changing the masker level. Panel (b) depicts the $N_\pi(-15\ \text{dB})N_\pi S_\pi$ configuration, where the interaural phase of the masker is held constant, but the masker level is lowered by 15 dB. The two lower panels (c) and (d) show the "backward masking" situations obtained by time reversing the "forward masking" situations sketched in the upper panel. Panel (c) depicts the $N_0 N_\pi S_\pi$ configuration, and panel (d) depicts the $(-15\ \text{dB})N_\pi N_\pi S_\pi$ configuration. The threshold of a 20-ms 500-Hz S_π probe tone is obtained as a function of the delay time between tone offset [panels (a) and (b)] or onset [panels (c) and (d)] and the switching of the masker.

In experiment II, the same maskers as in experiment I were used, but in a reversed temporal order: the $N_0 N_\pi S_\pi$ condition [Fig. 1(c)], a 375-ms segment of interaurally in-phase noise, followed by a 375-ms segment of interaurally phase-inverted noise, and the $(-15\ \text{dB})N_\pi N_\pi S_\pi$ condition [Fig. 1(d)], a 750-ms N_π masker, which is attenuated by 15 dB during the first 375-ms segment and is not attenuated during the second 375-ms segment. Again, the masked threshold of the S_π probe tone was determined. In this experiment, however, the threshold was measured as a function of the delay time between signal *onset* and the transition of the masker. The delay time varied between -300 ms and $+100$ ms.

In addition, three reference conditions, referred to as "nontransient conditions," were employed, where both the level of the masker and its interaural phase were held constant throughout the 750-ms duration of the masker: the $N_\pi S_\pi$ condition (both noise and signal interaurally phase inverted), the $N_0 S_\pi$ condition (noise interaurally in-phase and signal interaurally phase inverted) and the $(-15\ \text{dB})$ $N_\pi S_\pi$ condition (noise and signal interaurally phase inverted, but the masker attenuated by 15 dB such that the spectrum level equaled 25 dB/Hz). The signal was temporally centered in the masker.

We assume that performance in the binaural conditions [Fig. 1(a) and (c)] is determined by the change of activity in a binaural processing system and that performance in the monaural conditions [Fig. 1(b) and (d)] is determined by the change of activity in a monaural processing system. In both cases, the masked threshold of a short 500-Hz S_π probe tone is assumed to represent the masker activity in the respective channel during the presence of the probe tone. Since the condition sketched in Fig. 1(a) is similar to the monaural forward masking condition sketched in Fig. 1(b), we call this condition "binaural forward masking." Similarly, the condition sketched in Fig. 1(c) is called "binaural backward masking."

E. Trial structure and measurement procedure

A two-interval forced-choice (2IFC) procedure was employed. Each trial started with a 198-ms warning light, followed by a 198-ms pause and the two 750-ms observation intervals, separated by a 250-ms interstimulus interval. At the end of the second observation interval, a 1500-ms interval was allowed for responding. In both observation intervals, a 40-ms marking light was turned on 20 ms before the time when the signal might occur. An additional 198-ms feedback light was provided to mark the interval that actually contained the probe tone. Trial-by-trial feedback was provided.

An adaptive staircase algorithm was used to control the probe-tone level, following the recommendations of Kollmeier *et al.* (1988). At the beginning of a track, the signal level was set well above the expected threshold and lowered by 1.0 dB after each correct response. As soon as the first incorrect response was recorded, the signal level was increased by 1.0 dB and a "one up/three down" rule was adopted (Levitt, 1971), which lowered the signal level by 1.0 dB after three successive correct responses and increased the level by 1.0 dB after one incorrect response. Each measurement block consisted of 60 trials, and all four subjects were tested simultaneously. The threshold estimate was obtained as the average of the levels presented on all trials after the third reversal. Each data point represents the median threshold estimate of at least four independent tracks for each subject.

II. RESULTS

A. Experiment I

The individual results of the forward masking experiment [cf. Fig. 1(a) and (b)] are given in Fig. 2(a)–(d) for all four subjects. The open squares represent median thresholds for the $N_\pi N_0 S_\pi$ condition and the triangles the thresholds for the $N_\pi(-15\ \text{dB})N_\pi S_\pi$ condition. The abscissa denotes the time delay between the masker transition and the *offset* of the probe tone. The probe-tone level at threshold is plotted on the ordinate. The 0-dB point corresponds to each individual subject's nontransient $N_\pi S_\pi$ threshold. The signal-to-noise ratios associated with these individual thresholds are given in Table I. The arrows in the right-hand corner depict the individual nontransient $N_0 S_\pi$ thresholds (solid arrows) and $(-15\ \text{dB})N_\pi S_\pi$ thresholds (dotted arrows).

In the $N_\pi N_0 S_\pi$ condition [open squares in Fig. 2(a)–(d)], a relatively high threshold level (exceeding the nontransient $N_\pi S_\pi$ threshold by up to 2 dB) is obtained for

TABLE I. Signal-to-noise ratios (E/N_0 in dB) for the nontransient $N_\pi S_\pi$, $N_0 S_\pi$ and (-15 dB) $N_\pi S_\pi$ thresholds for each individual subject. The (-15 dB)$N_\pi S_\pi$ threshold is related to the unattenuated masker level used for the $N_\pi S_\pi$ and $N_0 S_\pi$ condition. In addition, the difference ΔL denotes the difference between the individual nontransient $N_\pi S_\pi$ threshold and the $N_\pi S_\pi$ threshold value that has been fit to the data in Fig. 2.

Subject	$N_\pi S_\pi$ (dB)	$N_0 S_\pi$ (dB)	(-15 dB)$N_\pi S_\pi$ (dB)	ΔL (dB)
1	11.6	-4.2	-3.7	-1.58
2	11.6	-5.1	-3.6	-0.96
3	11.7	-1.7	-2.9	-1.67
4	10.8	-2.8	-4.0	-1.40

delay times less than -100 ms. Conversely, a relatively low threshold level (approximating the nontransient $N_0 S_\pi$ threshold) is obtained for delay times greater than 200 ms. This difference in threshold level agrees with the $N_0 S_\pi$ MLD, that is, the difference between the nontransient $N_\pi S_\pi$ and $N_0 S_\pi$ thresholds averaged across the four subjects, which amounts to 15.0 dB. For probe-tone delays between 100 and 200 ms, a continuous transition from the high threshold at negative delays to the low threshold at large positive delays is observed.

In order to compare the rate of threshold decay in response to a binaural phase transition with the rate of a comparable threshold decay in response to a monaural level change, the data of the N_π(-15 dB)$N_\pi S_\pi$ condition are also shown in Fig. 2(a)–(d) as triangles. Similar to the "binaural" case, a relatively high threshold level (exceeding the nontransient $N_\pi S_\pi$ threshold by up to 2 dB) is obtained at delay times less than -50 ms. Conversely, a relatively low threshold level [approximating the nontransient (-15 dB)$N_\pi S_\pi$ threshold] is obtained at delay times greater than 100 ms. This difference in threshold level was selected to match the average $N_0 S_\pi$ MLD as closely as possible (i.e., 15.0 dB). Again, there is a continuous transition between the high threshold at negative delays and the low threshold at large positive delays. The transition is more rapid than that for the $N_\pi N_0 S_\pi$ condition.

B. Experiment II

The individual results of the backward masking experiment [cf. Fig. 1(c) and (d)] performed by time reversing the masker sequences of experiment I are given in Fig. 2(e)–(h). The open squares represent median thresholds for the $N_0 N_\pi S_\pi$ condition, and the triangles show the median thresholds for the (-15 dB)$N_\pi N_\pi S_\pi$ condition. In this experiment, however,the abscissa denotes the time delay between the masker transition and the *onset* of the probe tone. The probe-tone level at threshold is plotted on the ordinate. As in Fig. 2(a)–(d), the individual nontransient $N_\pi S_\pi$ reference level corresponds to 0 dB and the nontransient $N_0 S_\pi$ and (-15 dB)$N_\pi S_\pi$ threshold levels for each individual subject are given as a solid arrow or dotted arrow, respectively.

The conditions of our experiments are analogous to temporal masking conditions, in that periods of reduced masking precede or follow periods of greater masking. To a good approximation, the results depicted in Fig. 2(e)–(h) are a mirror image of those in Fig. 2(a)–(d). However, as with traditional temporal masking experiments (cf. Fastl, 1976; Small *et al.*, 1972), the transitions are more rapid for the backward masking case [Fig. 2(e)–(h)] than for the forward masking case [Fig. 2(a)–(d)]. In addition, the differences in the rate of threshold change between monaural and binaural cases suggest that the binaural system reacts more "sluggishly" to temporally varying stimuli, also compatible with the previous literature (e.g., Grantham and Wightman, 1979). The solid lines and dotted lines in Fig. 2 denote threshold functions fit to the data using the approach described in Sec. III.

III. INTEGRATION MODEL

To quantify the rate of change in threshold for the binaural and monaural forward and backward masking tasks described above, we fit functions, with a minimum number of free parameters, to the data. These functions are shown in Fig. 2 as solid lines for the "binaural" conditions and as dotted lines for the "monaural" conditions. In this section, the assumptions and the model used to generate these functions are specified, and the obtained time constants are compared to the "minimum binaural integration time" discussed by Grantham and Wightman (1979).

Our model incorporates elements of the "equalization–cancellation (EC) theory" (Durlach, 1972), as well as some of the assumptions described by Grantham and Wightman, to obtain a relationship between the amount of binaural masking and the delay time t: The threshold of the probe tone is determined by the instantaneous masking level $L(t)$ at the output of a binaural noise-reduction processor (e.g., the "cancellation" mechanism in the EC theory), operating in the critical band around the probe-tone frequency. The instantaneous masking level $L(t)$ is determined by the average interaural cross-correlation coefficient $r(t)$. This relation is taken from the EC theory:

$$L(t) = L_M - 10\log\{(K+1)/[K - r(t)]\}, \qquad (1)$$

where L_M is the monaural masked threshold level, K represents the internal noise and $r(t)$ is the time-varying average interaural cross-correlation coefficient. Here, $r(t)$ is obtained by a weighted integration of the instantaneous interaural cross-correlation $\rho(t)$ during the preceding instants of time ($-\infty < t' < t$) or, more generally, by integrating over an infinite range of preceding and successive instants of time:

$$r(t) = \int_{-\infty}^{\infty} w(t-t')\rho(t')dt'. \qquad (2)$$

Here, $w(t-t')$ denotes one of several temporal window functions described in the following paragraph. Although the integration window extends over an infinite range in the future, this should not imply that the subject's performance violates the causality principle by extracting information yet to be presented in the future. The contradiction can be resolved by setting the window to zero at a sufficiently large positive time (which can be done without significantly altering the results) and by allowing the subject to introduce an

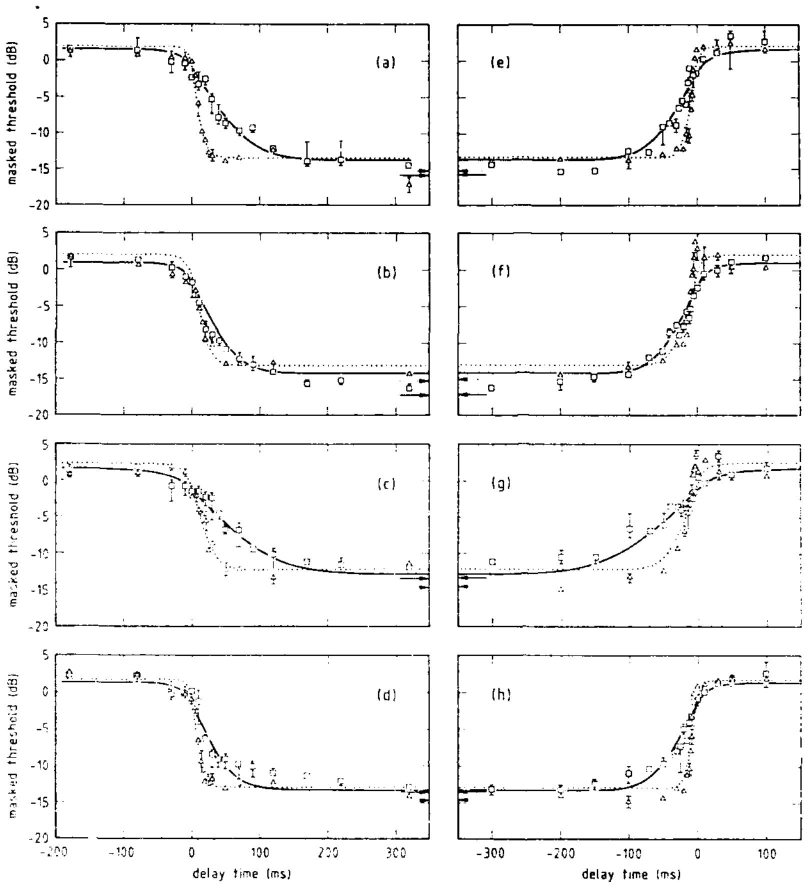

FIG. 2. Panels (a)–(d) show median estimates and interquartile ranges of the masked threshold for the $N_\pi N_0 S_\pi$ configuration (squares) and the N_π (− 15dB)N_0 S_π configuration (triangles) as a function of the delay time between offset of the probe tone and transition in the masker. Panels (e)–(h) show median estimates and interquartile ranges for the $N_0 N_\pi S_\pi$ configuration (squares) and the (− 15dB)N_0 N_π S_π configuration (triangles) as a function of the delay time between onset of the probe tone and transition in the masker. The individual results for subjects 1–4 are given in panels (a)–(d) and (e)–(h), respectively. The 0-dB point on the ordinate denotes the individual nontransient $N_\pi S_\pi$-thresholds, the solid arrows at the ordinate denote the nontransient $N_0 S_\pi$-thresholds, and the dotted arrows denote the nontransient (− 15 dB)$N_0 S_\pi$-thresholds. The solid lines represent model functions that have been fitted to the "binaural" data on the basis of a double-sided exponential temporal window. The dotted lines represent model functions fit to the "monaural" data with a similar algorithm (see text for details).

arbitrary delay between the sensory input and his judgment that is greater than or equal to this period of time.

Grantham and Wightman (1979) used a single-sided exponential temporal window and defined its time constant as the "binaural minimum integration time." In many psychophysical experiments, however, more complex temporal windows are required to account for the data (e.g., Moore *et al.*, 1988). On the other hand, if too many free parameters are needed to describe a particular temporal window, their fitted values will have little explanatory utility. For this reason, the discussion here will be restricted to the following temporal window functions (each of which requires only two parameters):

(1) rectangular window

$$w(t) = (\tau_1 + \tau_2)^{-1} \quad \text{for} \quad -\tau_1 \leqslant t < \tau_2$$
$$= 0 \quad \text{elsewhere;} \qquad (3)$$

(2) triangular window

$$w(t) = 2(\tau_1 + t)/(\tau_1^2 + \tau_1\tau_2) \quad \text{for} -\tau_1 < t < 0$$
$$= 2(\tau_2 - t)/(\tau_2^2 + \tau_1\tau_2) \quad \text{for} -0 < t < \tau_2$$
$$= 0 \quad \text{elsewhere}; \tag{4}$$

(3) Gauss window

$$w(t) = \frac{1}{(\tau_1 + \tau_2)\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\tau_1 + t}{\tau_1 + \tau_2}\right)^2\right]; \tag{5}$$

(4) double-sided exponential window

$$w(t) = \frac{\exp(t/\tau_1)}{(\tau_1 + \tau_2)} \quad \text{for} \quad t \leqslant 0$$
$$= \frac{\exp(-t/\tau_2)}{(\tau_1 + \tau_2)} \quad \text{for} \quad t > 0; \tag{6}$$

(5) rounded exponential window

$$w(t) = \frac{1 - (t/\tau_1)}{2(\tau_1 + \tau_2)} \exp\left(\frac{t}{\tau_1}\right) \quad \text{for} \quad t \leqslant 0$$
$$= \frac{1 + (t/\tau_2)}{2(\tau_1 + \tau_2)} \exp\left(\frac{-t}{\tau_2}\right) \quad \text{for} \quad t > 0. \tag{7}$$

Note that the integral of these window functions is normalized to unity. The "equivalent rectangular duration" is the inverse of the maximum value of the respective temporal window.[3]

The idealized time-dependent interaural cross-correlation $\rho(t)$ is derived from the idealized instantaneous interaural phase of the noise masker. For the binaural "forward masking" experiment $N_\pi N_0 S_\pi$ [Fig. 2(a)–(d)], it is given by

$$\rho(t) = -1 \quad \text{for} \quad t \leqslant 0$$
$$= +1 \quad \text{for} \quad t > 0. \tag{8}$$

The function $\rho(t)$ for the binaural "backward masking" experiment $N_0 N_\pi S_\pi$ [Fig. 2(e)–(h)] is obtained by time reversing Eq. (8).

By inserting a window function selected from Eqs. (3)–(7) and Eq. (8) into Eq. (2), $L(t)$ can be calculated from Eq. (1) for both the binaural forward and backward masking experiment. We do, however, make an additional assumption. Presumably, the subject can manipulate the temporal center of his window to increase the effective signal-to-noise ratio ("off-time listening"). To allow for this effect, we assume that the subject centers his window at the *onset* of the signal under the backward masking condition, and at the *offset* of the signal under the forward masking condition. The parameters that have to be fit to the data of each subject are the time constants τ_1 and τ_2 and the values of L_M and K. In principle, L_M is given by the nontransient $N_\pi S_\pi$ reference threshold and K is determined by the nontransient $N_0 S_\pi$ MLD for each individual subject. In the transient conditions with a change in the interaural masker phase, however, the threshold levels for very large positive and large negative values of t do not approach these nontransient thresholds, but are up to 2.0 dB higher.[3] Since this increase in threshold level is roughly the same before and after the transition of the masker, only the value of L_M has to be fit to the data, while the value of K is determined by the individual nontransient $N_0 S_\pi$ MLD. Further, the three free parameters L_M, τ_1, and τ_2 are fit to the data of the binaural forward masking experiment ($N_\pi N_0 S_\pi$) and the backward masking experiment ($N_0 N_\pi S_\pi$) simultaneously. Thus a single least-squares fit between the data and the two model functions is found using the Simplex method (Nedler and Mead, 1965). The time constants τ_1 and τ_2 and the normalized deviation measure B_{nl} obtained for each integration window are given in Table II for each individual subject.[4]

Figure 3 shows threshold functions for the five choices of the binaural integration window [Eqs. (3)–(7)] fit to the "binaural" data of subject 1. The data points are taken from Fig. 2(a) (condition $N_\pi N_0 S_\pi$, triangles) and Fig. 2(e) (condition $N_0 N_\pi S_\pi$, inverted triangles). The ordinate and the abscissa are the same as in Fig. 2.

Obviously, the dotted curves in Fig. 3, which were computed with a rectangular window, do not yield an adequate description of the experimental data. The long-dashed curves in Fig. 3 are based on a triangular window [Eq. (4)]. As with the rectangular window, these curves do not give an adequate description of the smooth transition in threshold for delays smaller than − 50 ms or greater than 50 ms. For this reason, integration windows such as the Gaussian, the double-sided exponential, and the rounded exponential win-

TABLE II. Time constants τ_1 and τ_2 (in ms), fit to the "binaural" forward and backward masked threshold data from Fig. 2, for five different shapes of the integration window and each individual subject. The normalized deviation measure B_{nl} (see footnote 4) is also included for the data obtained with each subject.

		Subject 1	Subject 2	Subject 3	Subject 4
Rectangular window	τ_1/τ_2 (ms)	53.1/56.4	52.8/53.9	83.2/92.8	33.2/41.4
	B_{nl}	0.894	0.864	0.895	0.922
Triangular window	τ_1/τ_2 (ms)	77.5/102	61.6/64.6	137/129	46.1/49.6
	B_{nl}	0.925	0.909	0.933	0.946
Gaussian window	τ_1/τ_2 (ms)	− 5.5/37.2	− 2.3/29.7	0.5/53.1	− 1.8/22.4
	B_{nl}	0.941	0.928	0.941	0.957
Exponential window	τ_1/τ_2 (ms)	22.2/27.6	19.8/21.0	42.7/40.5	16.1/17.1
	B_{nl}	0.963	0.966	0.956	0.968
Rounded exponential window	τ_1/τ_2 (ms)	15.0/18.9	13.6/14.6	28.8/27.5	10.6/11.4
	B_{nl}	0.958	0.952	0.952	0.965

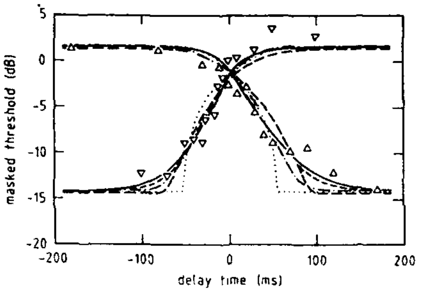

FIG 3. Comparison of different functions fitted to the binaural data of subject 1 from Fig. 2. The upright triangles denote the $N_\pi N_0 S_\pi$ condition, and the inverted triangles denote the $N_0 N_\pi S_\pi$ condition. Dotted line: rectangular window; long-dashed line: triangular window; dash-dotted line: Gaussian window; solid line: double-sided exponential window; short-dashed line: rounded exponential window.

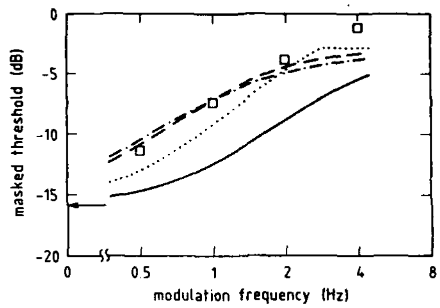

FIG. 4. Predictions for the data of Grantham and Wightman (1979) based on a rectangular temporal window (dotted curve: $\tau_1 + \tau_2 = 175$ ms), a single-sided exponential temporal window (dash-dotted curve: $\tau_1 = 0$, $\tau_2 = 181$ ms), and a double-sided exponential temporal window (solid curve: $\tau_1 = 42.7$ ms, $\tau_2 = 40.5$ ms, dashed curve: $\tau_1 = \tau_2 = 114$ ms). The solid and dotted curves are based on the time constants estimated for subject 3 from the data given in Fig. 2(c) and (g). The dash-dotted and dashed curves were obtained by a least-squares fit to the data provided in this figure. The abscissa denotes the modulation frequency of a noise masker with sinusoidally time-varying interaural correlation, and the ordinate denotes the masked threshold of an S_π probe tone presented simultaneously with the occurrence of a maximally positive interaural correlation of the masker. The 0-dB point denotes the stationary $N_\pi S_\pi$ threshold, and the arrow on the left-hand side denotes the stationary $N_0 S_\pi$ threshold. Squares denote average threshold values of the three subjects of Grantham and Wightman (1979).

dow, yield more appropriate model functions. The problem with the Gaussian window is that it produces the same slopes for forward and backward masking (dashed-dotted lines in Fig. 3), although the time course of the data clearly shows asymmetries between forward and backward masking. The double-sided exponential window [Eq. (6)] provides the best fit for the binaural data presented here (solid line in Fig.3) since the normalized "goodness of fit" measure B_{nl} (cf. Table II) takes its maximum value for all four subjects. The functions fit to each individual subject's data with this type of a temporal window are included as solid lines in Fig.2. The second best fit is obtained with the rounded exponential window [Eq. (7)]. The model functions fitted with this window (short-dashed lines in Fig. 3) deviate only marginally from those obtained with the double-sided exponential window.

Although the single-sided exponential temporal window was not specifically examined here (because its predictions for the backward masking case differ substantially from the data), the "equivalent rectangular durations" for all of our subjects and for all windows considered agree well with Grantham and Wightman's "binaural minimum integration time" estimates (between 44 and 243 ms).

Although Grantham and Wightman's experiments are rather different from ours and the range of their time constants exceeds the range of the time constants reported here, the relation between the two methods should be considered in greater detail to assure that a substantial disagreement does not exist between both sets of data. Therefore, we attempt to predict their data with our time constants for the rectangular and the double-sided exponential temporal window. In addition, the ability of a double- and a single-sided exponential temporal window to predict their data is examined. In the Appendix, the maximum MLD is calculated for Grantham and Wightman's maskers with sinusoidally varying interaural correlation as a function of the modulation frequency. The predictions for the rectangular and the double-sided exponential temporal window are plotted in Fig. 4 using the decay constants obtained for our subject 3, whose fitted decay constants for $N_\pi N_0 S_\pi$ and $N_0 N_\pi S_\pi$ were the largest (i.e., closest to the middle of the range reported by Grantham and Wightman). The dotted curve denotes theoretical values for a rectangular window ($\tau_1 + \tau_2 = 175$ ms) and the solid line denotes the curve for a double-sided exponential window ($\tau_1 = 42.7$ ms and $\tau_2 = 40.5$ ms). The open squares denote the mean MLD values for 500 Hz for the three subjects of Grantham and Wightman (1979), as supplied by one of the authors. The normalized deviation from these data for the curves predicted for our subject 3 is $B_{nl} = 0.840$ (rectangular window, dotted line in Fig. 4) and $B_{nl} = -0.363$ (double-sided exponential window, solid line in Fig. 4). Both curves are below the data supplied by Grantham and Wightman for all modulation frequencies, indicating that even the largest time constants found in our study are below the values required to predict Grantham and Wightman's average data. However, since the deviation between the dotted curve and Grantham and Wightman's data is small, the first-order prediction from our data based on a rectangular temporal window is in approximate agreement with Grantham and Wightman's findings.

The dashed-dotted curve and the dashed curve were fitted to the average data by using a single-sided exponential window ($\tau_1 = 0$, $\tau_2 = 181$ ms, $B_{nl} = 0.841$) and a double-sided exponential window ($\tau_1 = \tau_2 = 114$ ms, B_{nl}

TABLE III. Same as Table II for the "monaural" forward and backward masked thresholds from Fig. 2.

		Subject 1	Subject 2	Subject 3	Subject 4
Recantangular window	τ_1/τ_2 (ms)	13.3/15.8	20.8/17.0	32.0/26.4	12.6/12.5
	B_{nl}	0.943	0.888	0.910	0.961
Triangular window	τ_1/τ_2 (ms)	19.9/21.4	34.9/28.7	44.3/37.0	18.4/19.6
	B_{nl}	0.950	0.891	0.919	0.964
Gaussian window	τ_1/τ_2 (ms)	−0.1/8.7	2.6/9.6	3.4/13.0	−0.02/7.8
	B_{nl}	0.954	0.909	0.928	0.964
Exponential window	τ_1/τ_2 (ms)	6.5/6.4	11.4/8.7	14.7/11.3	5.9/6.0
	B_{nl}	0.947	0.900	0.911	0.950
Rounded exponential window	τ_1/τ_2 (ms)	4.4/4.5	7.5/5.9	9.8/7.8	4.0/4.2
	B_{nl}	0.905	0.905	0.918	0.957

$= 0.908$), respectively. As can be seen, Grantham and Wightman's average data are about equally well fit using a single- and a double-sided exponential window.

In terms of a systems analysis approach applied to the binaural unmasking mechanism (such as the EC device), our data describe something analogous to the "envelope step response," while the data of Grantham and Wightman (1979) and Grantham (1982) describe something analogous to the "modulation transfer function" (i.e., the modulation transfer function is defined as the amount of modulation detectable at the output of the system in response to an envelope-modulated signal, while Grantham and Wightman measured the response to a signal whose interaural correlation was modulated). If the system can be approximated as linear and time invariant, both descriptions would be equivalent and the formulas given in the Appendix would give the appropriate relation between our approach and Grantham and Wightman's approach: An exponentially rising and decaying step response would correspond to an exponential temporal window and a modulation transfer function given by Eq. (A5), which decreases with $(1/f_m)^2$. On the other hand, a modulation transfer function decreasing with $(1/f_m)$ would correspond to either a rectangular or a single-sided exponential temporal window. Hence, the slope of the modulation transfer function for large values of f_m could differentiate between different types of temporal windows. Unfortunately, this slope cannot be estimated precisely from experimental data because the MLD decreases rapidly as the modulation frequency increases. Therefore, no significant difference is observed for the single- and double-sided exponential temporal windows in predicting Grantham and Wightman's data (cf. Fig. 4). However, large differences between different types of temporal windows are observed for our data, which therefore appear to be more appropriate for describing the interaural correlation averaging process.

One significant difference between our data and Grantham and Wightman's data is the fact that the time constants of the double-sided exponential window required to fit their average data are higher than those obtained from our data (cf. Table II). This fact also holds for the time constants of the rectangular window. However, since the four modulation frequencies employed by Grantham and Wightman might not provide a sufficient database to discriminate among the different two-parameter windows, only a qualitative agreement with our data should be expected. This agreement is established by the fact that our range of rectangular window widths, and equivalent rectangular durations of the double-sided exponential windows (with the exception of subject 4), is completely contained within their range of time constants for the single-sided exponential window (44–243 ms).

In order to compare the time constants obtained for the binaural conditions $N_\pi N_0 S_\pi$ and $N_0 N_\pi S_\pi$ with those for the monaural conditions $N_\pi(-15\text{ dB})N_\pi S_\pi$ and $(-15\text{ dB})N_\pi N_\pi S_\pi$, the same algorithm was used to fit theoretical curves to the data. The dotted lines in Fig. 2 denote these functions generated with a double-sided exponential temporal window. However, since the interaural correlation of the masker did not change in these conditions, we rewrite Eq. (1) as

$$L(t) = L_{-15\text{ dB}} + 10\log[1 + b(t)], \tag{9}$$

where $L_{-15\text{ dB}} = L_M - 10\log(K+1)/(K-1)$ denotes the nontransient $(-15\text{ dB})N_\pi S_\pi$ threshold, and $b(t)$ is interpreted as the weighted average of that part of the masker power that exceeds the -15-dB level of the masker.[5] For this reason, Eq. (9) simply describes a power-law additivity of monaural thresholds: In the nontransient $N_\pi S_\pi$ case $[b(t) = 2/(K-1)]$ or $(-15\text{ dB})N_\pi S_\pi$ case $[b(t) = 0]$, the respective threshold values L_M or $L_{-15\text{ dB}}$ are obtained, whereas the transient properties of the threshold are determined by a moving weighted average of the masker power. Similar approaches have been described by Robinson (1974) and Moore *et al.* (1988). The time constants derived for different types of temporal windows are given in Table III. They are substantially smaller than those obtained for the binaural data with the same temporal window function, and tend to be higher than those obtained by Moore *et al.* (1988). In addition, the Gaussian temporal window appears to predict our monaural data better than the exponential window (which provides the best fit for our binaural data) and a rounded exponential temporal window (which provides the best fit for the data of Moore *et al.*, 1988). Since the dynamic range of the data presented here is small in comparison to the measurement accuracy of about 1 dB, only properties of the central part of the temporal window can be estimated, whereas its shape at times remote from the center and its total dynamic range cannot be predicted accurately. Therefore, a detailed discussion of the shape of the monaural temporal weighting window and the limited explanatory value of such temporal windows (Püschel, 1988) is beyond the scope of this paper.

IV. DISCUSSION

From the data presented here, it is evident that the binaural system reacts more sluggishly to a change in the "effective level" of the masker than does the monaural system. Similar findings of binaural sluggishness have been reported by several authors (Blauert, 1968, 1972; Grantham and Wightman, 1978, 1979; Zurek and Durlach, 1987; Grantham, 1982, 1984) using different psychoacoustic experiments and different binaural cues. There are at least two problems with comparing their time constants to those reported here.

First, the definition of a "time constant" varies from experiment to experiment. As shown, when appropriate corrections are made, our time constants agree to a first approximation with those of Grantham and Wightman (1979). Other authors have derived time constants between 200 and 500 ms from the dependence of the MLD on masker duration in simultaneous masking [Yost, 1985 (see footnote 6); Kohlrausch, 1986] and nonsimultaneous masking (Small *et al.*, 1972; Lakey, 1976). Unfortunately, all underlying model assumptions were not specified exactly in these studies, so that a quantitative analysis that relates our comparatively low time constants to their findings is impossible.

Second, since different experiments test different binaural abilities, i.e., localization (Blauert, 1968), lateralization (Blauert, 1972; Grantham and Wightman, 1978; Pollack, 1978; Grantham, 1982), and detection (Grantham and Wightman, 1979; Kollmeier, 1986; Kohlrausch, 1986), it is possible that many of these time constants estimate different properties of the binaural system, even though most binaural time constants are of the same order of magnitude.

Although the variability in the data within studies and between studies that use similar techniques might be used to argue against this second point, the possibility of multiple different sources of binaural sluggishness has some appeal. In our experiment, the slow decay of binaural masking might simply be attributed to forward masking in a comparatively sluggish binaural transmission channel (e.g., at the output of the "C" mechanism of Durlach, 1972), while the masking level decays more slowly than in a monaural channel (e.g., at either of the monaural inputs to the "decision unit" of Durlach, 1972). Since detection would occur whenever the level of the probe tone exceeds the temporary masking level in any of these channels, the performance in our "monaural" conditions would be determined by the fast decay in the monaural channels, whereas the slow decay of the binaural channel would only show up in our "binaural" conditions. For the present discussion, we denote the slow decay in the binaural channel as "binaural channel sluggishness." When predicting dynamic localization and interaural correlation discrimination tasks with the same kind of binaural model, a more sophisticated detection strategy has to be assumed. With Durlach's EC model, for example, the outputs of several hypothetical "EC" mechanisms tuned to different interaural delays and intensity differences have to be monitored over time, and discrimination would occur if any change in the distribution of output levels is observed. The delay associated with establishing and comparing these distributions from the multiple sluggish "binaural channels" over time might result in an additional binaural sluggishness that we denote as "binaural analyzer sluggishness." For example, binaural analyzer sluggishness might contribute to the time constants of about 200 ms reported by Blauert (1968, 1972), who studied the maximum detectable frequency of switching between different sound source locations.[7]

In most experiments on dynamic properties of the binaural system, both hypothetical sources of binaural sluggishness might influence the obtained time constants. In our experiment, however, the comparatively small time constants could suggest a negligible role of binaural analyzer sluggishness: The binaural system might attempt to detect the S_π probe tone by adopting a fixed binaural processing strategy (such as subtracting both monaural stimuli without interaural delay in the C mechanism of Durlach, 1972, or inspecting only one optimum interaural time delay "place" of Jeffress, 1948) and therefore monitoring the output of only one "sluggish" binaural channel. In this case, the detection process could be similar to the monaural case in that the binaural analyzer would not have to compare activity across channels of the binaural display or within channels as a function of time. The threshold is determined by the effective level of the masker at the output of the binaural channel and thus is influenced by binaural channel sluggishness, but not by binaural analyzer sluggishness. Within this view, the measurements reported here would provide a lower limit on estimates of binaural sluggishness.

ACKNOWLEDGMENTS

The authors wish to thank M. Kinkel, A. Kohlrausch, B. C. J. Moore, J. M. Weisenberger, D. W. Grantham, and two anonymous reviewers for helpful comments on earlier drafts of this paper. This work was supported by a grant from NINCDS (NS 03856) to the Central Institute for the Deaf, and by the Fulbright Commission of the Federal Republic of Germany, who supported the first author during his stay at Washington University in St. Louis. Additional support was provided by AFOSR (86-0298, 89-0302) and NSF (BNS-85-11768, BNS-87-20305).

APPENDIX

The stimulus used by Grantham and Wightman (1979) had a time-dependent interaural cross-correlation coefficient that varied sinusoidally at the modulation frequency $\omega = 2\pi f_m$:

$$\rho(t) = \sin \omega t. \tag{A1}$$

The average interaural cross-correlation coefficient $r(t)$ is obtained by substituting Eq. (A1) and the appropriate window function [Eqs. (3)–(7)] into Eq. (2), such that

$$r_\omega(t) = \int_{-\infty}^{\infty} w(t-t')\sin \omega t' \, dt' = \mathrm{Im}[e^{-j\omega t} \cdot W(\omega)]. \tag{A2}$$

In this equation, Im denotes the imaginary part of a complex number, j is $\sqrt{-1}$, and $W(\omega)$ is the Fourier transform of the window $w(t)$. The right-hand side of Eq. (A2) follows from the definition of the Fourier transform of a continuous function $w(t')$ and its transformation properties if the function is time reversed and shifted by an amount t. For a fixed modulation frequency f_m, the largest MLD with a short S_π probe tone is obtained for the largest possible positive value of $r(t)$, which is simply the modulus of $W(\omega)$:

$$r_{\max}(\omega) = \max_{-\infty < t < \infty} [r_\omega(t)] = |W(\omega)|. \tag{A3}$$

For a rectangular temporal window [Eq. (3)] and comparatively low modulation frequencies [$f_m < 1/2(\tau_1 + \tau_2)$], the largest possible value of $r(t)$ is

$$r_{\max}(\omega|\text{rectangular}) = \sin \omega(\tau_1 + \tau_2)/\omega(\tau_1 + \tau_2). \tag{A4}$$

Note that $r_{\max}$ decreases with $1/\omega$.

The largest possible value of $r(t)$ for a double-sided exponential window [Eq. (6)] is

$$r_{\max}(\omega|\text{exponential}) = [(1 + \omega^2\tau_1^2) \times (1 + \omega^2\tau_2^2)]^{-1/2}. \tag{A5}$$

The respective value for the single-sided exponential window is obtained by setting τ_1 to zero in Eq. (A5). Thus $r_{\max}$ for the single-sided exponential, like the rectangular window, decreases with $1/\omega$.

By inserting Eq. (A4) or (A5) into Eq. (1), the minimum binaural masked threshold, which corresponds to the maximum MLD, can be calculated for Grantham and Wightman's maskers. These predictions are plotted in Fig. 4 versus the modulation frequency $f_m = \omega/2\pi$.

[1]We use the term "monaural" to refer to conditions where performance *would not be* expected to be appreciably altered if the stimuli had been presented to only one ear. Note that many such conditions would not be truly monaural. Similarly, we refer to "binaural" conditions as those where performance *would be* expected to be appreciably altered if the signal and masking stimuli had been presented to only one ear. Note that not all conditions where stimuli are presented to both ears meet this criterion.

[2]There are two possible ways to normalize temporal windows: First, their integral can be normalized to unity, and, second, their maximum value can be normalized to unity. Our approach uses the first method, consistent with Eq. (2) and with the linear amplitude scale on which the window functions were defined. The second method would be appropriate if the window functions had been defined on a logarithmic amplitude scale with their maximum value equal to 0 dB. To compare different window functions normalized with the second method, the "equivalent rectangular duration" is defined as specified by Moore *et al.* (1988). That is, the integral of the window function is equated to the integral of a rectangular window with amplitude 1.0.

For our window functions normalized with the first method, the equivalent rectangular duration is computed by first scaling the maximum value to unity and then calculating the integral of the functions. Therefore, the equivalent rectangular duration is given by the inverse of the temporal windows maximum value. With a double-sided exponential window, for example, the equivalent rectangular duration is simply the sum of τ_1 and τ_2.

[3]One possible explanation for the 2-dB discrepancy between the thresholds for very large positive and negative values of t and the respective nontransient thresholds is that switching the interaural phase produces a transient auditory sensation that is reported by the subjects to sometimes sound like the probe tone. The perception of this transient, as well as the transient produced by a sudden increase or decrease in masker level, might confuse the subject even if the probe tone is presented considerably before or after switching the masker.

[4]The normalized, nonlinear deviation measure B_{nl} was computed as

$$B_{nl} = 1 - \sum_{i=1}^{J} [L_i - L(t_i)]^2 \Big/ \sum_{i=1}^{J} (L_i - \bar{L})^2,$$

where L_i ($i = 1,...,J$) are the J threshold values obtained at the time delays t_i ($i = 1,...,J$), $L(t_i)$ is the model function at time t_i, and $\bar{L}$ is the average of all J threshold values. Since B_{nl} is always less than or equal to one, the model function that fits the data best yields the highest value of B_{nl}.

[5]In the binaural conditions, $r(t)$ is interpreted as the average interaural correlation coefficient that is given in units of signal power and ranges between -1 and 1. For an analogous description of the monaural condition, the function $q(t)$ replaces $r(t)$ such that $[1 - q(t)]$ is interpreted as average masker power and ranges between 0 and 2. By normalizing $[1 - q(t)]$, the resulting expression $b(t) = [1 - q(t)]/(K - 1)$ can be interpreted as the weighted average of that part of the masker power that exceeds the -15-dB level of the masker. Hence, Eqs. (1) and (9) are equivalent, although their parameters would have different interpretations.

[6]It should also be noted that the data of Yost (1985) appear to conflict *qualitatively* with the data presented here. That is, Yost's results imply that the N_π segment of the noise should not influence the detectability of a probe tone that is presented in the N_0 segment of the noise. These contradictions are considered in greater detail by Gilkey *et al.* (1990), who tried and failed to replicate the results of Yost. One possibility is that the different findings are due to short gaps before switching the interaural phase in the binaural noise stimuli employed by Yost, whereas no artifact was present in the stimuli used here and by Gilkey *et al.*

[7]Alternatively, both hypothetical types of binaural sluggishness can be defined in terms of an interaural delay line model of the type first proposed by Jeffress (1948). Binaural channel sluggishness would correspond to the attack and decay rate of a sound image located at a certain interaural delay time; whereas binaural analyzer sluggishness would correspond to the temporal properties of the binaural analyzer that evaluates the patterns along the interaural delay line.

Berg, K., and Yost, W. A. (**1976**). "Temporal masking of a click by noise in diotic and dichotic listening conditions," J. Acoust. Soc. Am. **60**, 173-177.

Blauert, J. (**1968**). "Ein Beitrag zur Trägheit des Richtungshörens in der Horizontalebene," Acustica **20**, 200–205.

Blauert, J. (**1972**). "On the lag of lateralization caused by interaural time and intensity differences," Audiology **11**, 265–270.

Blodgett, H. C., Jeffress, L. A., and Taylor, R. W. (**1958**). "Relation of masked threshold to signal-duration of various interaural phase-combinations," Am. J. Psychol. **71**, 283-290.

Durlach, N. I. (**1972**). "Equalization and cancellation theory," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York).

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomena," in *Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), Vol. IV, pp. 365–466.

Fastl, H. (**1976**). "Temporal masking effects I: Broadband noise maskers," Acustica **35**, 287–302.

Gilkey, R. H., Robinson, D. E., and Frank, A. S. (**1988**). "A software pseudorandom noise generator," J. Acoust. Soc. Am. **83**, 829–831.

Gilkey, R. H., Simpson, B. D., and Weisenberger, J. M. (**1990**). "Masker fringe and binaural detection," J. Acoust. Soc. Am. (submitted).

Grantham, D. W. (**1982**). "Detectability of time-varying interaural correlation in narrow-band noise stimuli," J. Acoust. Soc. Am. **72**, 1178–1184.

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.

Grantham, D. W., and Wightman, F. L. (**1979**). "Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation," J. Acoust. Soc. Am. **65**, 1509–1517.

Green, D. M. (**1966**). "Interaural phase effects in the masking of signals of different duration," J. Acoust. Soc. Am. **39**, 720–724.

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

Jeffress, L. A. (1948). "A place theory of sound localization," J. Comp. Physiol. Psychol. 41, 35–39.

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

Kohlrausch, A., and Fassel, R. (1988). "Vergleich nichtsimultaner Mithörschwellen in monauralen und binauralen Maskierungsexperimenten," in *Fortschritte der Akustik* (DPG-Kongressgesellschaft, Bad Honnef, Federal Republic of Germany), pp. 593–596.

Kollmeier, B. (1986). "Entwicklung zeitoptimierter psychoakustischer Messverfahren und ihre Anwendung auf binaurale Zeiteffekte," Ph.D. Thesis, Universitat Gottingen.

Kollmeier, B., Gilkey, R. H., and Sieben, U. K. (1988). "Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model," J. Acoust. Soc. Am. 83, 1852–1862.

Lakey, J. R. (1976). "Temporal masking-level differences: The effect of mask duration," J. Acoust. Soc. Am. 59, 1434–1442.

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

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

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (1988). "The shape of the ear's temporal window," J. Acoust. Soc. Am. 83, 1102–1116.

Nedler, J. A., and Mead, R. (1965). "A simplex method for function minimization," Comput. J. 7, 308.

Pollack, I. (1978). "Temporal switching between binaural information sources," J. Acoust. Soc. Am. 63, 550–558.

Püschel, D. (1988). "Prinzipien der zeitlichen Analyse beim Hören," Ph.D. Thesis, Universität Göttingen.

Robinson, C. E. (1974). "Auditory running average hypothesis," J. Acoust. Soc. Am. 55, 645–648.

Robinson, D. E., and Trahiotis, C. (1972). "Effects of signal duration and masker duration on detectability under diotic and dichotic listening conditions," Percept. Psychophys. 12, 333–334.

Small, A. M., Bogges, J., Klich, R., Kuehn, D., Thelin, J., and Wiley, T. (1972). "MLDs in forward and backward masking," J. Acoust. Soc. Am. 51, 1365–1367.

Trahiotis, C., Dolan, T. R., Miller, T. H. (1972). "Effect of backward masker fringe on the detectability of pulsed diotic and dichotic tonal signals," Percept. Psychophys. 12, 335–338.

Viemeister, N. F. (1977). "Temporal factors in audition: A systems analysis approach," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 419–428.

Wightman, F. L. (1973). "Binaural detection in forward and backward masking," J. Acoust. Soc. Am. 54, 309(abs).

Yost, W. A. (1985). "Prior stimulation and the masking-level difference," J. Acoust. Soc. Am. 78, 901–907.

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

Zwicker, U., and Zwicker, E. (1984). "Binaural masking-level differences as a function of masker and test-signal duration," Hear. Res. 13, 215–221.

956

Masker fringe and binaural detection

Robert H. Gilkey, Brian D. Simpson, and Janet M. Weisenberger
Signal Detection Laboratory, Central Institute for the Deaf, 818 S. Euclid, St. Louis, Missouri 63110

(Received 5 July 1988; revised 2 September 1988; accepted 17 April 1990)

Yost [J. Acoust. Soc. Am. 78, 901–907 (1985)] found that the detectability of a 30-ms dichotic signal (Sπ) in a 30-ms diotic noise (No) was not affected by the presence of a 500-ms dichotic forward fringe (Nπ). Kollmeier and Gilkey [J. Acoust. Soc. Am. 87, 1709–1719, (1990)] performed a somewhat different experiment and varied the onset time of a 25-ms Sπ signal in a 750-ms noise that switched, after 375-ms, from Nπ to No. In contrast to Yost, they found that the Nπ segment of the noise reduced the detectability of the signal even when the signal was temporally delayed well into the No segment of the noise and suggested that the Nπ segment of noise acted as a forward masker. To resolve this apparent conflict, the present study investigated the detectability of a brief Sπ signal in the presence of an No masker of the same duration as the signal. The masker was preceded by quiet or an Nπ forward fringe and followed by quiet, an No, or Nπ backward fringe. The present study differs from most previous studies of the effects of the masker fringe in that the onset time of the signal was systematically varied to examine how masking changes during the time course of the complex fringe–masker–fringe stimulus. The results failed to replicate those of Yost in that an Nπ forward fringe reduced the detectability of the signal, and agreed with those of Kollmeier and Gilkey in that thresholds were elevated well after the offset of the Nπ segment of the noise. The addition of an Nπ backward fringe was also shown to reduce the detectability of the signal. Possible reasons for differences between the results of the present study and those of Yost are evaluated. Results are discussed in the context of models of binaural detection.

PACS numbers: 43.66.Dc, 43.66.Pn [WAY]

INTRODUCTION

McFadden (1966) investigated the detectability of a low-frequency (400-Hz), 125-ms, sinusoidal signal in the presence of wideband masking noise that was either continuous or that was pulsed on and off with the signal. Under diotic conditions, thresholds in continuous noise were slightly lower than thresholds in pulsed noise (about 0.4 dB).[1] However, under dichotic conditions, the difference in thresholds obtained with continuous and pulsed maskers was much larger, on the order of 4–6 dB. McFadden also examined conditions in which the masker onset occurred before the signal onset, but their offsets were simultaneous. He referred to the portion of the masker that occurred before the signal onset as a forward masker "fringe." Thresholds improved as the duration of this fringe was increased. Performance was comparable to the continuous masker condition when the duration of the fringe exceeded 600 ms.

Robinson and Trahiotis (1972) considered the influence of signal duration on this effect. Their results with a 261-ms signal replicated those of McFadden. With a 37-ms signal they found even larger differences between pulsed and continuous maskers (about 9 dB).[2]

Trahiotis *et al.* (1972) showed that the addition of a backward fringe after a pulsed masker has effects similar to, although not as strong as, those of a forward fringe (i.e., as the duration of a backward fringe is increased, performance approaches, but does not equal, that for a continuous masker).

The differences between pulsed and continuous maskers observed under dichotic conditions have been interpreted in the context of models that state that interaural differences between the parameters of the waveforms presented to the two ears provide the basis for binaural detection (e.g., the vector model of Jeffress, 1972). For example, if we consider the detection of an interaurally out-of-phase signal (Sπ) in the presence of an interaurally in-phase noise (No), on noise-alone trials the interaural differences are constant and equal to zero, while on signal-plus-noise trials there are ongoing fluctuations in the interaural parameters (e.g., interaural time differences) during the time the signal is on (the "ongoing" cue). If the noise is on continuously, there is also a shift in interaural parameters, from diotic to dichotic, that occurs at the onset of the signal (the "onset" cue). The absence of the onset cue with pulsed noise maskers is used to account for the decrease in detectability relative to continuous noise maskers. Similarly, because the addition of a forward fringe reintroduces the onset cue, performance comparable to that with the continuous masker is predicted.

Yost (1985), following earlier authors (e.g., McFadden, 1966; Robinson and Trahiotis, 1972), suggested that the presence of a forward and/or backward fringe allows the binaural system to "establish a baseline" against which changes in interaural parameters resulting from the signal onset or offset can be detected. Because the binaural system responds slowly to changes in interaural parameters, detection is facilitated when the duration of the masker fringe increases, providing more time for estimation of these baseline parameters. Based on previous findings and a number of additional manipulations performed in his study, Yost concluded that the presence of fringe activity anywhere during an interval approximately 300 to 500 ms in duration immediately before signal onset can have an effect on the measured threshold. He further suggested that the length of this interval corresponds to the time required for accurate estimation of interaural parameters.

Yost also found that, when the interaural parameters of the fringe were different from those of the masker (e.g., an $N\pi$ fringe preceding an No masker), thresholds were the same as with a pulsed noise masker (i.e., the fringe had no effect on detectability). Yost invoked the equalization–cancellation (E-C) model of Durlach (1972) to explain why a forward fringe with the same interaural parameters as the masker yielded lower thresholds than a forward fringe with interaural parameters different from those of the masker. He argued that, when an No fringe was presented before an No masker, the equalization stage of the model would adopt a transformation to cancel the masker fringe. Such a strategy would also be optimal for detecting the signal, canceling the masker and doubling the signal. Because it takes some time for the equalization stage to adopt the correct transformation, the longer the fringe (up to 500 ms), the lower the measured threshold. When an $N\pi$ fringe preceded an No masker, the equalization transformation would again be chosen to cancel the fringe. Because it takes time for the equalization stage to adopt a new transformation, the No masker would not be canceled, and the $S\pi$ signal, if present, would be reduced. Thus, within this interpretation, we would expect performance to be much worse under the $N\pi$ forward fringe condition than for a continuous No masker condition or for an No forward fringe condition.

Yost's interpretation does not specifically address the pulsed masker condition. However, it seems unlikely that a quiet period preceding an No masker would induce the equalization stage to adopt such a nonoptimal equalization transformation. Thus we would expect the threshold for the pulsed noise condition to be lower than the threshold for the $N\pi$ forward fringe condition. This prediction conflicts with Yost's findings.

Kollmeier and Gilkey (1990) were also interested in the temporal properties of the binaural system. However, their approach, at least on the surface, was somewhat different. They wished to compare the time course of temporal masking under monaural and binaural conditions. Kollmeier and Gilkey noted that, in previous studies where the masker was simply switched on and off, the results were difficult to interpret, because the resultant change in the level of the effective masker influenced not only the binaural channel, but also the monaural channel. They therefore measured the detectability of a brief 25-ms $S\pi$ signal[2] in the presence of a 750-ms noise masker whose interaural phase was switched from $N\pi$ to No after 375 ms, as a function of the delay between the transition in the noise and the onset of the signal (Δt). This masker configuration produces a change in the effective level of the masker at the phase transition within the binaural channel, but not within the monaural channels. The curve in Fig. 1 shows the results for this $N\pi$–No masker configuration. There is a gradual transition in the amount of masking, reminiscent of a temporal masking function. Note, however, that, when $\Delta t = 0$ (the point marked by the *F*) this temporal masking configuration is quite similar to some of the masker fringe conditions investigated by previous researchers. That is, there is both an $N\pi$ forward fringe and an No backward fringe. The C at $\Delta t = 0$ represents threshold with a "pseudo-continuous" No noise (i.e., a 750-ms No masker with no phase transition). Thus the two points indicated by the F and the C are very similar to the forward fringe condition and the continuous masker condition investigated by Yost. The 10-dB difference between the $N\pi$–No masker configuration (point F)and the continuous masker configuration (point C) at $\Delta t = 0$ is comparable to the difference obtained between pulsed and continuous maskers under dichotic conditions in other studies (e.g., Yost, 1985; Robinson and Trahiotis, 1972). Thus it might be suggested that the threshold would not change if the $N\pi$ segment of the $N\pi$–No masker were deleted, a conclusion in agreement with Yost's finding that the $N\pi$ forward fringe has no effect on the measured threshold.

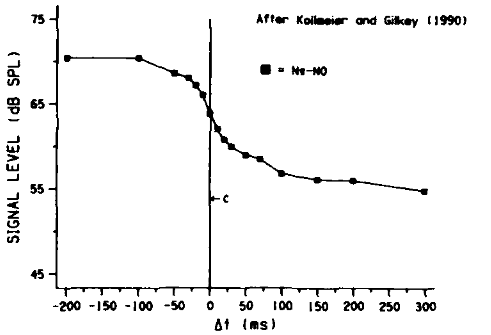

FIG. 1. Threshold level of a 20-ms $S\pi$ signal as estimated by a 3-down 1-up (79.4 percent correct) adaptive staircase procedure plotted as a function of Δt, the delay between the phase transition in the noise and the onset of the signal, for an $N\pi$-No masker configuration. The C shows the threshold in the center of a 750-ms No masker. The vertical line corresponds to a signal onset temporally aligned with the transition between the first and second segment of the noise. The data shown are the average thresholds of four subjects, based on the results of Kollmeier and Gilkey (1990).

A more careful consideration of the Kollmeier and Gilkey data indicates that the $N\pi$ segment of the noise is having a substantial effect. First, it should be recalled that, in previous studies, the presence of an No backward fringe had been found to improve detectability over the pulsed masker condition (e.g., Trahiotis *et al.*, 1972). When the duration of the backward fringe is similar to that present at point F, the detectability of the signal should only be about 3 dB worse than in the continuous masker condition. Thus, if the $N\pi$ segment of the noise really had no effect in the study of Kollmeier and Gilkey, we would expect the difference between point F and point C to have been around 3 dB, rather than the approximately 10 dB observed in Fig. 1. Second, the overall shape of the threshold function in Fig. 1, including the continued improvement in detectability as the signal is moved farther away from the $N\pi$ segment of the noise and into the No segment of the noise, was interpreted by Kollmeier and Gilkey (1990) to indicate that the $N\pi$ segment acts like a forward masker. (That is, its effects last long after its offset.)

Kollmeier and Gilkey also interpreted their results within the context of the E-C model, but assumed that the equalization stage adopts a strategy that is optimal to cancel the No segment of the masker. Therefore, the Nπ segment is effectively doubled by the cancellation stage, such that the level of the masker at the output of the cancellation stage is quite large during the Nπ segment, compared to the level during the No segment. If activity in the binaural channel does not decay instantaneously (e.g., see Grantham and Wightman, 1979), activity associated with the Nπ segment of the noise will continue even after the phase transition in the noise. Kollmeier and Gilkey argued that the function shown in Fig. 1 describes the time course of this decay of activity. Although Kollmeier and Gilkey did not make specific predictions for the pulsed masker condition, it can be seen that, while the masking functions for an No masker with and without an Nπ forward fringe might intersect at $\Delta t = 0$, presumably they are quite different functions elsewhere.

Kollmeier (1986) directly replicated Yost, measuring the detectability of a brief Sπ signal in a brief No masker. He reported the difference between the pulsed masker and continuous masker conditions to be about 3.5 dB and the difference between the Nπ forward fringe and pulsed masker conditions to be almost 10 dB. These results also suggest that the Nπ forward fringe is having a substantial effect.[3]

To resolve these apparent differences between the studies of Yost and of Kollmeier and Gilkey (1990) and Kollmeier (1986), a series of experiments was conducted to reexamine the effects of forward and backward masker fringe on dichotic tone-in-noise masking.

I. GENERAL METHODS

Both signal and noise stimuli were output at a 20-kHz sampling rate through separate 16-bit digital-to-analog converters and passed through 7.8-kHz low-pass antialiasing filters. The signal and noise stimuli were added with an analog mixer and presented via TDH-49 headphones mounted in circumaural cushions (Grason–Stadler model 001A) to subjects in individual sound-attenuating booths. Stimulus generation and presentation and response collection were controlled by an SMS-1000 minicomputer with a PDP 11/73 processor.

The masking stimulus was pseudorandom noise generated with a 33-bit software shift register (Gilkey *et al.*, 1988) and bandpass filtered between 100 and 3000 Hz by a Krohn-Hite filter (model 3750) with slopes set to 24 dB/oct. The noise was turned on and off essentially instantaneously. Thus the rise/fall time was determined by the filters and the headphones. The spectrum level of the noise was 40 dB SPL/Hz. The phase of the noise in the left ear was held constant, while the phase of the noise in the right ear was changed by 180° for Nπ presentations. It should be noted that the generated noise can be at either of only two instantaneous amplitudes. Therefore, instantaneous changes in the interaural phase of the noise do not introduce transients into the monaural waveform.

The signal was a computer-generated 500-Hz sinusoid with a total duration of 20 ms (measured from the beginning of the rise to the end of the fall) and was shaped with 5-ms linear rise/fall ramps. It was always presented 180° out of phase interaurally (Sπ). The level of the signal was controlled by separate programmable attenuators for each subject.

Figure 2 shows a schematic diagram of the stimulus conditions examined in these experiments. The noise masker can be thought of as being divided into three segments. The first and last segments had a duration of 350 ms, while the middle segment had a duration of 20 ms. The middle segment always contained No noise. The first and third segments could contain No noise, Nπ noise, or quiet. Threshold for detection of the signal was measured as a function of the delay (Δt) from the beginning of the middle segment of the noise to the onset of the signal. For the particular case when $\Delta t = 0$, the successive panels of this figure show stimulus configurations analogous to: (a) "pseudo-continuous" masker; (b) pulsed masker; (c) Nπ forward fringe; (d) No backward fringe; (e) Nπ forward fringe and No backward fringe; (f) Nπ backward fringe; and (g) Nπ forward fringe and Nπ backward fringe conditions. For convenience, we often refer to the first segment of the noise masker as the "forward fringe" and the third segment of the noise masker as the "backward fringe." These terms are correct when $\Delta t = 0$, but something of a misnomer when $\Delta t \neq 0$. Note that

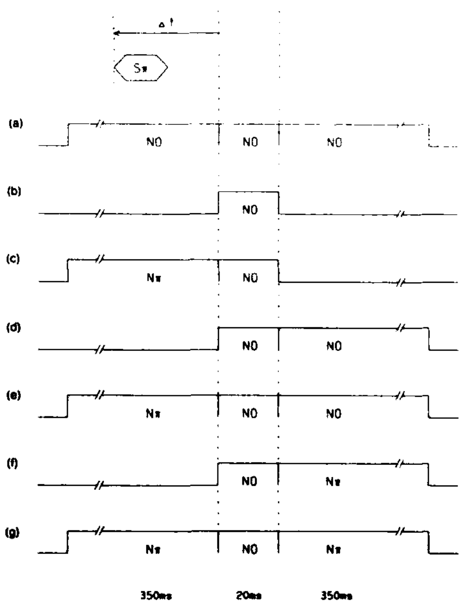

FIG. 2. Schematic diagram showing the seven masker configurations investigated in this paper: (a) pseudo-continuous masker; (b) pulsed masker; (c) Nπ forward fringe; (d) No backward fringe; (e) Nπ forward fringe and No backward fringe; (f) Nπ backward fringe; and (g) Nπ forward fringe and Nπ backward fringe conditions. See text for details.

the stimulus configurations shown in panels (a)–(c) are analogous to those studied by Yost, and the stimulus configurations shown in panels (a) and (e) are analogous to those investigated by Kollmeier and Gilkey. The specific conditions tested in each experiment are described with those experiments.

II. EXPERIMENT 1—Nπ FORWARD FRINGE

The first experiment was performed to determine if the masking functions for the pulsed masker and Nπ forward fringe conditions did indeed intersect when $\Delta t = 0$, or if the presence of the Nπ forward fringe caused a further elevation of thresholds above those observed for pulsed maskers.

A. Methods

1. Subjects

Four subjects, two males and two females between the ages of 19 and 21, were paid for their participation. All had clinically normal hearing as measured by audiometric testing. Subjects received extensive training before data collection began.

2. Procedure

A two-alternative, forced-choice (2AFC) procedure was used to obtain thresholds. The beginning of each trial was marked by a 60-ms warning light. A second light with a duration of 60 ms was turned on 760 ms after the offset of the warning light to mark the first observation interval, and 760 ms after the offset of this light a third light was turned on for 60 ms, marking the second observation interval. The signal, if present, was turned on 20 ms after the light marking the appropriate observation interval. The 760-ms interval between the observation intervals was set such that the offset of the masker in the first interval and the onset of the masker in the second interval would be separated by at least 100 ms. Subjects pressed a button to indicate the interval containing the signal. Trials were self-paced, and subjects received visual feedback on a trial-by-trial basis.

The typical experimental session consisted of four sets of four blocks. Masker configuration and Δt, the interval from the beginning of the second segment of the noise to the onset of the signal, were randomly selected across blocks, but were not varied within blocks. Thresholds for each block were estimated by a 3-down, 1-up adaptive staircase procedure (Levitt, 1971), corresponding to a target percent correct of 79.4. Initial signal levels were set at approximately 15 dB above the anticipated threshold and varied in 4-dB steps for the first two reversals, 2-dB steps for the next two reversals, and 1-dB steps thereafter. A block was terminated after all subjects had completed a minimum of ten reversals. Thresholds were estimated by discarding the first four reversals and averaging the signal level at the remaining reversals, up to the highest even number of reversals.

Three stimulus configurations were tested, corresponding to the first three configurations shown in Fig. 2, panels (a)–(c): No–No–No masker configuration, QUIET–No–QUIET masker configuration, and Nπ–No–QUIET masker configuration. For the QUIET–No–QUIET condition and the Nπ–No–QUIET masker configuration, Δt was varied from -185 to 185 ms. For the No–No–No masker configuration, Δt was always equal to 0. Between 8 and 11 blocks of trials were conducted for each value of Δt with each masker configuration.

B. Results and discussion

Figure 3 shows the threshold signal level for each subject under each of the three masker configurations for $\Delta t = 0$, corresponding to "pseudo-continuous" masker, pulsed masker, and Nπ forward fringe conditions. Consistent with previous findings, thresholds for the pulsed masker condition are 3.7–12.0 dB higher than thresholds in the pseudo-continuous masker condition. However, in contrast to the findings of Yost (1985), thresholds for the Nπ forward fringe condition are 2.3–10.5 dB higher than thresholds for the pulsed masker condition, indicating that the Nπ forward fringe has a substantial effect on detectability. These results are similar to those reported by Kollmeier (1986).

Thresholds as a function of Δt for the Nπ–No–QUIET and for QUIET–No–QUIET masker configurations are shown in Fig. 4. Each data point represents the average of four subjects. The C indicates the pseudo-continuous masker condition at $\Delta t = 0$. It can be seen, as suggested in Fig. 3, that the functions do not intersect at $\Delta t = 0$. Further, these functions do not overlap at any point, and it is clear that the effects of the Nπ forward fringe extend well beyond its offset.

The results of experiment 1 failed to replicate those of Yost (1985) and suggest instead that the presence of an Nπ forward fringe does reduce detectability. Further, the fact that the effects of the Nπ forward fringe continue long after its offset is compatible with the interpretation of Kollmeier and Gilkey (1990) that the Nπ segment of the noise acts as a forward masker. Note that in their study an No backward

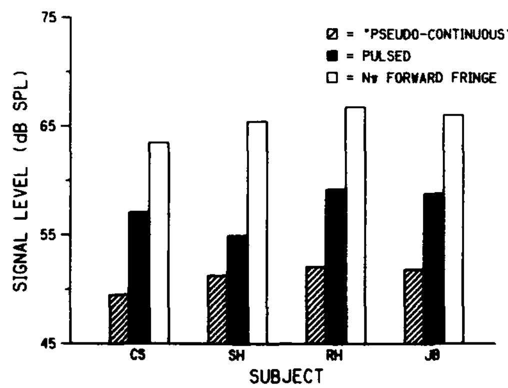

FIG. 3. Threshold level of a 20-ms Sπ signal whose onset is simultaneous with the beginning of the second segment of the noise for each of four subjects in three different masker conditions, as estimated by a 3-down 1-up (79.4 percent correct) adaptive staircase procedure. The pseudo-continuous masker is a 720-ms No noise, the pulsed masker is a 20-ms No noise, and the Nπ forward fringe masker is a 20-ms No noise preceded by a 350-ms Nπ forward fringe.

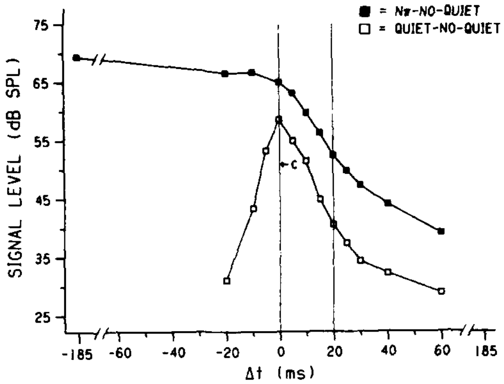

FIG. 4. Threshold level of a 20-ms Sπ signal, as estimated by a 3-down 1-up (79.4 percent correct) adaptive staircase procedure, is plotted as a function of Δt, the time from the beginning of the second segment of the noise to the onset of the signal for the Nπ-No-QUIET masker configuration and the QUIET-No-QUIET masker configuration. The C shows the threshold in the pseudo-continuous masker condition at $\Delta t = 0$. The two vertical lines correspond to signals whose onsets are simultaneous with the beginning and end of the middle segment of the noise. The data shown are averaged across four subjects.

fringe was also present, and may have influenced the obtained results. The next experiment provides a more systematic investigation of the effects of a backward masker fringe.

III. EXPERIMENT 2—BACKWARD FRINGE

In experiment 2, the Nπ forward fringe, pulsed masker, and continuous masker conditions were examined again for a different set of subjects and compared with several additional conditions, including the QUIET–No–No masker configuration [Fig. 2, panel (d)], which at $\Delta t = 0$ is an No backward fringe condition, and Nπ–No–No masker configuration [Fig. 2, panel (e)], a condition similar to the No–Nπ masker configuration of Kollmeier and Gilkey (1990). As discussed in the Introduction, the presence of an No backward fringe has also been shown to increase the detectability of the signal relative to the pulsed masker condition (Trahiotis *et al.*, 1972). If so, under the QUIET–No–No condition the facilitative influence of the No backward fringe should produce a threshold at $\Delta t = 0$ that is close to the value for the continuous masker condition. Further, if the Nπ forward fringe has no effect on detectability, the threshold for the Nπ–No–No masker configuration at $\Delta t = 0$ should be the same as that for the QUIET–No–No masker configuration. On the other hand, if the Nπ forward fringe is acting as a forward masker, the thresholds for Nπ–No–No masker configuration should be above those for the QUIET–No–No masker configuration at $\Delta t = 0$, and for other values of Δt as well.

The effect on detectability of an Nπ backward fringe was also measured in two additional conditions: the QUIET–No–Nπ masker configuration [Fig. 2, panel (f)], which at $\Delta t = 0$ contains a backward fringe whose interaural parameters differ from the preceding noise; and the Nπ–No–Nπ masker configuration [Fig. 2, panel (g)], which includes both a forward and backward Nπ fringe. Given that an No forward fringe or an No backward fringe increases detectability, but an Nπ forward fringe decreases detectability, it might be expected that an Nπ backward fringe would also decrease detectability.

A. Method

1. Subjects

Three subjects (two males and one female, all 22 years of age) were paid for their participation. None had participated in experiment 1. All had clinically normal hearing as measured by audiometric testing. Subjects received extensive training on the task before data collection began.

2. Procedure

General procedural details are similar to those of experiment 1. All of the conditions shown in Fig. 2 were examined in this experiment. The No–No–No masker configurations [Fig. 2(a)], QUIET–No–QUIET masker configurations [Fig. 2(b)], and Nπ–No–QUIET masker configurations [Fig. 2(c)] were tested only at $\Delta t = 0$; the remaining masker configurations were tested at values of Δt between -185 and 185 ms. Between four and ten blocks of trials were run for each masker configuration at each value of Δt.

B. Results and discussion

Figure 5 plots threshold signal levels as a function of Δt, averaged across the three subjects for the Nπ–No–No and QUIET–No–No masker configurations. The C, P, and π plotted at $\Delta t = 0$ represent thresholds for the pseudo-continuous masker condition, pulsed masker condition, and Nπ forward fringe condition, respectively. Thresholds for the

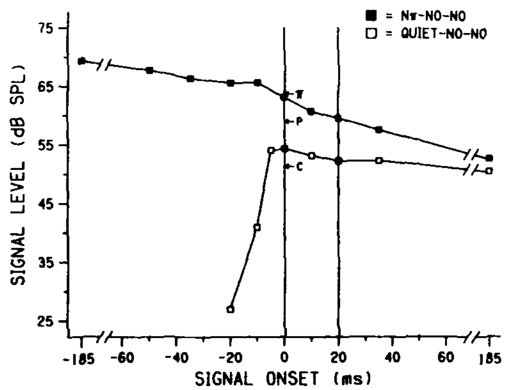

FIG. 5. Threshold level of a 20-ms Sπ signal is plotted as a function of Δt for the QUIET–No–No masker configuration and the Nπ–No–No masker configuration. The C, P, and π show the thresholds in the pseudo-continuous masker, pulsed masker, and Nπ forward fringe conditions, respectively, at $\Delta t = 0$. The data shown are averaged across three subjects. Other details are as in Fig. 4.

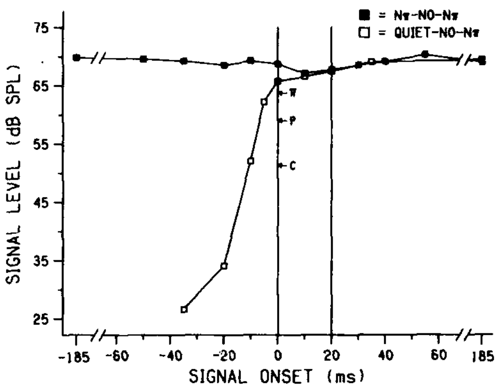

FIG. 6. Threshold level of a 20-ms $S\pi$ signal is plotted as a function of Δt for the QUIET–No–$N\pi$ masker configuration and the $N\pi$–No–$N\pi$ masker configuration. The C, P, and π show the thresholds in the pseudo-continuous masker pulsed masker, and $N\pi$ forward fringe conditions, respectively, at $\Delta t = 0$. The data shown are averaged across three subjects. Other details are as in Fig. 4.

pseudo-continuous masker and pulsed masker conditions differ by approximately 7.7 dB, consistent with previous findings and with the results of experiment 1. Thresholds for the pulsed masker and $N\pi$ forward fringe conditions differ by about 4.7 dB, also in agreement with experiment 1. In addition, it can be seen that threshold for the QUIET–No–No masker configuration at $\Delta t = 0$ (No backward fringe condition) is within 3.0 dB of the value for the pseudo-continuous masker, replicating the findings of Trahiotis *et al.* (1972).

The most important aspect of the results shown in Fig. 5 is the fact that there is no overlap between the functions for the QUIET–No–No and $N\pi$–No–No masker configurations. The effect of the $N\pi$ forward fringe raises thresholds well above those for the QUIET–No–No masker configuration for all values of Δt. At $\Delta t = 0$, the two functions differ by approximately 8.8 dB. Further, the function for the $N\pi$–No–No masker configuration does not approach the function for the QUIET–No–No masker configuration until the signal is well into the No segment of the noise.

In Fig. 6, the thresholds for the QUIET–No–$N\pi$ and the $N\pi$–No–$N\pi$ masker configurations are plotted as a function of Δt. If the $N\pi$ backward fringe had no effect on detectability, it would be expected that the value for the QUIET–No–$N\pi$ masker configuration at $\Delta t = 0$ would equal that for the pulsed masker condition. However, the $N\pi$ backward fringe raises thresholds above the pulsed condition by approximately 6.7 dB. By comparing the functions for the $N\pi$–No–$N\pi$ masker configuration and the QUIET–No–$N\pi$ masker configuration, it can be seen that there is a detrimental effect of the $N\pi$ forward fringe when Δt is equal to zero (3.0 dB). For values of Δt greater than zero the two curves are essentially identical. The limited effect of the $N\pi$ forward fringe for positive values of Δt is probably a ceiling effect (i.e., thresholds are approximately equal to the expected monaural threshold).

The results of experiments 1 and 2 indicate that the introduction of an $N\pi$ forward fringe before an No masker elevates thresholds above those for the pulsed masker condition. This finding is in conflict with Yost's (1985) result that the $N\pi$ fringe had no effect. The remaining three experiments were undertaken in an attempt to reconcile the discrepancies between the results of the present study and those of Yost.

IV. EXPERIMENT 3—OVERTRAINING

Subjects may have become confused by the number of experimental conditions under which they were tested in experiments 1 and 2 (Δt and the masker configuration changed randomly from block to block), and may have performed more poorly under some conditions than they might have if fewer conditions had been tested. It is also possible that they may have found the phase transition in the noise phenomenologically distracting (the perceptual effect of this shift in phase is a movement of the auditory image within the head), and that this distraction affected performance. In experiment 3, the possibility that these factors were influencing the results was investigated by "overtraining" subjects on only a few conditions. That is, if the results observed in experiments 1 and 2 were due to confusion effects or attention effects, it should be possible to reduce these problems by overtraining the subjects and limiting the number of conditions.

A. Method

1. Subjects

The subjects from experiment 1 participated in experiment 3. All subjects received 22–24 blocks of additional training in the specific experimental conditions tested in experiment 3 before data collection began.

2. Procedure

General procedural details were similar to those of experiments 1 and 2. The pulsed masker condition and $N\pi$ forward fringe condition were investigated with $\Delta t = 0$ for all stimulus presentations. The experimental condition was held constant within sets of four blocks and alternated between sets of four blocks. Between 28 and 35 blocks of trials were obtained under each configuration.

B. Results and discussion

The top two panels of Table I show the average difference in threshold between the pulsed masker and the $N\pi$ forward fringe conditions for each subject in experiments 1 and 3. As can be seen, the average values for the two experiments differ by only about 1.0 dB, indicating that, for most subjects, exposure to a large number of values of Δt did not substantially affect the results of experiment 1.

V. EXPERIMENT 4—SINGLE-INTERVAL TASK

Yost (1985) assumes that, when the interaural parameters of the fringe are the same as those of the masker, the binaural system uses the fringe in order to estimate the interaural parameters of the masker. He further assumes that the

TABLE I. Differences in threshold between pulsed masker and Nπ forward fringe masker conditions at $\Delta t = 0$, by subject.

Subject	Experiment 1 multiple delays Difference (pulsed-Nπ fringe)
CS	6.4 dB
SH	10.5
RH	7.6
JB	2.3
Average	6.7
Subject	Experiment 3 overtraining at $\Delta t = 0$ Difference (pulsed-Nπ fringe)
CS	5.9 dB
SH	7.5
RH	6.1
JB	3.1
Average	5.7
Subject	Experiment 4 single-interval task Difference (pulsed-Nπ fringe)
CS	8.2 dB
SH	6.6
RH	5.9
JB	3.5
Average	6.1

binaural system takes a considerable period of time in order to estimate the interaural parameters of a waveform. A "fringe" waveform present in the approximately 500-ms interval before the onset of the signal can influence the obtained thresholds. Therefore, it may be inappropriate to estimate dichotic thresholds using techniques such as multiple interval forced-choice procedures. That is, the masker waveform in one interval may act as a fringe to the masker in the other interval unless the time between the observation intervals is quite large. Indeed, under the NoSπ condition, McFadden (1966) observed a reduction of the difference between pulsed and continuous maskers when a 2AFC procedure was used rather than a single-interval procedure. In his study, Yost used a single-interval, yes/no procedure, whereas experiments 1–3 of the present study employed a 2AFC procedure. Even though the time between the two observation intervals in our procedure was relatively large (760 ms), it might be argued that the stimulus in one interval may have influenced the processing of the stimulus in the other interval and thus affected the results (e.g., a 500-ms pause between the observation intervals was not sufficient to make 2AFC and single-interval procedures equivalent in the experiment of McFadden, 1966). In experiment 4, the pulsed masker and Nπ forward fringe conditions were reexamined using a single-interval, yes/no procedure.

A. Method

1. Subjects

The four subjects from experiments 1 and 3 participated in experiment 4.

2. Procedure

General procedural details were similar to those of experiments 1–3, except as described below. Thresholds were measured with a single-interval, yes/no procedure. The signal level was fixed for all trials within a block. The beginning of each trial was marked with a 60-ms warning light, followed after 760 ms by an observation interval marked by a 60-ms light, which was turned on 20 ms prior to the onset of the signal, if present. As before, trials were self-paced, and visual feedback was provided on a trial-by-trial basis.

Experimental sessions were arranged in sets of three 100-trial blocks. Each set was preceded by a 20-trial practice block. Within each set of blocks, the experimental condition was held constant and three signal levels were tested on successive blocks. The signal levels differed in 2.5-dB steps, with the middle value corresponding to approximately 75% correct performance for that subject. Values of d' were calculated for each subject for each block and averaged across blocks. Three-point psychometric functions were fit as straight lines to the logarithmic transform of

$$d' = m(E/N_0)^k, \tag{1}$$

as described by Egan *et al.* (1969). Threshold was estimated as the level corresponding to $d' = 1.16$ $[P(C)_{2AFC} = 0.794]$.

The two conditions tested in experiment 3, QUIET–No–QUIET masker configuration [Fig. 2(b)] and Nπ–No–QUIET configuration [Fig. 2(c)] were investigated with $\Delta t = 0$ (the pulsed masker and Nπ forward fringe conditions). At least six blocks of trials at each of the three signal levels were obtained for each condition.[4]

B. Results and discussion

Table I shows the difference in threshold between the pulsed masker and Nπ forward fringe conditions for each subject. The middle panel shows values from experiment 3 (2AFC procedure), and the bottom panel shows values from experiment 4 (single-interval procedure). The average difference between the conditions with the 2AFC procedure in experiment 3 was 5.7 dB, whereas the average difference between the conditions with the single-interval procedure in experiment 4 was 6.1 dB. While this slight effect of the psychophysical procedure is in the same direction as the 1.9 dB effect reported by McFadden, it seems safe to conclude that the use of a 2AFC procedure in experiments 1–3 did not significantly influence the results.

VI. EXPERIMENT 5—SIGNAL DURATION

In Yost's (1985) study, the signal duration was 30 ms, as measured from the beginning of the rise to the end of the fall, whereas in experiments 1–4 the signal duration was 20 ms. Robinson and Trahiotis (1972), in measuring the influence of signal duration on forward fringe effects, found considerable differences in results obtained for short (37-ms) and long (261-ms) signals.[5] Specifically, they found a much larger difference between continuous and pulsed maskers under dichotic conditions when the signal duration was short. The seemingly related "overshoot effect" is also dependent on signal duration (e.g., Zwicker, 1965; Fastl, 1976). However, the overshoot effect when measured for dichotic conditions is usually quite small (e.g., Trahiotis *et al.*, 1972; McFadden, 1988). Nevertheless, it is possible that the shorter duration of the signal used in the present study led to a larger difference between pulsed masker and $N\pi$ forward fringe conditions than would have been found with a longer duration signal. This possibility was examined more closely in experiment 5.

A. Method

1. Subjects

Four subjects (two males and two females between the ages of 19 and 22 years) were paid for their participation in the experiment. All had clinically normal hearing as measured by audiometric testing. Two subjects (JB and CS) had participated in experiments 1, 3, and 4. All subjects were extensively trained on the experimental task before data collection began.

2. Procedure

Procedural details were the same as in experiment 4, with the exception that signal durations of both 20 and 30 ms, as measured from the beginning of the rise to the end of the fall, were used. The QUIET–No–QUIET and $N\pi$–No-QUIET masker configurations with $\Delta t = 0$ were tested (i.e., pulsed masker and $N\pi$ forward fringe conditions).[5] Between 13 and 27 blocks of trials for each signal duration and level under each condition were obtained for each subject.[4]

B. Results and discussion

Table II shows the difference between the pulsed masker and $N\pi$ forward fringe conditions for each of the four subjects. As can be seen, there is little difference in the size of the effect obtained with the 20- and 30-ms signals, with an average difference for the 30-ms signal of 3.1 dB and an average difference for the 20-ms signal of 4.2 dB.

TABLE II. Differences in threshold between pulsed masker and $N\pi$ forward fringe masker conditions at $\Delta t = 0$, by subject.

Subject	Experiment 5 30-ms signals Difference (pulsed-$N\pi$ fringe)
CS	3.7 dB
KR	3.1
RQ	2.6
JB	2.9
Average	3.1
Subject	Experiment 5 20-ms signals Difference (pulsed-$N\pi$ fringe)
CS	6.6 dB
KR	3.2
RQ	4.5
JB	2.4
Average	4.2

Although the size of the effect is somewhat smaller for the longer duration signal, the presence of the $N\pi$ forward fringe still appreciably reduces detectability for all four subjects. Thus it seems unlikely that the use of a 20-ms rather than 30-ms signal in experiments 1–4 substantially altered the results.

VII. GENERAL DISCUSSION

Yost (1985) found that, if the interaural parameters of the forward fringe differed from the interaural parameters of the portion of the masker that overlapped with the signal, the presence of the fringe had no effect on detectability. In contrast, Kollmeier and Gilkey (1990) argued that an $N\pi$ segment of noise preceding an No masker acted like a forward masker.

The results of experiment 1 of the present study failed to replicate those of Yost (1985). That is, when Δt is equal to zero, thresholds in the $N\pi$ forward fringe condition are 6.7 dB higher than in the pulsed masker condition. Moreover, thresholds for the $N\pi$–No–QUIET masker configuration were above those for the QUIET–No–QUIET masker configuration for all values of Δt, demonstrating that the $N\pi$ segment of the noise has effects that last long after its offset. In experiment 2, these effects were reexamined in conditions where an No backward fringe was also present. Again, thresholds were about 8.8 dB higher when the $N\pi$ forward fringe was added. Also, the function that related threshold to Δt for the $N\pi$–No–No masker configuration was always above that for the QUIET–No–No masker configuration, again suggesting that the $N\pi$ segment of the noise has a long-lasting effect. Further, the shape of the $N\pi$–No–No function resembled that of a forward masking function and was similar to the function obtained by Kollmeier and Gilkey (1990) for a similar condition. The addition of an $N\pi$ backward fringe was also shown to raise threshold at $\Delta t = 0$ over that for the pulsed masker condition. Experiments 3–5 indicate that the difference between the results of the present study and those of Yost cannot be readily explained by training effects, the psychophysical procedures employed, or the duration of the signal.[6]

Based on the previous literature, it appears that, under dichotic presentations, the addition of a forward or backward masker fringe that has the same interaural parameters as those of the masker will enhance the detectability of the signal. On the other hand, the results of the present study, as well as those of Kollmeier and Gilkey (1990) and Kollmeier (1986), indicate that the addition of a forward or backward fringe whose interaural parameters are different from those of the masker will decrease the detectability of the signal.

There are at least three possible approaches to explaining these results. First, a phase transition in the noise produces a distinct sensation, which, to a first approximation, can be described as "movement" of the auditory image. Perhaps the subject's ability to detect the signal when it is close to the transition in the noise is hampered because he is distracted by this change in the auditory image. As the signal is moved away from the transition in the noise, the subject is better able to focus his attention on the signal. One problem with this argument is that it provides no specific explanation for the difference between the pulsed masker and continuous masker conditions, except perhaps to assume that the onset of the masker is also distracting, although less distracting than the phase transition in the noise. A second problem with this argument is that we might expect the distracting effects of the transition in the noise to raise thresholds during the $N\pi$ segment of the masker as well. However, instead of getting worse, thresholds actually seem to improve slightly as the signal is brought closer to the transition in the noise within the $N\pi$ segment. Indeed, when the signal offset and the transition are simultaneous, thresholds appear to be at least slightly lower than when the signal is well within the $N\pi$ segment of the noise.

The second explanation assumes, as did Kollmeier and Gilkey, that the $N\pi$ forward fringe acts as a forward masker. If we assume an E-C model (Durlach, 1972) that selects an equalization transformation that maximizes the $S\pi$ signal at the output of the cancellation stage, then the output will be large during the $N\pi$ segment of the noise (i.e., the noise will be doubled in amplitude) and small during the No segment of the noise (i.e., the noise will be canceled to the limit determined by the internal noise). Further, if activity in the binaural channel decays slowly, then activity associated with the $N\pi$ segment of the noise might continue after the transition in the noise and act as a forward masker. Thus the gradual increase in detectability as the signal is moved from the $N\pi$ segment to the No segment in Fig. 5 might represent the time course of temporal masking. One problem with this interpretation is that it does not specifically address the difference between the No forward fringe condition and the pulsed masker condition, but only states that the $N\pi$ forward fringe condition should be worse than either.

The third explanation assumes, as did McFadden (1966), Robinson and Trahiotis (1972), and Yost (1985), that the masker fringe provides a reference or baseline against which the signal is detected. For example, in an $NoS\pi$ detection task an No forward fringe provides a reference of diotic cues. On noise-alone trials, the stimulus remains diotic throughout, while on signal-plus-noise trials there are ongoing dichotic cues during the time the signal is on (the ongoing cue), and there is a transition in interaural parameters from diotic to dichotic at the signal onset (the onset cue). When no fringe is present (the pulsed masker condition), the onset cue is not available, and the subject must rely on the ongoing cue alone. An $N\pi$ fringe, on the other hand, would provide a reference of dichotic cues. On noise-alone trials, the transition would be from dichotic interaural cues to diotic interaural cues, and on signal-plus-noise trials the transition would be from dichotic interaural cues to dichotic interaural cues of reduced average magnitude. It could well be argued that this condition would provide a less effective onset cue than is provided in the No forward fringe condition, but probably no less effective than the nonexistent onset cue in the pulsed masker condition. However, while in the No forward fringe condition and the pulsed masker condition there are never dichotic cues on noise-alone trials, under the $N\pi$ forward fringe condition there would be interaural difference cues present on both noise-alone and signal-plus-noise trials. Thus it could be argued that the ongoing cue is less reliable than under either the No forward fringe condition or the pulsed masker condition. If so, one might expect performance to be best for the No forward fringe condition (both onset and ongoing cues), and worst for the $N\pi$ forward fringe condition (poor onset cue and poor ongoing cue). Thus this explanation is consistent with the pattern of results observed here.

As suggested by Yost, a similar argument can be made in the context of the E-C model. If we assume that the equalization stage of the model adopts an equalization transformation that attempts to minimize the noise at each instant in time,[7] then under the No forward fringe condition, the subject would choose an equalization transformation that maximally canceled the No fringe. Given that this equalization transformation is also optimal for detecting the signal, performance will be good. Under the $N\pi$ forward fringe condition, the subject would choose an equalization transformation that maximally cancels the $N\pi$ fringe. Such a transformation would double the No portion of the masker and cancel the signal, a very ineffective strategy for detecting the signal. If we assume that the system can change its equalization transformation only slowly, then this transformation will still be in effect during the No segment of the noise, and performance will be poor. Although it is unclear what equalization transformation the model would adopt in the quiet before the onset of a pulsed masker, it is not unreasonable to assume that the system adopts some random transformation, which, on average, is more effective than the transformation adopted in the $N\pi$ forward fringe condition, but less effective than the transformation adopted in the No forward fringe condition. Thus this argument predicts the ordering of the three conditions, if not the quantitative details. Within this view, the gradual decrease in threshold after the $N\pi$ segment of the noise indicates the time required to establish the correct equalization transformation.

In summary, it is not clear why the results obtained in the present study differ from those of Yost (1985). The elevation of thresholds resulting from addition of an $N\pi$ fringe appears to be robust (all nine subjects who participated in the experiments reported here show the effect). However, it should be noted that there was some intersubject variability in the size of the effect. Overall, investigations into the effects of adding forward and backward masker fringes suggest that the binaural system is sensitive to activity occurring during a considerable period before and after the presentation of a signal, and that any such activity may have long-lasting effects. None of the models presented here or in the literature provides a very quantitative description of the data. However, a model which assumes that the masker fringe provides a baseline or reference of interaural information is not inconsistent with the data presented here.

ACKNOWLEDGMENTS

The present study was supported by grants from the Air Force Office of Scientific Research (86-0298 and 89-0302), the National Science Foundation (BNS-85-11768 and BNS-87-20305), and from the National Institutes of Health (DC00306). The authors wish to thank Mervyn Pow, W. Michael Mudrovic, and Rafael Quiñónez for their assistance with this research and William Yost, Dennis McFadden, and Constantine Trahiotis for their many useful comments and suggestions on a earlier version of this manuscript.

[1] This result agrees with most later studies using moderate- to long-duration, narrow-band, low-frequency signals masked by broadband maskers (Tucker *et al.*, 1968; Robinson and Trahiotis, 1972; Wier *et al.*, 1977; Kohlrausch, 1986). However, substantial effects are sometimes observed when the signal duration is reduced, the masker bandwidth is narrowed, or the signal frequency is raised. These effects, when observed under diotic conditions, may be related to the overshoot effect (Zwicker, 1965; Fastl, 1976; Carlyon, 1987). However, while substantial differences exist between burst and continuous masking for dichotic conditions, dichotic overshoot is usually small (Trahiotis *et al.*, 1972; McFadden, 1988).

[2] The durations shown here are the total duration of the signal (i.e., from the beginning of the rise to the end of the fall), and thus are different from the original durations reported by the authors.

[3] Bell (1972) measured the detectability of an $S\pi$ signal in a masker that was switched from NU (interaurally uncorrelated noise) to No within a trial. The signal was always presented during the No portion of the noise, but the duration of the No fringe either before or after the signal was varied. He found that performance was worse when no No fringe was present and suggested that this case was comparable to the pulsed masker condition. However, because he did not actually measure thresholds under the pulsed masker condition, it is not possible to determine whether an *NU* forward fringe affects detectability.

[4] For a few blocks, at the highest signal levels, the obtained response matrices had empty cells. The data for these blocks were discarded.

[5] When the duration of the signal was 30 ms, the duration of the No segment of the noise was increased to 30 ms, and the duration of the $N\pi$ forward fringe was decreased to 345 ms.

[6] Recently, Yost (1988) has pointed out another potentially significant difference in the stimuli between the two studies. The transition between the fringe and masker segments of our generated noise (i.e., before filtering) was essentially instantaneous and produced no change in intensity within the monaural channels. In Yost's (1985) study the transition between the fringe and the masker was more gradual. The fringe ended with a 5-ms linear decay followed by an approximately 1-ms gap and then by the 5-ms linear onset of the masker or masker plus signal. Hence, any phase transition occurred over an 11-ms interval. In addition, there was a gap of 7 ms (if measured from the half power points) between the fringe and the masker. The exact implications of these stimulus differences are unclear, but there is at least some possibility that they may be responsible for the observed differences. For example, the gap may be significant in light of the work of Hafter and Buell (1983), indicating that a gap in a sequence of stimuli appeared to have the effect of "restarting" the binaural system. Thus it is at least possible that the gap between the fringe and the masker may have allowed the binaural analyzer to escape the influence of the $N\pi$ fringe. However, in this case it is not clear why there is still a facilitating effect of the No fringe.

[7] The data of McFadden (1967) would indicate that subjects can do something equivalent to adopting different equalization transformations on different trials. He investigated detectability under the NoSo condition and the $N\pi$So condition and found that subjects did no worse when the two types of trials were randomly mixed within the same block than when the two types of trials were presented in separate blocks. Unpublished data from our laboratory indicate that even when subjects are overtrained to listen for NoSo or $NoS\pi$ trials, their performance is not significantly hampered on unexpected probe trials containing the other condition. These results allow for the possibility that the subject can alter his processing strategies within a trial. On the other hand, a more reasonable interpretation might be to assume that there are multiple E-C channels, each tuned to a particular equalization transformation. In this case, the question of how rapidly the equalization transformation can be changed within a channel would be restated as how the binaural analyzer chooses the appropriate channel to monitor, and how rapidly that choice can be made.

Bell, D. W. (1972). "Effect of fringe on masking-level difference when gating from uncorrelated to correlated noise," J. Acoust. Soc. Am. 52, 525–529.

Carlyon, R. P. (1987). "A release for masking by continuous, random, notched noise," J. Acoust. Soc. Am. 81, 418–426.

Durlach, N. I. (1972). "Binaural signal detection: Equalization and cancellation theory," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. II, pp. 369–462.

Egan, J. P., Lindner, W. A., and McFadden, D. (1969). "Masking-level differences and the form of the psychometric function," Percept. Psychophys. 6, 209–215.

Fastl, H. (1976). "Temporal masking effects: I. Broad band noise masker," Acustica 35, 287–302.

Gilkey, R. H., Robinson, D. E., and Frank, A. S. (1988). "A software pseudorandom noise generator," J. Acoust. Soc. Am. 83, 829–831.

Grantham, D. W., and Wightman, F. L. (1979). "Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation," J. Acoust. Soc. Am. 65, 1509–1517.

Hafter, E. R., and Buell, T. N. (1983). "Recovery from saturation in lateralization of high-frequency stimuli," J. Acoust. Soc. Am. Suppl. 1 74, S40.

Jeffress, L. A. (1972). "Binaural signal detection: Vector theory," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. II, pp. 349–368.

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

Kollmeier, B. (1986). "Entwicklung zeitoptimierter psychoakustischer Messverfahren und ihre Anwendung auf binaurale Zeiteffekte," Doctoral dissertation, Georg-August-Universität, Göttingen, West Germany.

Kollmeier, B., and Gilkey, R. H. (1990). "Binaural forward and backward masking: Evidence for sluggishness in binaural detection," J. Acoust. Soc. Am. 87, 1709–1719.

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

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

McFadden, D. (1967). "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.

McFadden, D. (1988). "Absence of overshoot in a dichotic masking condition," J. Acoust. Soc. Am. 83, 1685–1687.

Robinson, D. E., and Trahiotis, C. (1972). "Effects of signal duration and masker duration on detectability under diotic and dichotic listening conditions," Percept. Psychophys. 12, 333–334.

Trahiotis, C., Dolan, T. R., and Miller, T. H. (1972). "Effect of 'backward' masker fringe on the detectability of pulsed diotic and dichotic tonal signals," Percept. Psychophys. 12, 335–338.

Tucker, A., Williams, P. I., and Jeffress, L. A. (1968). "Effect of signal duration on detection for gated and for continuous noise," J. Acoust. Soc. Am. 44, 813–816.

Wier, C. C., Green, D. M., Hafter, E. R., and Burkhardt, S. (1977). "Detection of a tone burst in continuous- and gated-noise maskers; defects of signal frequency, duration, and masker level," J. Acoust. Soc. Am. 61, 1298–1300.

Yost, W. A. (1985). "Prior stimulation and the masking-level difference," J. Acoust. Soc. Am. 78, 901–907.

Yost, W. A. (1988). Personal communication.

Zwicker, E. (1965). "Temporal effects in simultaneous masking by white-noise bursts," J. Acoust. Soc. Am. 37, 653–663.