SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

REPORT DOCUMENTATION PAGE

DTIC ELECTE DEC 03 1990 S B D

AD-A229 743

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	
2a. SECURITY CLASSIFICATION AUTHORITY	Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 90 1164
6a. NAME OF PERFORMING ORGANIZATION: HEARING RESEARCH CENTER, DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF FLORIDA; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NL
6c. ADDRESS (City, State and ZIP Code): GAINESVILLE, FL 32611	7b. ADDRESS (City, State and ZIP Code): BUILDING 410, BOLLING AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-89-0021
8c. ADDRESS (City, State and ZIP Code): Bld 410, Bolling AFB, DC 20332	10. SOURCE OF FUNDING NOS.
11. TITLE (Include Security Classification): AUDITORY PATTERN MEMORY	

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

12. PERSONAL AUTHOR(S): ROBERT D. SORKIN

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
FINAL TECHNICAL	FROM 88-10-1 TO 90-9-30	90-10-31	52

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	AUDITORY PERCEPTION, AUDITORY SEQUENCE DISCRIMINATION, TEMPORAL PATTERN PERCEPTION

19. ABSTRACT

Three studies of temporal pattern perception were conducted. The listener's task was to determine whether or not two arrhythmic, tonal sequences formed the same temporal pattern. The first study tested the Pattern Correlation model. According to this model, the listener estimates the correlation between the pattern of time intervals marked by the tones in each sequence. Listener performance was characterized by an internal noise that is dependent on the average time between marker tones. The second study tested the effects of temporally compressing or expanding the stimuli. These transformations are common in speech and music; an important feature of temporal pattern perception is the ability to recognize patterns as similar, despite such transformations. An additional internal noise, proportional to the magnitude of the difference between the pattern transformations, was postulated to describe performance. In the third study, the stimuli were presented at different time delays and to separate earphone channels. At very short delays, waveform correlation is the likely comparison mechanism. At longer delays, pattern correlation is the probable mechanism, but this process is ineffective when the temporal patterns overlap.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. Genevieve Haddad	22b. TELEPHONE NUMBER (Include Area Code): (202) 767-5021; 22c. OFFICE SYMBOL: NL

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE: UNCLASSIFIED

AUDITORY PATTERN MEMORY

Mechanisms of Temporal Pattern Discrimination by Human Observers

Robert D. Sorkin
Hearing Research Center
Department of Psychology
University of Florida
Gainesville, Florida 32611

31 October 1990

Final Technical Report for Period 1 October 1988 - 30 September 1990

AFOSR-89-0021

Prepared for

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base, DC 20332-6448

PROJECT SUMMARY

How do listeners determine whether two tonal sequences have the same or different temporal patterns? Three studies of temporal pattern discrimination were conducted. Listener performance in these experiments was evaluated using a mathematical model of temporal pattern discrimination. Analyses of these experiments allow specification of the temporal pattern discrimination mechanisms employed by the human auditory system.

The first study consisted of experiments that tested how listeners discriminate between arrhythmic, tonal sequences approximately one half-second in duration. Performance was modeled by the Pattern Correlation Model. According to the model, the listener extracts a list of marker interonset times from each pattern, and then computes the correlation between the pattern of time intervals marked by the tones in each sequence; other information about the input waveforms (such as absolute timing or spectra) is discarded. The experiments tested how listener performance depended on basic parameters of the task, such as sequence correlation, and number, duration, and variability of pattern elements. Listener performance was consistent with the predictions of the Pattern Correlation model, but was limited by an internal time jitter or noise that was a function of the average intermarker interval.

The second study evaluated the human listener's ability to discriminate between word-length tonal sequences that were subjected to uniform temporal transformations, such as time compression and expansion. One of the most intriguing features of temporal pattern perception is the ability to recognize patterns as similar, despite such compression or expansion manipulations. Examples of such time normalization abound in speech and music perception and we are normally unaware of such temporal changes, even when they occur during relatively brief stimuli, such as words. The experiments in the second study also tested how well the Pattern Correlation model could predict the effects of time compression and expansion on listener performance. The model proved useful in describing performance in a variety of different conditions that employed multiplicative and additive time transformations. Listener performance dropped when one of the sequences was compressed or expanded in time. In order for the model to describe this performance, it was necessary to postulate an additional, internal noise component that was proportional to the magnitude of the difference between the sequence transformations.

The purpose of the third study was to evaluate the possibility that different pattern comparison mechanisms operate under different task conditions. The experiments evaluated discrimination when the two patterns began at delayed starting times. The patterns were presented at different frequencies and to different ears, and were subjected to multiplicative compressions and expansions. Listeners performed well even when

the patterns contained tones of different frequency and in spite of the patterns being presented to separate earphone channels.

Performance was good when the sequences were presented either (near) simultaneously or at relatively long time delays. When the time between pattern onsets was less than 10-ms, discrimination was very sensitive to the expansion or compression manipulation, indicating that discrimination in this region was based on the process of waveform correlation. At longer time separations, performance was relatively insensitive to such transformations, consistent with the Pattern Correlation hypothesis. Thus the results support a two-phase mechanism: when the sequence delay is less than 20-ms, the binaural waveform correlator is the active mechanism; when the sequence delay is greater than 20-ms, the pattern correlator is the active mechanism. Morever, the efficiency of the pattern correlation mechanism is very poor when the sequences overlap in time. It appears that the sequential presentation of stimulus patterns may be a requirement for the pattern correlator to function.

TABLE OF CONTENTS

page

I. Perception of temporal patterns defined by tonal sequences

Robert D. Sorkin
Department of Psychology, University of Florida, Gainesville, Florida 32611

(Received 23 June 1989; accepted for publication 6 November 1989)

This experiment tested how listeners discriminate between the temporal patterns defined by two sequences of tones. Two arrhythmic sequences of n tones were played successively ($n = 8$, 12, or 16, tone duration = 35 ms, frequency = 1000 Hz), and the listener reported whether the sequences had the same or different temporal patterns. In the first sequence, the durations of the intertone gaps were chosen at random; in the second sequence, the gaps were either (a) the same as the first sequence or (b) chosen at random. Discrimination performance increased with the variability of the gap sequences and decreased with the size of the correlation between the sequences. A discrimination model based on computation of the sample correlation between the sequences of gaps, but limited by an internal variability of approximately 15 ms, described observer performance in a variety of conditions.

PACS numbers: 43.66.Mk, 43.66.Ba [WAY]

INTRODUCTION

How do listeners discriminate between the temporal patterns defined by two tonal sequences? The answer to this question may be relevant to important issues in the perception of speech and musical patterns. We report on some experiments and propose a model for describing behavior in tasks in which a listener must decide whether two arrhythmic, equitone sequences have the same or different temporal patterns.

Several investigators have studied the perception of partially unstructured or arrhythmic temporal sequences. Lunney (1974) showed that the discrimination of irregularity in tempo, introduced into the fourth click of the output of a metronome, was an exponential function of the period, in a range of period durations from 30–3200 ms. Pollack studied the perception of temporal gaps within trains of very brief pulses (Pollack, 1967, 1968a) and the perception of periodicity and jitter in pulse trains (Pollack, 1968b, c, d). Pollack found that the threshold for gap discrimination increased with the interpulse interval, for interpulse intervals greater than 10 ms. In general, performance was best when the pulse trains contained large numbers of intervals and had very short interpulse intervals. Pollack suggested that the processing of trains with very short interpulse intervals probably involved a spectral mode of processing, while long interpulse intervals (> 10 ms) probably required a temporal processing mode.

Sorkin *et al.* (1982) studied the perception of tone sequences with randomly jittered temporal patterns. Their subjects heard two sequences of n tones: One sequence had a fixed intertone interval and the other had jitter added to the intertone intervals. Subjects had to detect which sequence had the added jitter. Sorkin *et al.* found that discrimination improved with the number of intervals and decreased with the average duration of the intervals (the durations ranged from 20–110 ms). Their results were consistent with temporal discrimination data employing single, marked time intervals (Creelman, 1962; Getty, 1975; Divenyi and Danner, 1977; Divenyi and Sachs, 1978; and Allen, 1979).

Sorkin *et al.* (1982) proposed a statistical model of jitter detection, in which the timing of different frequency tones was monitored (and compared) across separate critical band channels; discrimination of time jitter within a critical band channel was much better than across channels. Performance increased in the expected way with the number of tones in each sequence and with the different regular frequency patterns employed. However, when the frequency patterns were random, listener performance was very much below the model's predictions.

In a similar experiment, Halpern and Darwin (1982) presented subjects with a sequence of four clicks which marked three intervals; their subjects had to indicate whether the last interval was shorter or longer than the preceding two. Halpern and Darwin tested base durations ranging from 400–1450 ms. Discrimination performance, as measured by the standard deviation of the resulting psychometric functions, was an increasing function of the base duration; the resulting Weber fraction was about 0.05, consistent with that reported by Getty (1975).

Recently, Schulze (1989) reported a variation of the Halpern and Darwin experiment in which subjects were asked to report whether the last of n intervals marked by tones was longer or the same as the $n - 1$ preceding intervals. Schulze used base durations of from 50 to 400 ms and from two to six intervals in each sequence. Schulze tested an hypothesis similar to that of the Sorkin *et al.* (1982) model about the expected improvement in discriminability with number of intervals. Discrimination improved with the number of intervals, for most of the subjects. Schulze failed to find evidence for a Weber's law effect; for his subjects, the discrimination limen was between 5 and 15 ms and independent of the base duration.

In the present experiment the listener was asked to compare two arryhthmic tonal sequences and report whether the temporal patterns were the same or different. The two sequences were either identical or had partially correlated temporal envelopes. This task is a generalization of the Sorkin *et al.* (1982) jitter-detection paradigm. An advantage of these paradigms is that the information carrying aspects of the sequences are distributed throughout the sequence, rather than concentrated on one judged interval as in the Halpern and Darwin (1982) and Schulze (1989) experiments. The goal of the present experiment was to test whether a listener's ability to perform sequence comparison can be described by a process in which the listener computes the correlation between the sequence temporal envelopes.

I. METHOD

Listeners compared pairs of tone sequences composed of n 1000-Hz tone bursts of 35-ms duration and approximately 71-dB sound-pressure level. Tone bursts were shaped by a 4-ms linear rise and decay envelope. After listening to the pair of tone sequences presented on each experimental trial, the subject had to respond whether or not the temporal pattern of tones was the same or different. There were two types of experimental trials: trials on which the identical sequence of tones and intertone intervals (gaps) were presented (*SAME* trials) and trials on which the pattern of intertone gaps was different in the two presented sequences (*DIFFERENT* trials). On trials when the sequences were different, the only difference between the sequences was in the pattern of intertone gaps and tone onsets. The first part of Fig. 1 illustrates a *SAME* trial; the second part illustrates a *DIFFERENT* trial. The type of trial was chosen at random, with $p(SAME) = 0.5$.

The intertone gaps were generated by a process that enabled the experimenter to control the mean gap duration μ_{gap}, the standard deviation of the gaps, σ_{gap}, and the correlation ρ_{ex} between the two gap sequences on trials when the sequences were different. The intertone gaps were constructed by combining three independently generated normal deviates, with one deviate common to the two sequences (see Appendix). Gap durations of less than 2 ms were not allowed. The sequence correlation is given by the ratio of two variances, the variance common to the two sequences divided by the sum of the common and unique variances (Jeffress and Robinson, 1962):

$$\rho_{ex} = \sigma^2_{com}/(\sigma^2_{com} + \sigma^2_{un}) \quad (1)$$

and

$$\sigma^2_{gap} = (\sigma^2_{com} + \sigma^2_{un}), \quad (2)$$

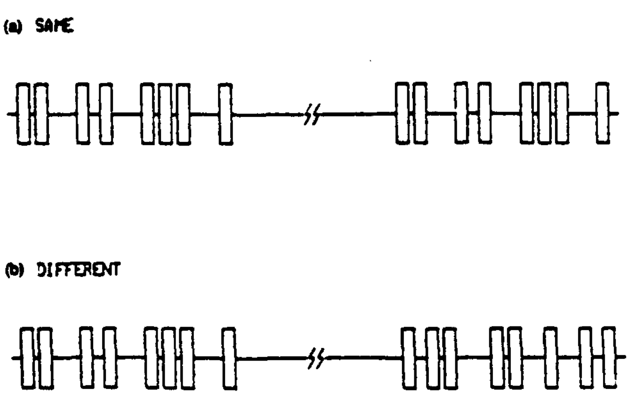

FIG. 1. The envelopes of typical tone sequences are shown for (a) same and (b) different trials.

where com and un refer, respectively, to the common and unique portions.

One male and three female undergraduate students at the University of Florida served as observers; they were paid an hourly wage plus an incentive for correct responses. Listeners had normal hearing and performed the tasks for approximately 2 h per day, 3 days per week. Listeners were seated in a double-walled acoustically insulated chamber. The stimuli were presented monaurally via TDH-39 headphones. The conditions were tested in 100 trial blocks; typically, 8 blocks were completed in a session. The two sequences to be discriminated on each trial were presented with a 750-ms intersequence separation; full feedback about the correct response was provided after each trial.

II. CORRELATION MODEL

A straightforward model of observer performance in the temporal pattern discrimination task follows from the assumption that the observer computes the correlation between the two sequences of gaps presented on each trial. Suppose that the observer's response is based on the value of the Pearson product-moment correlation coefficient statistic r_{12} computed on the sample of intertone gaps defined by the pair of sequences $\langle t_{1,1}, t_{1,2}, \ldots, t_{1,n} \rangle$ and $\langle t_{2,1}, t_{2,2}, \ldots, t_{2,n} \rangle$. A transformation of the correlation coefficient, known as the Fisher r-to-Z transformation, is defined as

$$Z = \tfrac{1}{2} \ln[(1 + r_{12})/(1 - r_{12})]. \quad (3)$$

The sampling distribution of Z is distributed approximately normally, for gaps drawn from a normal distribution and for n of at least moderate size ($n \approx 10$). If ρ is the population correlation coefficient, the mean and standard deviation of Z are then given by (Brunk, 1960)

$$\mu_z \cong \frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right) + \frac{\rho}{2n-1} \quad (4)$$

and

$$\sigma_Z \cong (n-3)^{-1/2}. \quad (5)$$

Discrimination performance can be obtained from the normalized difference between the means of the Z statistic, given the possible hypotheses on a trial: *SAME* when $\rho = 1.0$ and *DIFFERENT* when $\rho = \rho_{ex}$. The discriminability d' is given by the difference between the means of the Z statistic divided by the standard deviation of Z. [The contribution of the right-hand term of Eq. (4) is very small.]

For a human observer, the effective correlation between the sequences on *DIFFERENT* trials will depend on ρ_{ex}, σ_{gap}, and the magnitude of internal variability in the observer's encoding and storage of the gaps. We assume that the observer's observation of the gaps is subject to a temporal jitter σ^2_{in}, and that this jitter is uncorrelated across the gap sequences. Adding this uncorrelated jitter σ^2_{in} to Eqs. (1) and (2), yields

$$\rho_{DIFF} = \frac{\sigma^2_{com}}{\sigma^2_{com} + \sigma^2_{un} + \sigma^2_{in}} = \frac{\rho_{ex}}{1 + (\sigma_{in}/\sigma_{gap})^2}, \quad (6)$$

and from Eqs. (1) and (2) and $\rho = 1.0$, the effective correlation on *SAME* trials,

$$\rho_{SAME} = [1 + (\sigma_{in}/\sigma_{gap})^2]^{-1}. \quad (7)$$

The magnitude of the internal temporal jitter σ_{in} is the single parameter of the model. Because the internal jitter is independent between the two sequences, it acts to reduce the effective correlation of the sequences.

Discrimination performance can be calculated using Eqs. (4), (6), and (7) to compute the difference between the means of the Z statistic on *DIFFERENT* and *SAME* trials divided by the standard deviation of Z:

$$d' = \left[\frac{1}{2}\ln\left(\frac{1+\rho_{SAME}}{1-\rho_{SAME}}\right) + \frac{\rho_{SAME}}{2n-1} - \frac{1}{2}\ln\left(\frac{1+\rho_{DIFF}}{1-\rho_{DIFF}}\right) - \frac{\rho_{DIFF}}{2n-1}\right] \times (n-3)^{-1/2}. \quad (8)$$

III. EXPERIMENT 1: EFFECT OF SEQUENCE CORRELATION AND VARIABILITY

The purpose of this experiment was to examine how discrimination performance depended on the correlation between the sequences ρ_{ex} (as specified on *DIFFERENT* trials, since $\rho = 1$ on *SAME* trials) and the standard deviation of the intertone gaps σ_{gap}, and to estimate the magnitude of the internal noise σ_{in}.

A. Procedure

Observers were run in conditions using a range of different gap sequence correlations (from 0 to 0.8) at a fixed-gap standard deviation of 20 ms (experiment 1a), and then at gap standard deviations of 10, 20, 30, and 40 ms at a gap correlation of 0.6 (experiment 1b). The gap correlation and gap standard deviation were fixed within a block of 100 trials. The conditions were run in sequences of blocks having different gap correlations and a fixed-gap standard deviation or in sequences of blocks having different gap standard deviations and a fixed gap correlation. Table I summarizes the values for the different variables in the experiment. The order of gap correlation or gap standard deviation was randomized over the sequence of blocks. Listeners ran approximately 9000 trials before data collection was begun: no effects of practice were evident after this training period. The data indicated no strong response biases and no apparent relationship between the listeners' criteria and the conditions run. Sorkin (1962) extended detection theory to the same–different paradigm and considered some of the methodological questions involved.

TABLE I. Summary of conditions and variables for the pattern discrimination experiments. (All durations in milliseconds.)

Experiment	ρ_{DIFF}	Gap number	Gap mean	Gap standard deviation	Sequence duration
1a	0, 0.2, 0.35, 0.4				
	0.5, 0.6, 0.65, 0.8	11	50	20	970
1b	0.6	11	50	10, 20, 30, 40	970
2a	0.35	11	19	20	629
	0.35	11	50	20	970
	0.35	11	81	20	1311
2b	0.35	7	81	20	847
	0.35	11	39	20	849
	0.35	15	19	20	845

B. Results and discussion

Figure 2(a)–(d) shows the data from four observers at a mean gap duration of 50 ms and a gap standard deviation of 20 ms. Figure 3 shows the data averaged over the four observers at a gap mean of 50 ms. The vertical bars in the figures indicate plus and minus one standard error of the mean; in Fig. 3, these are the average of the standard errors for the four listeners in each condition. The solid lines in Fig. 2 are least-squares fits of the model to each observer's average data; the value of the internal jitter parameter is shown in each section of the figure. In Fig. 3, the model is fit to the average data.

The observed drop in listener performance with increases in the correlation of the sequences is consistent with the model. Discrimination performance should drop as the sequence correlation is increased, since the magnitude of any observable differences between the sequences must decrease as their temporal envelopes become more highly correlated. The value of the (single) internal temporal jitter parameter was 14.75 ms, for the fit of the model to the average data from the four listeners. This value for the internal jitter is at the high end of the range of values obtained in duration discrimination experiments employing single and multiple judged intervals (Lunney, 1974; Getty, 1975; Divenyi and Danner, 1977; Halpern and Darwin, 1982; Sorkin *et al.*, 1982; and Schulze, 1989). This value will be used for all subsequent fits of the model.

Figure 4 shows how average performance depended on the standard deviation of the gap duration. The vertical bars indicate plus and minus one standard error of the mean; the average standard errors for the four observers are shown for each condition. The solid line is the prediction of the correlation model, using the value of the internal jitter (based on the average data) of Fig. 3. According to the model, as the level of external variability in the gaps increases, the contribution of internal and (assumed) uncorrelated variability is reduced, and performance should improve. It is apparent that the model overestimates performance at high standard deviations of the gap.

IV. EXPERIMENT 2: EFFECT OF GAP DURATION AND NUMBER

The purpose of the second experiment was to examine how discrimination performance depended on the mean gap duration μ_{gap} and on the number of intertone gaps, n, and to compare these observations to the predictions of the model.

A. Procedure

Listeners ran two conditions in which the gap sequence correlation was fixed at 0.35, the gap standard deviation was fixed at 20 ms, and the mean and number of intertone gaps were varied. As in experiment 1, the gap sequence correla-

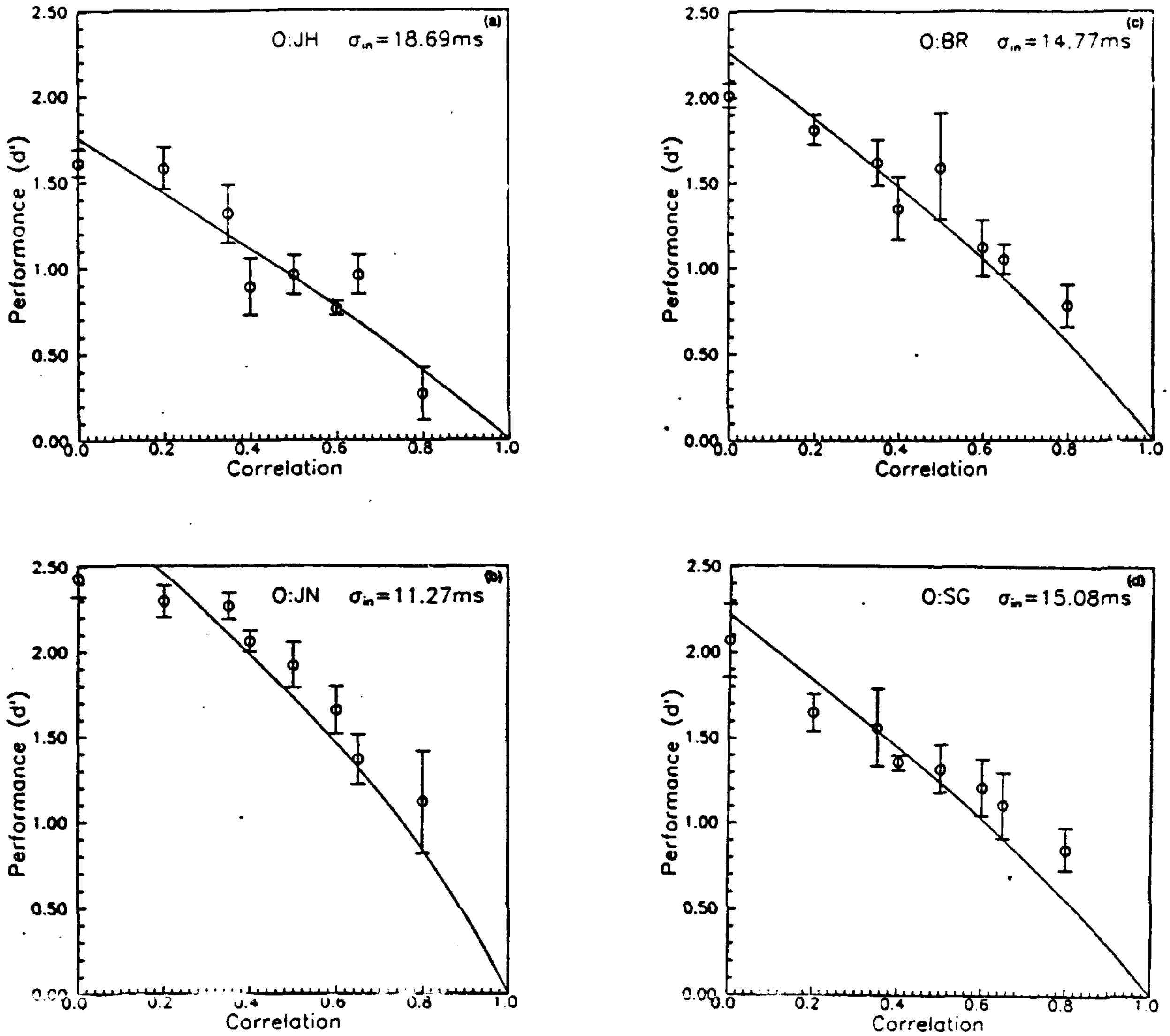

FIG. 2. Performance (d') is plotted as a function of the sequence correlation, for each of four observers. The solid lines show the performance of the correlation model with the internal noise standard deviation shown (see text).

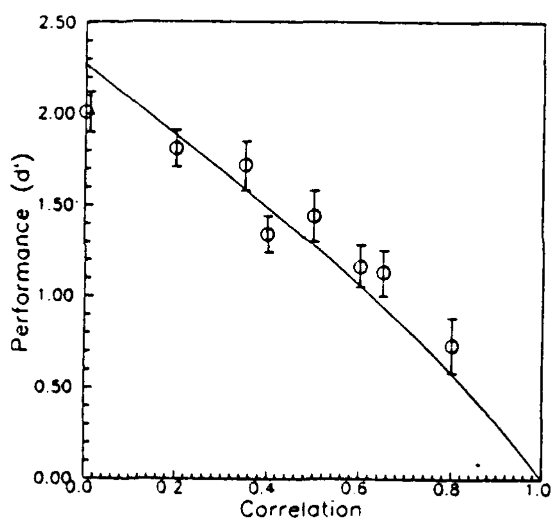

FIG. 3. The average performance of four observers (d') is plotted as a function of the sequence correlation. The solid line is the prediction of the correlation model with an internal noise of 14.75 ms.

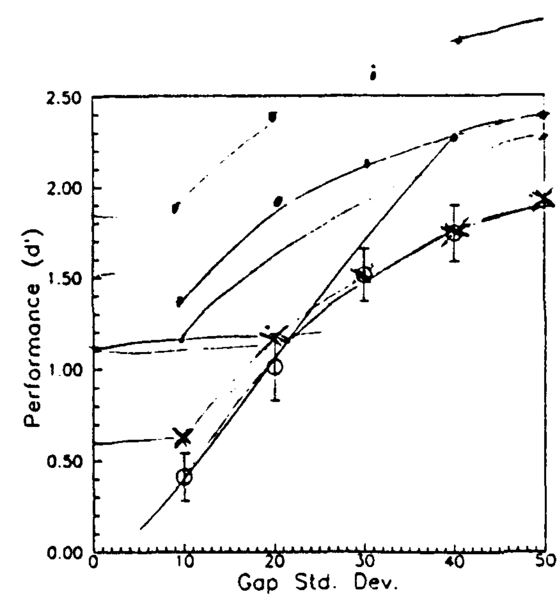

FIG. 4. The average performance of four observers (d') is plotted as a function of the standard deviation of the gaps. The solid line is the prediction of the correlation model with an internal noise of 14.75 ms.

tion, gap standard deviation, mean gap, and number of gaps were fixed within a block of 100 trials. The observers were run in sequences of blocks of fixed mean gap duration (or fixed gap and number); the order of conditions was randomized over blocks. Table I summarizes the experimental conditions. In experiment 2a, the mean gap condition, the number of gaps was fixed at 11 and the mean gap was either 19, 50, or 81 ms. In experiment 2b, the number of gaps condition, there were three gap-number–mean-gap pairings: 7 gaps with a mean of 81 ms, 11 gaps with a mean of 39 ms, and 15 gaps with a mean of 19 ms. These values were chosen so that the total duration of the gap sequence would be fixed at approximately 850 ms. The values of n and μ_{gap} were chosen to allow testing of a range of gap durations, subject to the constraint of avoiding excessively long stimulus sequences.

B. Results and discussion

Figure 5 shows the average performance in the mean gap condition as a function of the magnitude of the mean gap. As the mean gap was increased, observer performance decreased at an increasing rate. The model, as defined by Eqs. (6)–(8), made no assumption about the dependence of performance upon μ_{gap}. However, it is reasonable to expect that a Weber's law relationship would hold, such that the magnitude of the internal jitter σ_{in} would increase with the duration of the intervals to be judged. Such a relationship, where σ_{in} increases in proportion to μ_{gap}, has been found by Lunney (1974), Getty (1975), Divenyi and Danner (1977), Halpern and Darwin (1982), and Sorkin *et al.*, (1982).

In order to quantify the contribution of a Weber's law dependence of performance on gap duration in the present experiment, we set the internal jitter equal to a linear function of the mean gap duration:

$$\sigma_{in} = A + B\mu_{gap} , \qquad (9)$$

where A and B are constants. To estimate the parameters of the function, we reexamined the jitter discrimination data reported in our earlier study of sequence discrimination

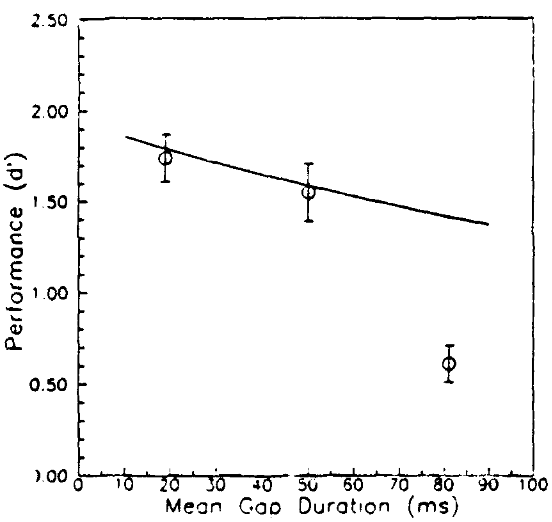

FIG. 5. The average performance of four observers (d') is plotted as a function of the mean gap duration. The solid line is the prediction of the correlation model revised to incorporate the effect of mean gap (see text).

(Sorkin *et al.*, 1982). In that study, listeners had to detect the presence of jitter added to equitone or binary tone sequences. That is, let

$$\sigma_{d'=1.0} = A + B\mu_{gap} , \qquad (10)$$

where $\sigma_{d'=1.0}$ is the standard deviation of the jitter discriminable at a $d' = 1.0$. The value of A in the Sorkin *et al.* study varied depending on the type of sequences tested. However, the slope B was relatively constant, at least for the equitone and alternating tone conditions. The slope was approximately 0.04 and 0.07 for subject P and S, respectively (see Figs. 6 and 7 in Sorkin *et al.*, 1982, for the equitone and alternating tone conditions, $n = 10$, and mean durations of 20–110 ms). For the current purpose, we chose an intermediate value of 0.05 for the B parameter. This value closely agrees with the Weber fractions obtained by Lunney (1974), Getty (1975), Divenyi and Danner (1977), and Halpern and Darwin (1982).

To estimate the value for the A parameter in the current experiment, we substituted $B = 0.05$, $\mu_{gap} = 50$, and $\sigma_{in} = 14.75$ ms in Eq. (9) (recall that $\sigma_{in} = 14.75$ ms is the value of the internal noise obtained in experiment 1 at $\mu_{gap} = 50$ ms). This yielded a value for A of 12.25 ms. The resulting expression for σ_{in} was then employed in Eqs. (6) and (7) for the computation of d'.

The prediction of the revised model is plotted as the solid line in Fig. 5; although the model's performance drops with increasing gap size, the drop is much less than that shown by the human observers at 80 ms. Some part of this performance drop at long gap means may be attributable to the fact that as the mean gap is increased, the total duration of the sequences becomes quite long. At mean gap durations of 19, 50, and 81 ms, the sequence spans are approximately 0.6, 1, and 1.3 s. An observer also must hold the information in the first sequence over the intersequence interval of 750 ms. It is possible that spans approaching 1 s or longer exceed the capacity of the observer's auditory memory, and hence the effective number of intervals being processed is much smaller than assumed by the model (see Watson, 1987).

Figure 6 shows the average performance of the observers as a function of the number of intertone gaps. Both the number of tones (and gaps) and the mean gap were manipulated, in order that the total duration of the sequence span would be held constant at approximately 0.85 s. Performance increased between 7 and 11 gaps and then leveled off. The solid curve shows the prediction of the revised model, using Eq. (9) and the values of A and B specified in the preceding paragraphs. The dashed curve is the model prediction based on an internal jitter that is independent of the mean gap (set equal to the prediction of the former model at $n = 7$). Both versions of the model overpredict performance at n equal to 15.

V. GENERAL DISCUSSION

I have tried to show that the discrimination of differences between temporally perturbed tone sequences may be described as a process in which the listener computes the correlation between the temporal envelopes of the sequences. This computation appears to be limited by an inter-

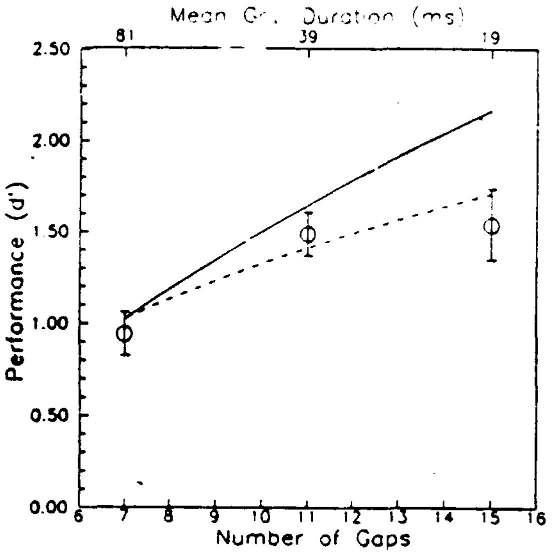

FIG. 6. The average performance of four observers (d') is plotted as a function of the number of gaps (average sequence duration is fixed). The solid line is the prediction of the correlation model revised to incorporate the effect of mean gap ($\sigma_{in} = 12.25 + 0.05\mu_{gap}$). The dashed line is the prediction of the correlation model with a fixed internal noise of $12.25 + (0.05)(81) = 16.3$ ms.

nal temporal variability, or noise, in the listener's encoding and storage of the stimulus information. In this study, the magnitude of the internal noise was approximately 15 ms. This is about 5–10 ms higher than difference thresholds obtained using two interval duration discrimination tasks. Consistent with the results of other studies, the level of the internal noise was dependent on the magnitude of the base duration to be discriminated. Performance was degraded when the time span of the sequences to be compared was longer than 1 s. Performance also was degraded when the listener was required to compare sequences having more than 12 intervals. These latter two effects probably are related to limitations in memory capacity or to the listener's use of a temporal window that is not uniform over the sequences.

The idea that a listener can compare auditory patterns by computing the correlation between temporal or spectral aspects of the patterns is not novel. Many models of the binaural detection mechanism have assumed a process that computes the interaural correlation between the left and right auditory channels (Durlach, 1963; Osman, 1971; Lindemann, 1986; and cf. Sorkin, 1965, and Pohlmann and Sorkin, 1974). Several investigators have studied the binaural discrimination of changes in the interaural whole-waveform correlation of the signals (e.g., for wideband noise, Pollack and Trittipoe, 1959; for pulse train polarity agreement, Pollack, 1971; and for wideband, narrow-band, and low-pass noise, Gabriel and Colburn, 1981). These studies have reported a dependence of discrimination on interaural correlation that is consistent with the hypothesized correlation process.

Recently, Richards (1987) reported an experiment on the discrimination of differences between simultaneously presented noise stimuli having partially correlated amplitude (and spectral) envelopes. Richards postulated a correlation discrimination process that is essentially identical to the one proposed in the present study. Her noise stimuli had bandwidths of 100 Hz and center frequencies of 2500 and 2750 Hz. For any given stimulus, these two noise bands had, on average, a specified correlation. The observers had to discriminate which of two such stimuli contained the higher correlation across the spectral bands. Richards tested her observers' ability to discriminate between a reference stimulus, containing either a zero or unit noise correlation, and target stimuli having a range of noise correlations. In general, her results supported the model: The observers' sensitivity to changes in envelope correlation was a monotonic function of the computed Z statistic and was essentially independent of the specific reference correlation.

In the binaural studies and in Richard's noise study, one assumes that the listener can compute the correlation between the transduced, critical-band-filtered signals; the signals are assumed to undergo minimal processing prior to the correlation operation. A similar process could be operating in the present study: The signals in each sequence are transduced, subjected to windowing and filtering operations, and then stored; finally, the correlation is computed between the resulting waveforms. An alternative, more cognitive, conception is that the listener processes each sequence so that only the magnitudes of the time intervals between tone onsets are encoded and stored. The listener then computes the correlation between the two lists of interonset times. This view of the correlation process implies different relationships between performance and the task characteristics. In contrast to the whole-waveform correlation, the computation of correlation based on two lists of stored numbers should be less sensitive to certain transformations of the sequences such as temporal compression or expansion. A future experiment will examine this idea.

The listener's subjective impression of the present task is of trying to recall and compare two briefly heard rhythmic patterns. That observation, and the relatively long interonset intervals employed in the current experiment, support the idea that the listener is using a temporal rather than spectral processing mode. In addition, changing the frequency of all of the tones in the second sequence has a negligible effect on performance. Even so, we would expect the simple correlation model to fail when the sequences are composed of tones of more than a single frequency. Many studies of the perception and production of temporal patterns have demonstrated the influence of sequence temporal structure on spectral pattern discrimination (Deutsch, 1980; Jones, 1981; Jones *et al.*, 1981; Jones *et al.*, G., 1982; and Monahan, 1987) as well as the influence of sequence spectral pattern on temporal pattern discrimination (Woods *et al.*, 1979; Handel and Lawson, 1983; Espinoza-Varas and Jamieson, 1984; Espinoza-Varas and Watson, 1986; and Sorkin, 1987).

The model of temporal jitter detection supported by the Sorkin *et al.* (1982) study assumed that best performance would occur when the tones marking the intervals were within a critical band in frequency. In that experiment, the detection of jitter in sequences containing different frequency tones was predictably poorer than with equitone sequences. It is possible that a similar assumption would enable the correlation model to describe pattern comparisons between multiple-frequency tone sequences.

For example, the listener might compute the correlation between the temporal envelopes of tone subsequences defined only within a single critical band. Correlations computed within separate critical bands then could be combined, in order to arrive at a composite estimate of the temporal similarity of the sequences.

ACKNOWLEDGMENTS

This research was sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant AFOSR-89-0021. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. Portions of this article were presented at the 118th meeting of the Acoustical Society of America in St. Louis. I thank Mark Pezzo and Dr. Greg C. Elvers for their help running the experiments, and Dr. David M. Green, Dr. Virginia M. Richards, and two anonymous reviewers, for helpful comments on an earlier draft of this article.

APPENDIX

The gap mean, standard deviation, and correlation were controlled by generating the gap durations in the following manner: Three independent normal deviates, x_a, x_b, and x_c, were generated and their absolute values added to arrive at random variables with a correlation of ρ_{ex}

$$x_1 = u|x_a| + c|x_c|, \quad \text{(A1)}$$

$$x_2 = u|x_b| + c|x_c|, \quad \text{(A2)}$$

where u and c are constants defined by

$$c = \rho_{ex}^{1/2}, \quad u = (1-\rho_{ex})^{1/2}. \quad \text{(A3)}$$

The resulting x_1 and x_2 values were limited to values between zero and 2.5 ($p < 0.02$) and then linearly transformed to arrive at gap sequences $\{t_{1,i}\}$ and $\{t_{2,i}\}$ with gap mean equal to μ_{gap} and standard deviation equal to σ_{gap}. To check these procedures, we computed the sample correlation coefficients r_{12} and the distributions of Z [Eq. (3)]; the t_1 and t_2 sequences had an average correlation equal to ρ_{ex}, and the Z distributions were approximately normal.

Allen, L. G. (1979). "The perception of time." Percept. Psychophys. 26, 340–354.

Brunk, H. D. (1960). *An Introduction to Mathematical Statistics* (Ginn, Boston).

Creelman, C. D. (1962). "Human discrimination of auditory duration." J. Acoust. Soc. Am. 34, 582–593.

Espinoza-Varas, B., and Jamieson, D. G. (1984). "Integration of spectral and temporal cues separated in time and frequency," J. Acoust. Soc. Am. 76, 732–738.

Espinoza-Varas, B., and Watson, C. S. (1986). "Temporal discrimination for single components of nonspeech auditory patterns," J. Acoust. Soc. Am. 80, 1685–1694.

Deutsch, D. (1980). "The processing of structured and unstructured tonal sequences," Percept. Psychophys. 28, 381–389.

Divenyi, P. L., and Danner, W. F. (1977). "Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra." Percept. Psychophys. 21, 125–142.

Divenyi, P. L., and Sachs, R. M. (1978). "Discrimination of time intervals bounded by tone bursts," Percept. Psychophys. 24, 429–436.

Durlach, N. I. (1963): "Equalization and cancellation theory of binaural masking level differences," J. Acoust. Soc. Am. 35, 1206–1218.

Gabriel, K. J., and Colburn, H. S. (1981). "Interaural correlation discrimination: I. Bandwidth and level dependence," J. Acoust. Soc. Am. 69, 1394–1401.

Getty, D. J. (1975). "Discrimination of short temporal intervals: A comparison of two models," Percept. Psychophys. 18, 1–8.

Handel, S., and Lawson, G. R. (1983). "The contextual nature of rhythmic interpretation," Percept. Psychophys. 34, 103–120.

Halpern, A. R., and Darwin, C. (1982). "Duration discrimination in a series of rhythmic events," Percept. Psychophys. 31, 86–89.

Jeffress, L. A., and Robinson, D. E. (1962). "Formulas for the coefficient of interaural correlation for noise," J. Acoust. Soc. Am. 34, 1658–1659.

Jones, M. R. (1981). "A tutorial on some issues and methods in serial pattern research," Percept. Psychophys. 30, 492–504.

Jones, M. R., Boltz, M., and Kidd, G. (1982). "Controlled attending as a function of melodic and temporal context," Percept. Psychophys. 32, 211–218.

Jones, M. R., Kidd, G., and Wetzel, R. (1981). "Evidence for rhythmic attention," J. Exp. Psychol. Human Percept. Perform. 7, 1059–1073.

Lindemann, W. (1986). "Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," J. Acoust. Soc. Am. 80, 1608–1622.

Lunney, H. W. M. (1974). "Time as heard in music and speech," Nature 249, 592.

Monahan, C. B., Kendall, R. A., and Carterette, E. C. (1987). "The effect of melodic and temporal contour on recognition memory for pitch change," Percept. Psychophys. 41, 576–600.

Osman E. (1971). "A correlation model of binaural masking level differences," J. Acoust. Soc. Am. 50, 1494–1511.

Pohlmann, L. D., and Sorkin, R. D. (1974). "Binaural masking level differences for pulse train signals of differing interaural correlation," J. Acoust. Soc. Am. 55, 1293–1298.

Pollack, I. (1967). "Asynchrony: The perception of temporal gaps within periodic auditory pulse patterns," J. Acoust. Soc. Am. 42, 1335–1340.

Pollack, I. (1968a). "Asynchrony II: Perception of temporal gaps within periodic and jittered pulse patterns," J. Acoust. Soc. Am. 43, 74–76.

Pollack, I. (1968b). "Detection and relative discrimination of auditory 'jitter'," J. Acoust. Soc. Am. 43, 38–315.

Pollack, I. (1968c). "Discrimination of mean temporal interval within jittered auditory pulse trains," J. Acoust. Soc. Am. 43, 1107–1112.

Pollack, I. (1968d). "Periodicity discrimination for auditory pulse trains." J. Acoust. Soc. Am. 43, 1113–1119.

Pollack, I. (1971). "Interaural correlation detection for auditory pulse trains," J. Acoust. Soc. Am. 49, 1213–1217.

Pollack, I., and Trittipoe, W. J. (1959). "Binaural listening and interaural cross correlation," J. Acoust. Soc. Am. 31, 1250–1252.

Richards, V. M. (1987). "Monaural envelope correlation perception," J. Acoust. Soc. Am. 82, 1621–1630.

Schulze, H. H. (1989). "The perception of temporal deviations in isochronic patterns," Percept. Psychophys. 45, 291–296.

Sorkin, R. D. (1962). "Extension of the theory of signal detectability to matching procedures in psychoacoustics," J. Acoust. Soc. Am. 34, 1745–1751.

Sorkin, R. D. (1965). "Uncertain signal detection with simultaneous contralateral cues," J. Acoust. Soc. Am. 38, 207–212.

Sorkin, R. D. (1987). "Temporal factors in the discrimination of tonal sequences," J. Acoust. Soc. Am. 82, 1218–1226.

Sorkin, R. D., Boggs, G. J., and Brady, S. L. (1982). "Discrimination of Temporal Jitter in Patterned sequences of Tones," J. Exp. Psychol.: Human Percept. Perform. 8, 46–57.

Watson, C. S. (1987). "Uncertainty, informational masking, and the capacity of immediate auditory memory," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 267–277.

Woods, D. D., Sorkin, R. D., and Boggs, G. J. (1979). "Stimulus context and duration discrimination," Percept. Psychophys. 26, 127–132.

II. EFFECT OF TIME COMPRESSION AND EXPANSION ON THE DISCRIMINATION OF TONAL PATTERNS.

ABSTRACT

This experiment tested how well human listeners can discriminate between temporal patterns that are compressed or expanded in time. The listener's task was to determine whether two arrhythmic, tonal sequences had the same or different temporal patterns. According to the Pattern Correlation model [Sorkin, J. Acoust. Soc. Am., 87, (1990)], listeners perform this task by computing the correlation between the pattern of time intervals marked by the tones in each sequence. Listener performance dropped when one of the sequences was compressed or expanded in time. In order for the model to describe the observed performance, it was necessary to postulate an internal noise component that was proportional to the magnitude of the difference between the sequence transformations.

INTRODUCTION

One of the most intriguing features of temporal pattern perception is the ability to recognize patterns as similar, despite time compression or expansion. Examples of such time normalization abound in speech and music perception and we are normally unaware of such temporal changes, even when they occur during relatively brief stimuli, such as words. Our experiments evaluated the human listener's ability to discriminate between word-length tonal sequences that were subjected to such transformations. We also tested how well the Pattern Correlation model (Sorkin, 1990) would predict the effects of time compression and expansion on listeners' performance.

A number of investigators have examined the human listener's sensitivity to the temporal properties of non-repeating, multi-tone, arrhythmic, sequences. In these experiments, the listener is asked to detect a small temporal difference or jitter in the patterns. Performance is a Weber-like function of the time intervals between marker tones (Halpern and Darwin, 1982; Hirsch, Monahan, Grant, and Singh, 1990; Lunney, 1974; Pollack, 1967, 1968a,b,c; Schulze, 1989; and Sorkin, 1990). The Weber ratio (delta-t/T) typically varies from approximately 5% to 20%, depending on the particular task conditions. Similar results have been obtained in experiments using single marked intervals (Abel, 1972a,b; Creelman, 1962; Getty, 1975; Divenyi and Danner, 1977; Divenyi and Sachs, 1978; Espinoza-Varas and Jamieson, 1984; and the review by Allan, 1979).

Temporal pattern discrimination also depends on the number of marked intervals (Schulze, 1989; Sorkin,1982), the spectral structure of the marker pattern (Bregman and Campbell, 1971; Bregman and Dannenbring, 1973; Espinoza-Varas and Watson, 1986; Preusser, 1972; Royer and Garner, 1966, 1970; Sorkin, 1982; and Woods et al., 1979), the temporal structure of the markers (Bharucha and Pryor, 1986; Bregman, 1990; Deutsch, 1980; Monahan and Hirsch, 1990; Jones et al., 1981; Jones et al., 1982; Sturges and Martin, 1974; Monahan, Kendall, and Carterette, 1987), and the temporal location of the information in the stimulus sequence (Espinoza-Varas and Watson, 1986; Hirsch, Monahan, Grant, and Singh, 1990; Watson et al., 1975, Watson et al., 1976).

Kidd and Watson (1988) tested listeners' ability to detect frequency changes in 5-tone random patterns that were exposed to frequency transpositions and or temporal expansions. These transformations involved multiplying the frequency or duration of one of the patterns by a constant factor between 1.12 and 2. They reported that performance decreased as a function of the magnitude of the transformation, with frequency transposition producing a larger degradation than temporal expansion. Sorkin and Snow (1987) reported a similar experiment in which listeners had to indicate whether two 8-tone sequences of 50-ms tones had the same or different frequency patterns. All tone and gap durations in the second sequence were expanded or compressed by up to 40%. They reported a drop in performance that was dependent on the magnitude of the time transformation.

In the present experiments, the listener is presented with two, successively played, arrhythmic tonal sequences. The series of time intervals between the tone onsets in each sequence define the temporal patterns to be discriminated. On half of the trials (SAME trials) these two temporal patterns are identical, and on half of the trials (DIFFERENT trials) the patterns are different; the listener must report which condition exits. An important experimental variable is the correlation, ρ_{ex}, between the two series of tone interonset times, $\{X_1\}$ and $\{X_2\}$. On SAME trials, $\rho_{ex} = 1.0$, and on DIFFERENT trials, ρ_{ex} is set at a constant value, less than 1.0, that depends on the particular condition. The task is easiest when $\rho_{ex} = 0$, and becomes more difficult as ρ_{ex} approaches unity.

We take, as a working assumption, that the perceived difference between temporal patterns (absent other cues, such as changes in amplitude or frequency) is dependent on the listener's estimate of the correlation between the two series of interonset times. The listener could estimate this correlation by computing the Pearson product-moment correlation coefficient on the lists of (transduced and encoded) interonset times. Distortions of the tonal sequences, such as temporal expansion or compression, should affect listener performance only to the extent that such transformations produce changes in the estimated correlation.

The ability to estimate the correlation between brief temporal patterns is suggested by experiments on speech perception. Apparently, listeners can calibrate or normalize an incoming speech signal for the unique timing-related properties of the speaker; this process enables the listener to modify the interpretation of important phonetic cues, such as voice onset time (Miller and Liberman, 1979; Miller and Dexter, 1988; Miller and Grosjean, 1981; and see the review by Miller, 1987). Miller and Volaitis (1989) showed that information about speaking rate (syllable duration) may be used to shift the category boundary that defines whether consonants are identified as voiced or voiceless (i.e. /b/ vs /p/). The listener could implement this process by estimating and comparing phonetic intervals in a relative fashion, by computing ratios of critical intervals to reference intervals. These relative intervals would then be compared to the list of relative intervals that characterize the phonetic prototypes. This is analogous to a correlation computation.

One aspect of our ordinary experience with musical rhythm is its apparent insensitivity to small changes in rate or tempo. The importance of relative, rather than absolute, time in musical perception was stressed by Handel (1989), in a discussion of factors determining the rhythmic character of auditory sequences. However, Handel pointed out that musical rhythm may not be independent of tempo, since changes in tempo can produce changes in the perceived dissimilarity of melodies. For example, Gabrielsson (1973) performed a multidimensional scaling analysis on a number of sets of different rhythmic patterns. Subjects had to rate the similarity between pairs of patterns drawn from each set. He found that changes in metronomic tempo produced effects on the subjects' similarity space at least as large as those produced by differences in the meter or temporal pattern of the stimuli. Gabrielsson's task is quite different from the current sequence discrimination tasks. A comparable scaling task would require listeners to rate the similarity of rhythmic patterns while under instructions to <u>ignore</u> differences in tempo.

We report on two experiments in this paper: In the first experiment we evaluated the effects of multiplicative time transformations to the tonal patterns to be discriminated. That is, we test the effects of multiplying all time intervals in both sequences, or in the second sequence alone, by a fixed constant. In the second experiment we test the effects of adding a fixed time interval to all times in both sequences, or to the second sequence alone. We also derive predictions for the behavior of the Pattern Correlation model, under the assumption of internal, uncorrelated noise.

METHOD

Two groups of subjects participated in the experiment. The first group consisted of one male and three females; the second consisted of two females and two males. One of the female subjects in the first group (MW) also served in the second group. All subjects were students at the University of Florida. They were paid an hourly wage plus an incentive for correct responses. All the subjects had normal hearing and performed the tasks for approximately 2 h per day, 3 days per week. Subjects were seated in a double-walled acoustically insulated chamber. The stimuli were presented monaurally via TDH-39 headphones. The conditions were run in 100 trial blocks; typically, 8 blocks were completed in a session. All independent variables (such as correlation and magnitude of time transformation) were held constant within a block of trials. Full feedback about the correct response was provided after each trial.

The subjects compaired pairs of tone sequences composed of 8, 1000 Hz tone bursts presented at 71-dB sound-pressure level. The tone bursts were shaped by a 4-ms linear rise and decay envelope. An interval of 825-s separated the pair of tone sequences. After listening to the pair of sequences on each trial, the subject indicated whether or not the temporal pattern of tones was the same or different. On a random half of the experimental trials, the temporal patterns were the same (SAME trials), that is, the sequence pattern correlation, ρ_{ex}, was 1.0. On half of the trials the patterns were different (DIFFERENT trials); that is, ρ_{ex} was fixed at either 0.2, 0.4, or 0.6. The average time interval between tone onsets varied from 60-ms to 120-ms, depending on the condition. The minimum interval between tones (offet to onset) was 2 ms. The first part of Figure 1 illustrates a SAME trial; the second part illustrates a DIFFERENT trial.

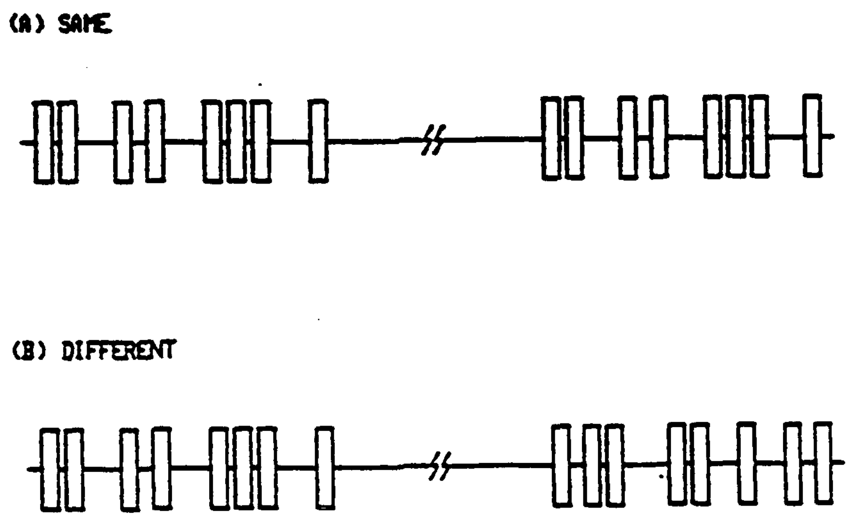

Figure 1. The envelopes of typical tone sequences are shown for same (a) and different (b) trials.

The time intervals between tones were generated by a process that enabled control of the mean and standard deviation of the intertone intervals. The sequences were generated by combining three independent normal random variables, X_a, X_b, X_c, where $\sigma_a^2 = \sigma_b^2 = \sigma_u^2$. The random variables were combined to form the two sequences of interonset times $\{X_1\}$ and $\{X_2\}$ in the following manner:

$$X_1 = X_a + X_c \text{ and } X_2 = X_b + X_c \quad (1)$$

The variance of the interonset times, σ_{ex}^2, is

$$\sigma_{ex}^2 = \mathrm{Var}[X_1] = \mathrm{Var}[X_2] = \mathrm{Var}[X_a + X_c] = \sigma_u^2 + \sigma_c^2 \quad (2)$$

The correlation between the sequences is determined by the ratio of the variance common to the two sequences, divided by the sum of the common and unique variances (see Section V and also Jeffress and Robinson, 1962):

$$\rho_{ex} = \sigma_c^2 / \sigma_{ex}^2 \quad (3)$$

For an ideal listener, e.g. one having no internal noise, the actual correlation on SAME trials would be equal to 1.0 and on DIFFERENT trials would be equal to ρ_{ex} (see Section V).

PATTERN CORRELATION MODEL

Sequence discrimination can be modeled by a process that estimates the correlation between the two series of tone interonset times (Sorkin, 1990). The main assumption of this model is that the listener's decision is based on the Pearson product-moment correlation computed on the interonset intervals and that the listener's performance is limited by internal noise. In the study by Sorkin (1990), the listener's performance was specified by a single parameter: the magnitude of the temporal jitter in the listener's encoding of the time between tone markers. This jitter was approximately 15-ms when the average time interval between tone onsets of 85-ms. Performance dropped when the intertone interval was increased, indicating that the internal noise was an increasing function of the duration of the intertone interval.

One of the goals of the present experiment was to evaluate the effects of uniform time expansions or compressions on the performance of the pattern correlation model. Suppose that we have two lists of numbers, $\{X_1\}$ and $\{X_2\}$, and that we wish to estimate the correlation between the lists, $\rho_{X1,X2}$. Our estimate of $\rho_{X1,X2}$ should not be affected by multiplying all items in $\{X_1\}$ by the factor k_1 and all items in $\{X_2\}$ by the factor k_2. The same would be true if all the $\{X_1\}$ were increased by the additive constant, t_1, and all the $\{X_2\}$ were increased by the additive constant, t_2.

These predictions may change if our estimates of the X_i are degraded by internal noise, because the nature and magnitude of the internal noise influence how accurately we can estimate $\rho_{X1,X2}$.

For example, suppose that there is a fixed internal noise that is independent of the magnitude of the intervals to be judged. An expansive transformation to both sequences, e.g. multiplying all the elements of $\{X_1\}$ and $\{X_2\}$ by the same factor, $k_1 = k_2 = k_m$, where k_m is greater than 1, will improve the accuracy of the correlation estimate. This is because increasing the element magnitudes, prior to adding the internal noise, reduces the influence of the internal noise on the estimate of ρ. The opposite result would obtain if k_m were less than unity. On the other hand, additive transformations, such as t_1 and t_2, should have no effect on performance, because such transformations have no effect on the variances of the element lists.

The effects of duration transformations to the first and second sequences are derived in Section V. The derivation results in the following equations which describe the effective correlation between the sequences on SAME and DIFFERENT trials:

$$\rho_{SAME} = k_1 k_2 [k_1^2 + (\sigma_{in}/\sigma_{ex})^2]^{-\frac{1}{2}} [k_2^2 + (\sigma_{in}/\sigma_{ex})^2]^{-\frac{1}{2}} \quad (4)$$

and

$$\rho_{DIFF} = \rho_{ex} \rho_{SAME} \quad (5)$$

where σ_{ex} is the experimental variable defined in equation 2, and σ_{in} is the internal noise. The additive constants t_1, t_2 have no effect on the correlation.

We assume that the listener estimates $\rho_{x1,x2}$ by computing the Pearson product-moment correlation between the two patterns of interonset times. The Fisher r-to-Z transformation (see Section V) yields a normal decision variable Z with known mean and standard deviation. We can predict the listener's performance by taking the normalized difference between the means of the Z statistic on SAME and DIFFERENT trials:

$$d' = (n-3)^{\frac{1}{2}} \left[\frac{1}{2} \ln\left(\frac{1+\rho_{SAME}}{1-\rho_{SAME}}\right) + \frac{\rho_{SAME}}{2n-1} - \frac{1}{2} \ln\left(\frac{1+\rho_{DIFF}}{1-\rho_{DIFF}}\right) - \frac{\rho_{DIFF}}{2n-1} \right] \quad (6)$$

where n is the number of interonset intervals and ρ_{SAME} and ρ_{DIFF} are defined by equations 4 and 5.

EXPERIMENT 1. EFFECT OF TIME COMPRESSION OR EXPANSION

The purpose of this experiment was to examine the effect of uniform time compression or expansion of the sequences to be discriminated.

A. Procedure

In order to test the effects of multiplicative transformations on discrimination, listeners were run under two experimental conditions: (a) control conditions in which both of the sequences on each trial received the same multiplicative time transformation, $k_1 = k_2 = k_m$; and (b) test conditions, in which only the second sequence of the pair on each trial was compressed or expanded, $k_1 = 1.0$, $k_2 = k_m$. The test and control conditions were run under three values of pattern correlation; for all values of k_m, ρ_{ex} was set equal to 0.2, 0.4, or 0.6. The control conditions were run with k_m equal to 0.6, 0.8, 1.0, 1.2, and 1.4, and the test conditions were run with k_m equal to 0.6, 0.8, 0.9, 1.1, 1.2, and 1.4. The correlation and transformation levels were fixed within each block of trials. The nominal duration of the tones was 25-ms and the nominal duration of the mean interonset interval (μ_{IOT}) was 75-ms. The nominal value of σ_{ex} was 25-ms. These durations were scaled proportionately by the value of the multiplicative constant, k_m. At least four blocks of 100 trials were run at each experimental condition. Listeners ran several thousand trials before data collection was begun; no effect of practice was evident on discrimination performance. The data indicated no consistent relationship between the listeners' response criteria and the duration transformation condition.

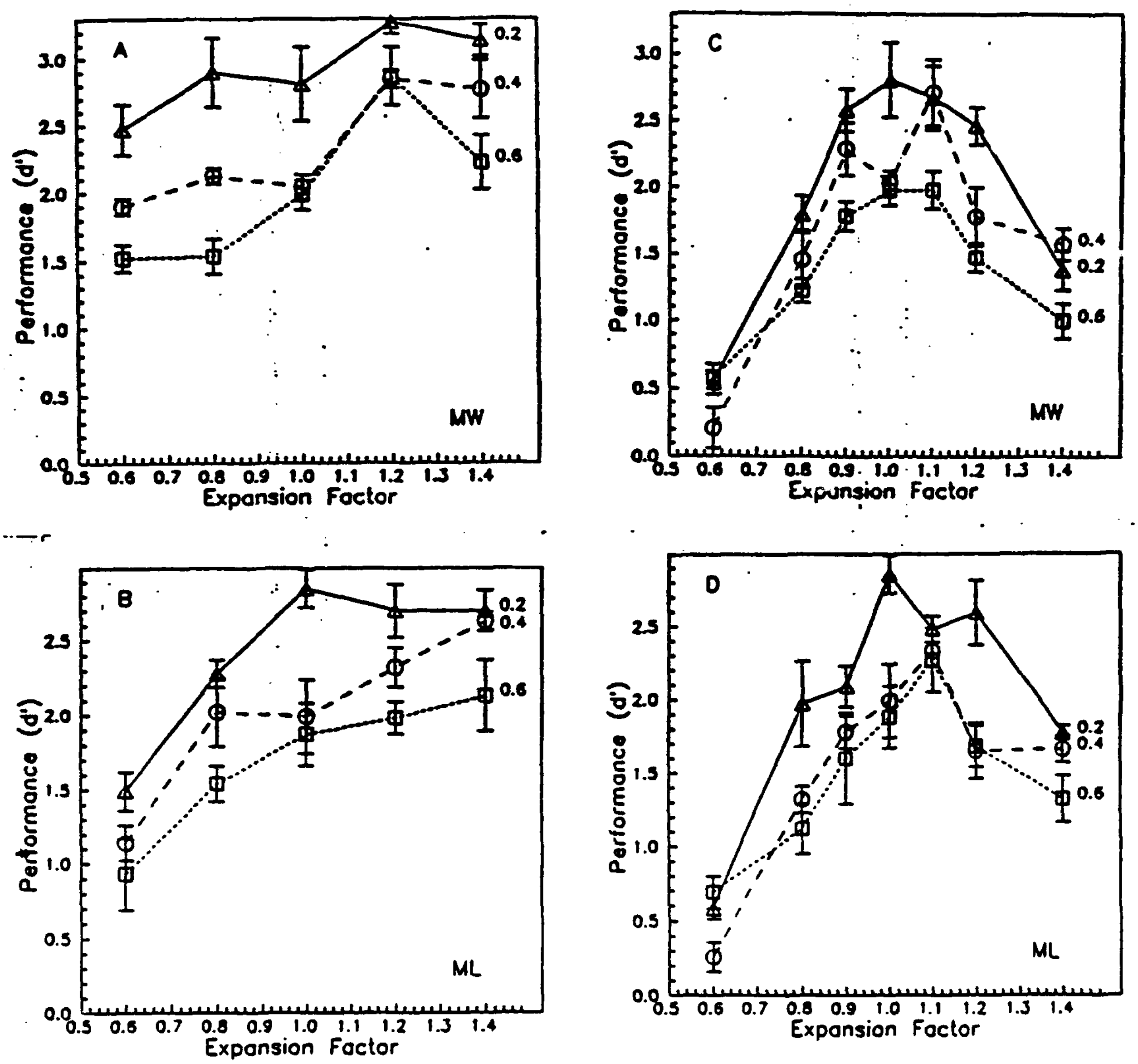

Figure 2. The performance (d') of two listeners (MW: panels A,C; ML: panels B,D) is plotted as a function of the uniform time expansion of both sequences (panels A,B) or the second sequence alone (panels C,D). The triangle, circle, and square symbols show, respectively, the data for the ρ_{ex} = 0.2, 0.4 and 0.6 conditions. The brackets show plus and minus one standard error of the mean.

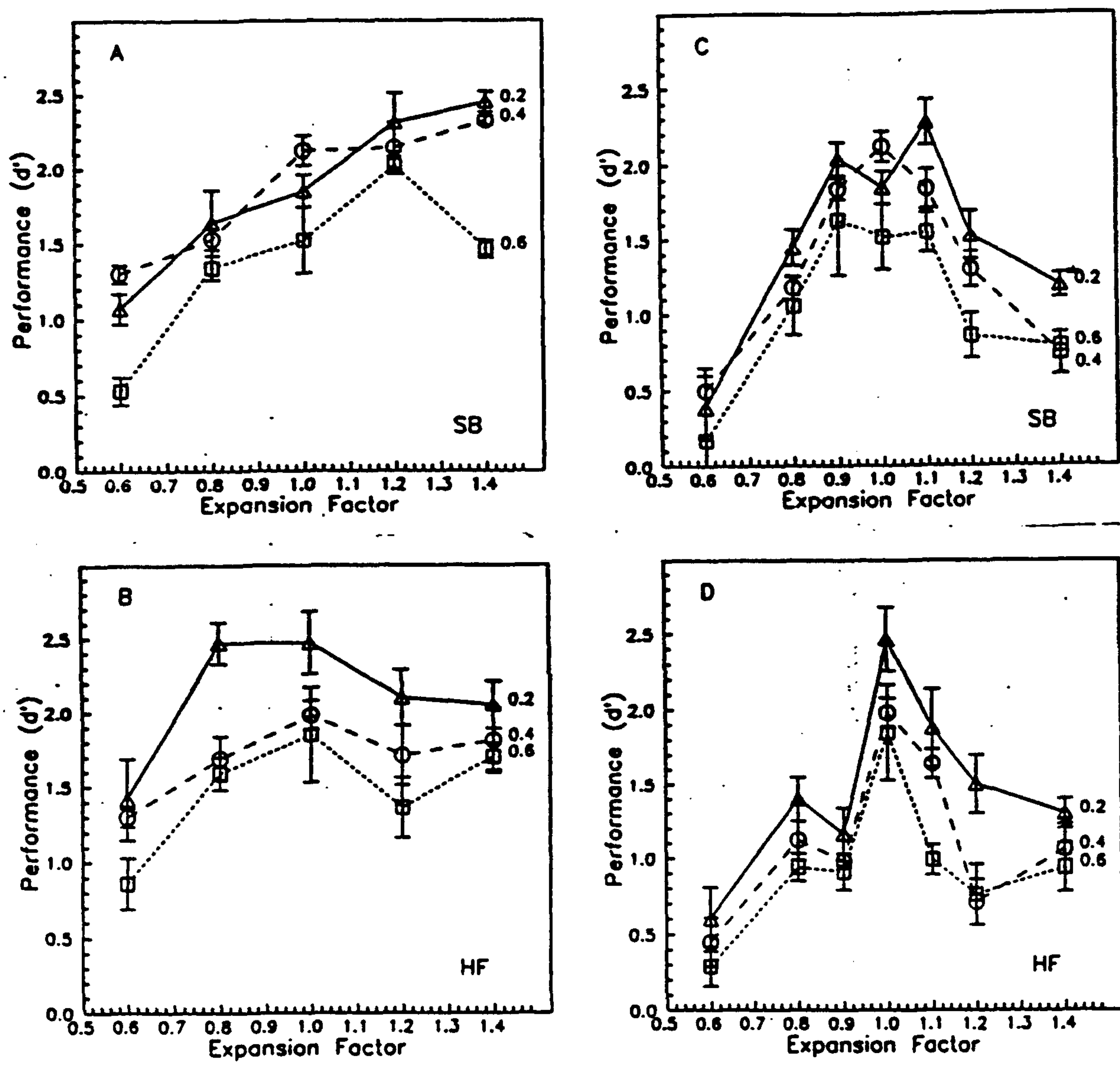

Figure 3. The performance (d') of two listeners (SB: panels A,C; HF: panels B,D) is plotted as a function of the uniform time expansion of both sequences (panels A,B) or the second sequence alone (panels C,D). The triangle, circle, and square symbols show, respectively, the data for the p_{ex} = 0.2, 0.4 and 0.6 conditions. The brackets show plus and minus one standard error of the mean.

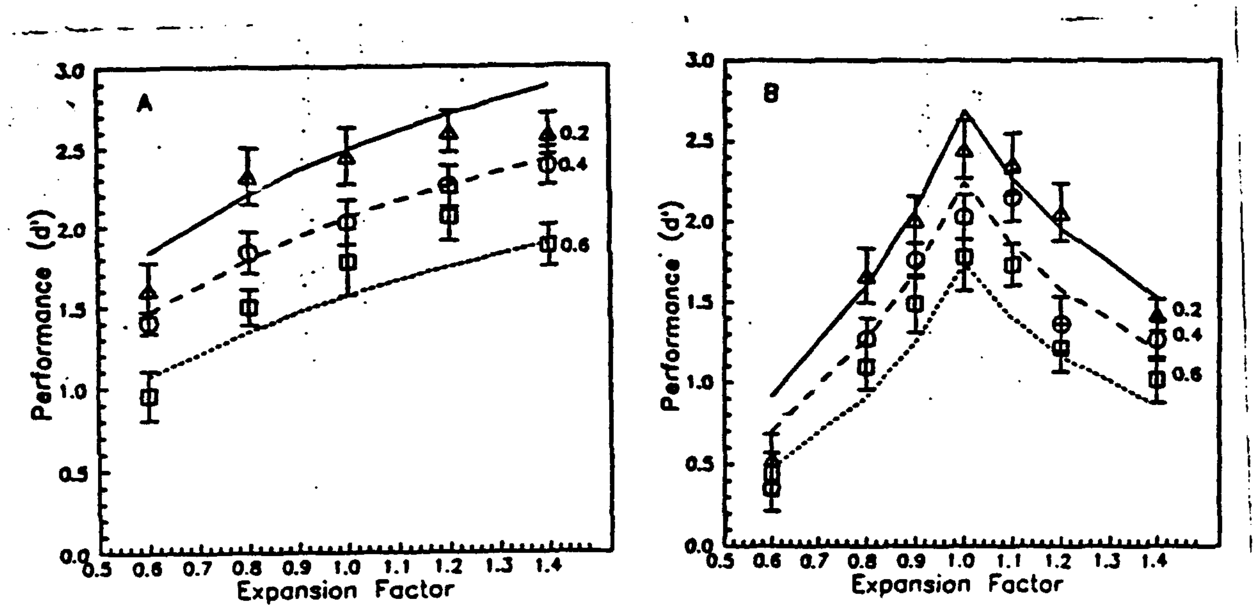

Figure 4. The average performance (d') of four listeners is plotted as a function of the uniform time expansion of both sequences (panel A) or the second sequence alone (panel B). The triangle, circle, and square symbols show, respectively, the data for the ρ_{ex} = 0.2, 0.4 and 0.6 conditions. The brackets show the average standard errors of the mean for the four listeners. The solid and dashed lines in panel A are the predictions of the correlation model fit to the averaged data, assuming $\sigma_{in} = A + B\mu_{IOT}$, where A=6.7373 and B=0.0528 (see text). The solid and dashed lines in panel B are the predictions of the pattern correlation model fit to the averaged data, assuming $\sigma^2_{in} = A + B\mu_{IOT} + C|k-1|$, where A=9.665, B=0, and C=28.93 (see text).

B. Results and Discussion

Figures 2, 3 and 4 show the effects on performance of expanding or compressing the sequences. Figures 2 and 3 show the data for individual listeners; the vertical bars indicate plus and minus one standard error of the mean. Figure 4 shows the average performance of the four listeners; the vertical bars are the average of the standard errors of the four subjects in each condition. The left-hand panels (A, B) of the figures show the results obtained in the control conditions ($k_1 = k_2 = k_m$); the right-hand panels (C, D) show the results obtained in the test conditions ($k_1 = 1$, $k_2 = k_m$). All conditions showed the predicted dependence on ρ_{ex} (Sorkin, 1990).

When both sequences were transformed, performance increased with the magnitude of the expansion; that is, performance improved with expansion ($k_m > 1$) and decreased with compression ($k_m < 1$). This is consistent with the prediction of the pattern correlation model with a fixed internal noise component: expansion of the interonset duration increases the external variance and thus reduces the decorrelating effect of internal noise; compression of the interonset duration decreases the external variance and increases the decorrelating effect of internal noise.

In order to evaluate the correlation model, we postulated an internal noise having both a constant noise component, A, and a Weber's-law component, B:

$$\sigma_{in} = A + B\mu_{IOT} \quad (7)$$

A least-squares fit of equations 4, 5, 6 and 7 to the averaged data from the multiplicative condition yielded the values: A=6.73-ms and B=0.053. The Weber component contributed about 37% of the internal noise. The predictions of the model are shown as the curves on the left-hand panel of figure 4. Except for the 0.4 condition, the fit is not very good, however, the increase of performance with the magnitude of expansion is approximated by the model.

The right-hand panels of the figures show that a different effect was produced by transforming only the second sequence of the pair ($k_1 = 1$, $k_2 = k_m$). Under these conditions, performance was a peaked function of the absolute magnitude of the expansion or compression. Listener performance at compressions of 0.6 were near to the chance level. Attempts to fit the model (not shown in figure 4b) to this data were not satisfactory, so we considered an additional assumption about the nature of the internal noise.

Consider an internal noise having a component that depends on the magnitude of the difference between the transformations to the patterns:

$$\sigma_{in} = A + B\mu_{IOT} + C|k_1 - k_2| \quad (8)$$

where A and B are as previously defined, and where C determines the magnitude of the noise component that is attributable to the absolute difference between the pattern transformations. The solid and dashed lines of figure 4b were generated by fitting the data from the three correlation conditions to a pattern correlation model that incorporated this noise assumption. The resulting parameters were: A = 9.66-ms, B = 0, and C = 28.93. The contribution of the pattern difference factor to the total internal noise ranged from zero at k=1.0, to approximately 60% at k=1.5.

EXPERIMENT 2. EFFECT OF ADDITIVE TRANSFORMATIONS

The purpose of this experiment was to examine the effects of uniform additions or reductions in the interonset durations of the sequences to be discriminated, and to observe any differences between the effect of additive and multiplicative transformation of the sequences.

A. Procedure

In order to test the effects of additive time transformations on discrimination, listeners were run under two experimental conditions: (a) control conditions in which a constant time, t_a, was added to (or subtracted from) the interonset times of both pair of sequences on each trial ($t_1=t_2=t_a$); and (b) test conditions, in which the constant time, t_a, was added to (or subtracted from) the interonset intervals only in the second sequence of the pair on each trial ($t_1=0$, $t_2=t_a$). The test and control conditions were run under one value of pattern correlation ($\rho_{ex} = 0.2$).

The control conditions were run with t_a equal to -15, -7, 0, 7, 15, 30, and 45-ms, and the test conditions were run with t_a equal to -7, 0, 7, 15, 30, and 45-ms. These values were chosen in order to produce the same interonset durations employed in experiment 1, at different values of k_a. That is, these times produced expansions to the interonset times equivalent to, respectively: 0.8, 0.9, 1.0, 1.1, 1.2, 1.4, and 1.6, for the control conditions and 0.9, 1.0, 1.1, 1.2, 1.4, and 1.6, for the test conditions. Because it was impossible to have interonset times smaller than 2-ms, we were concerned that the truncations of the interonset time required by the use of conditions with large negative values of t_a would distort the distributions of $\{X_1\}$ and $\{X_2\}$. Hence, conditions that would have required subtracting a constant time larger than 7-ms in the test (or 15-ms in the control) conditions were avoided. In addition to the additive conditions, the test and control conditions of experiment 1 were repeated with this group of listeners. Thus, a total of four transformation conditions were run:

(1) Multiplicative to both sequences

(2) Multiplicative to the second sequence

(3) Additive to both sequences

(4) Additive to the second sequence.

Three or four blocks of 100 trials were run at each experimental condition. As in experiment 1, the correlation and transformation levels were fixed within each block of trials.

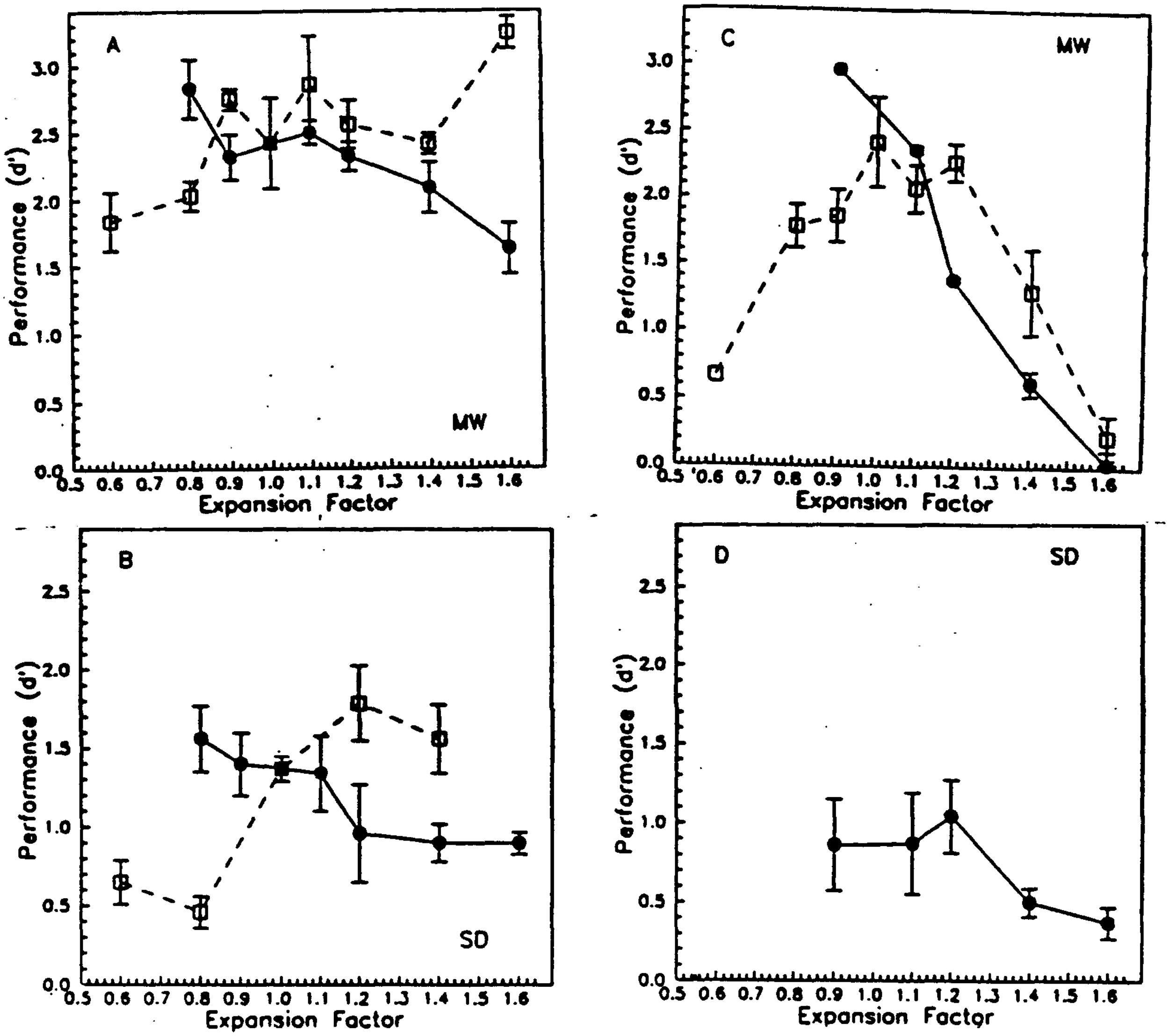

Figure 5. The performance (d') of two listeners (MW: panels A,C; SD: panels B,D) is plotted as a function of the time expansion applied to both sequences (panels A,B) or the second sequence alone (panels C,D). The filled circle symbols (and solid lines) show the data for conditions when the sequences are expanded by adding a fixed time to the interonset times. The open square symbols (and dashed lines) show the data for conditions when the expansion is implemented by multiplication by a constant factor. The brackets show plus and minus one standard error of the mean. Listener SD did participate in the multiplicative, second-sequence-alone condition.

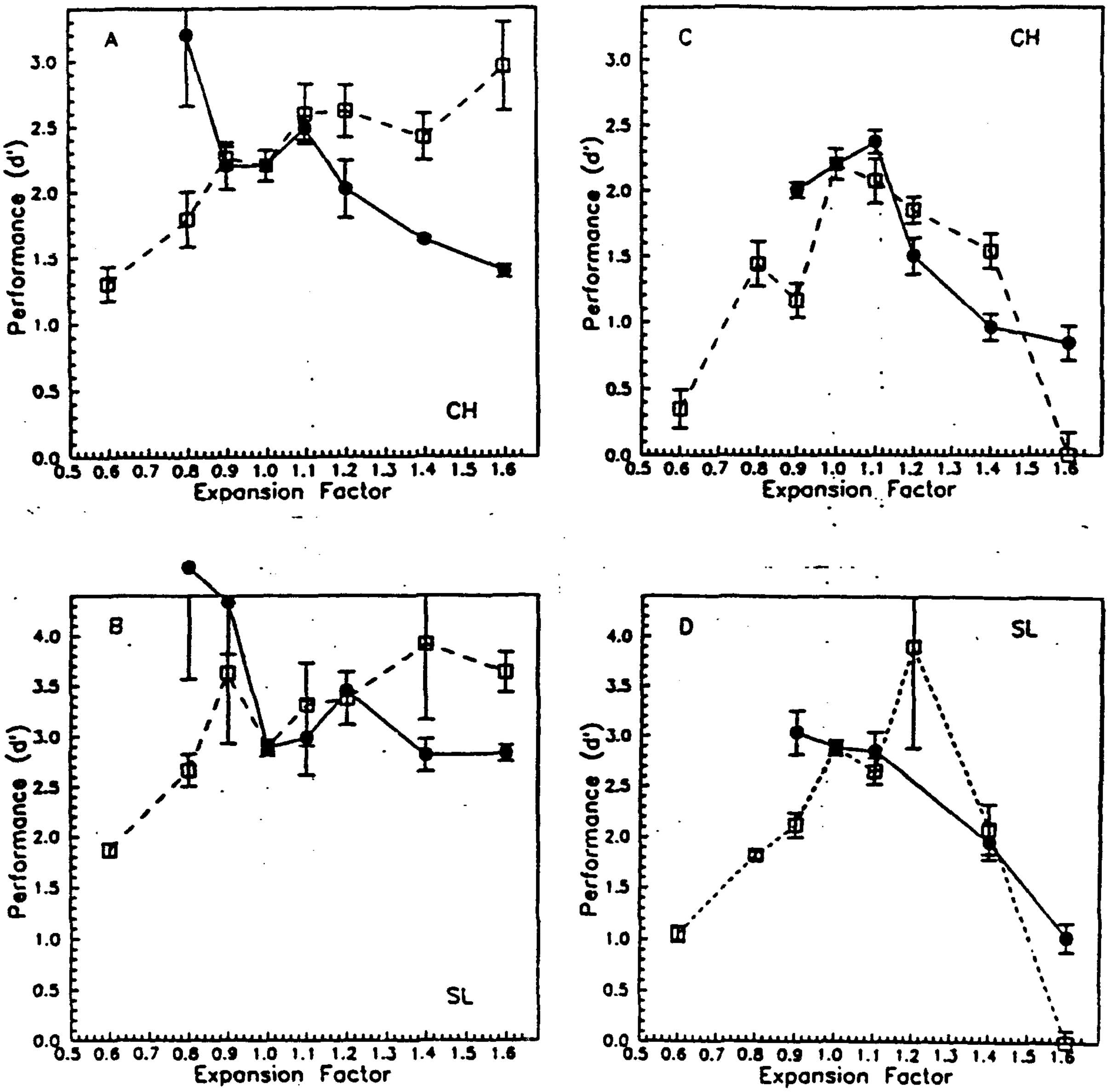

Figure 6. The performance (d') of two listeners (CH: panels A,C; SL: panels B,D) is plotted as a function of the time expansion applied to both sequences (panels A,B) or the second sequence alone (panels C,D). The filled circle symbols (and solid lines) show the data for conditions when the sequences are expanded by adding a fixed time to the interonset times. The open square symbols (and dashed lines) show the data for conditions when the expansion is implemented by multiplication by a constant factor. The brackets show plus and minus one standard error of the mean.

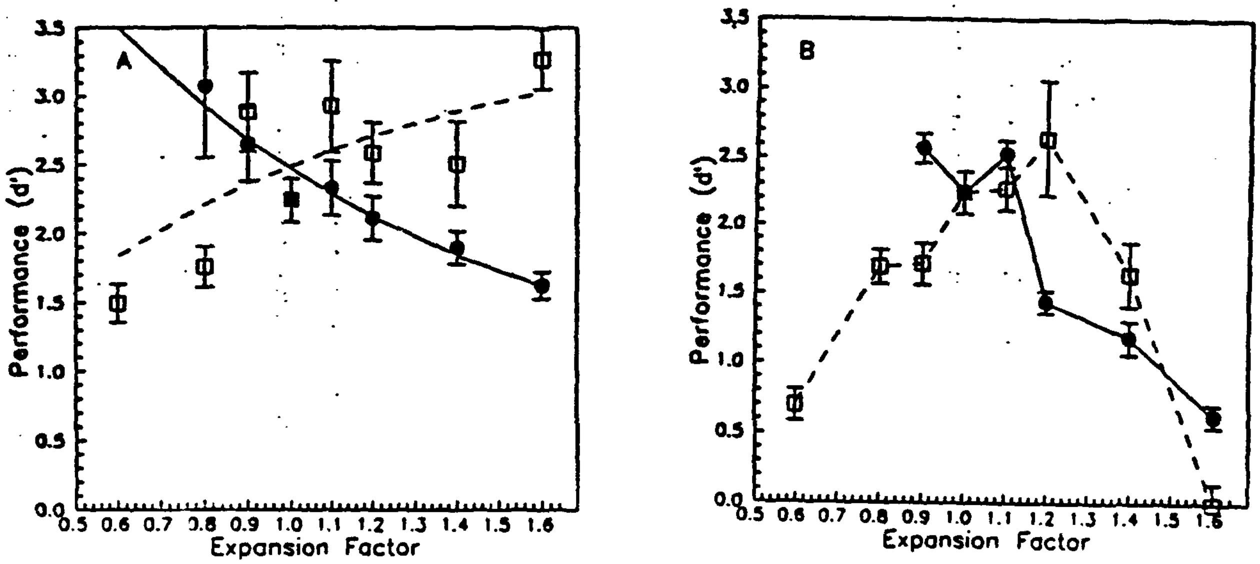

Figure 7. The average performance (d') of four listeners is plotted as a function of the uniform time expansion of both sequences (panel A) or the second sequence alone (panel B). The filled circle symbols show the data for conditions when the sequence times are expanded by adding a fixed time to the sequence times. The open square symbols show the data for conditions when the expansion is implemented by multiplication by a constant factor. The brackets show the average standard error of the mean for the four listeners. The dashed line is the prediction of the correlation model, using the parameters derived from the data of figure 4 (σ_{in}=6.7373+0.0528μ_{IOT}). The solid line is the prediction of the correlation model fit to the averaged additive transformation data (A=0, B=0.144; see text).

B. Results and Discussion

The left-hand panels (A, B) of figures 5 and 6 show the results obtained in the control conditions (multiplying or adding a fixed time to both sequences) on the performance of individual listeners. The right-hand panels (C, D) show the results obtained in the test conditions. The dashed lines (and square symbols) show the data from the multiplicative transformation conditions and the solid lines (and filled circle symbols) show the data from the additive conditions. Figure 7 shows the data averaged over the four listeners (the data in figure 7b are averaged over three listeners).

The data obtained in the multiplicative transform conditions replicated the results obtained in experiment 1, for the control (both-sequences) and test (second-sequence-alone) conditions. In figure 7a, the square symbols are the data points from the multiplicative both-sequences condition. The dashed line in figure 7a is the prediction of the pattern correlation model, using the parameters obtained from the model fit to the data of figure 4a ($\sigma_{in} = 6.7373 + 0.0528\ \mu_{IOT}$).

The filled circle symbols in figure 7a show the data obtained when both sequences received the additive transformation. The additive transformation produced an effect markedly different from that of the multiplicative transformation: instead of an increase in performance, there was a performance decrease of more than one d' unit. Adding a constant time interval increases the average interonset interval without increasing the sequence correlation. Thus, if the internal noise has a Weber's-law component, performance will decrease under (positive) additive transformations.

The solid line of figure 7a is a fit of the model to the additive data; the parameters are: $\sigma_{in} = 0 + 0.144\mu_{IOT}$; in this case, the Weber's law component contributed 100% of the internal noise. We cannot say why the Weber's law contribution dominates the internal noise in this case. Although the variance of the interonset interval is not changed by the additive transformation, the mean time interval between tone markers (offset to onset) and the duty cycle of the tone marker (the duration of the tone relative to the interonset time) are changed. Apparently, both of these variables have an effect on the nature of the internal noise.

When the additive transformation was applied to the second sequence alone, discrimination performance (for three of the four listeners) dropped more than 2 d' units. Since we did not test large negative time intervals, we could not determine whether or not the performance function had a maximum near k=1. As in the multiplicative case, performance decreased as a function of the magnitude of the additive transformation. This performance drop was steeper than that produced when both sequences were transformed.

GENERAL DISCUSSION

When both patterns were transformed by a multiplicative constant, performance increased with the amount of expansion. When both patterns received an additive transformation, performance decreased with the size of the additive constant. These results are generally consistent with a pattern correlation mechanism limited by internal noise. However, our attempt to characterize the performance functions by a single description of the internal noise was not successful. The contribution of the Weber's law noise component was greater in the additive case than the multiplicative case. This difference may be related to the differences between the two conditions in marker duty cycle and mean time interval between markers.

When only the second of the two sequences was transformed, performance in the multiplicative condition was a peaked function of the magnitude of the transformation. The fit of the model to this data was improved by assuming an internal noise component proportional to the magnitude of the transformation difference between the patterns. The existence of an internal noise component of this type implies that there is a processing cost associated with certain differences between the stimuli to be compared. Such costs have been noted in temporal discrimination tasks when the interval markers have different spectral properties (e.g. Divenyi and Danner, 1977; Hirsch et al. 1990; Sorkin et al., 1982) and in intensity discrimination tasks when the two signals are of different frequency (Lim et al., 1977). Some conversion or normalization is required when there are differences between the stimuli that are not relevant to the particular pattern comparison; there may be internal noise associated with the additional processing. It is interesting that in the present case, this cost is approximately a symmetric function of $|k_1-k_2|$.

An alternative explanation is that the listener's use of information from the temporal pattern(s) is not a uniform function of the position of the information within the sequence patterns, as noted recently by Hirsch et al. (1990). The listener may utilize information from certain regions of each stimulus sequence more than from others. Temporal transformation of the sequences may upset this temporal position effect, in that regions of maximum attention in two differentially transformed sequences, no longer coincide. The transformation results in a misalignment of the sequence weighting functions.

The current experiments indicate that the ability to normalize time is somewhat limited. Is the observed sensitivity to time scaling inconsistent with our expectations about rate normalization in speech perception? This question involves the complex issue of the nature of the rate normalization mechanism in speech (e.g., see Diehl and Walsh, 1989; and Pisoni et al., 1983). We restate two hypotheses that are relevant to the question: First, it is possible that a listener can implement an efficient time re-scaling process only for speech-like signals. That is, performance with random tonal sequences might be improved if the listener somehow could be induced to process the inputs as if they were speech signals. The second hypothesis is that the listener's use of relative timing information is no more efficient in the sequence experiments than it is for speech signals--but that the speech signal provides a richer source of time-scaling information that can be used to augment the basic timing data. Some of this information is carried by the higher order structure of the speech signal.

Finally, it is tempting to try to generalize the results of the present experiments to the case of repeated sequences. The listener's ability to discriminate between two rhythmic patterns may parallel the ability to discriminate between the patterns played singly. However, caution is advised. Although the pattern correlation hypothesis may be related to the perception of rhythmic stimuli, the present stimuli are not rhythmic (or metric). Repetition is generally considered to be a necessary condition for rhythmic perception (Handel, 1989; Sturges and Martin, 1974). Rhythmic percepts are said to emerge from the acoustic (and subjective) context of repetitive stimuli and act to segment and organize stimuli (Handel, 1989). Studies of rhythm and meter generally have been confined to temporal patterns that are repetitive. In our experiments, there was no repetition of the random temporal pattern on DIFFERENT trials, and there was only a single repetition of the stimulus pattern on SAME trials. An interesting question is whether the present results with time transformations (and those reported previously by Sorkin, 1990) will hold for repetitive patterns.

III. EFFECT OF INTERSEQUENCE DELAY INTERVAL ON THE DISCRIMINATION OF TONAL PATTERNS.

ABSTRACT

According to the Pattern Correlation model [Sorkin, J. Acoust. Soc. Am., 87, (1990)], listeners discriminate between arrhythmic tonal sequences by computing the correlation between the serial pattern of time intervals marked by the tones in each sequence. The present experiments evaluated discrimination when the sequences were presented at different frequencies and to different ears. The sequences began at delayed starting times and were subject to random time expansions. When the delay between sequence onsets was less than 10-ms, discrimination appeared to be based on comparison of the envelope of the stimulus waveforms. At longer time separations, however, performance was consistent with the Pattern Correlation hypothesis.

INTRODUCTION

In Studies I and II (Sorkin, 1990; Sorkin and Montgomery, submitted), we proposed a Pattern Correlation model of how listeners discriminate between the temporal patterns formed by two, arrythmic tonal sequences. The primary assumption of the model is that a listener discriminates differences between the two patterns by estimating the correlation between the serial (temporal) structure of the patterns. The present experiments compared the predictions of the pattern correlation model and an alternative, the waveform correlation model, when the tonal sequences were presented at different frequencies and to different earphone channels.

The basic experimental paradigm is the same as in the previous studies. The listener is presented with two, successively played, arrhythmic tonal sequences. The series of time intervals between tone onsets in each sequence define the two temporal patterns to be discriminated. On half of the trials these two patterns are identical, and on half of the trials the temporal patterns are different; the listener must report whether the patterns were the same or were different. The experiment variable is the correlation, ρ_{ex}, between the sequences on trials when the sequences are different; the task is easiest when ρ_{ex} equals 0 and increases in difficulty as ρ_{ex} approaches one.

A. Comparison of Pattern Correlation Model and Waveform Correlation Model

The basic assumption of the Pattern Correlation model is that the listener estimates the correlation between temporal patterns by computing the Pearson product-moment correlation coefficient, r_{12}, on the transduced and encoded series of marker interonset times. The performance of the listener is given by equations A7, A21 and A22 in Section V . Transformations or distortions of the tonal sequences, such as time expansions or compressions, should affect discrimination to the extent that the transformations produce differences in the listener's estimate of the correlation.

Study II tested the effects of constant additive and multiplicative transformations to the sequence time scales. In those experiments all tones were 1000 Hz and the sequences were presented monaurally, at a time separation of either 750-ms or 825-ms. Performance decreased when one of the sequences was compressed or expanded in time. The decrement was a function of the magnitude of the discrepancy in time compression between the two sequences; the amount of the performance drop ranged from 0 to 2 d' units over a range of compressions of from 0.6 to 1.6. Adding an internal noise component proportional to the absolute magnitude of the transformation difference, enabled the pattern correlation model to describe the obtained data.

The major assumption of the pattern correlation model is that the listener encodes and processes a list of interonset times from each sequence; the listener discards other information about the sequence waveforms, such as the absolute timing of signals or the signals' spectra. An alternative to this mechanism is a comparison process based on cross-correlation of the two sequence waveforms or their envelopes. A waveform correlation process can provide a very sensitive measure of the difference between the waveforms (or waveform envelopes) of the two sequences.

Because it involves a point-for-point comparison of the sequence waveforms, a waveform correlator may be very sensitive to time transformations such as compression or expansion. Such transformations would result in temporal misalignments of the patterns. Temporal misalignments that occur early in the sequences would produce even greater decorrelations between the sequences at later times. As a consequence, the performance of a waveform correlator may be seriously degraded by time transformations.

The correlator could deal with random (but uniform) temporal transformations made to one of the two patterns, by computing a correlation function. For example, a number of different expansions, τ, could be applied to the first sequence, and then each transformed sequence could be correlated with the second sequence. Provided that the sequences were correlated, the resulting function would yield a well-defined peak at the value of τ that matched the time transformation to the second sequence.

Under some conditions, a listener may use the binaural system to compute the cross-correlation between two input waveforms or their envelopes. For example, in many binaural detection situations, the auditory system behaves as though it computes a correlation between the inputs to the respective earphone channels (Colburn and Durlach, 1978). In general, the binaural mechanism can be used if the sequences are presented separately and almost simultaneously, to the two ears. Under these conditions, the listener can make very precise determinations of differences between the sequences. A delay in the second signal of longer than about 15 milliseconds would be expected to exceed the limits of this system (Bilsen and Goldstein, 1974). In addition, there is evidence that binaural comparisons can be performed when the signals are at high (and different) frequencies in the two ear channels (McFadden and Pasanen, 1974, 1975, 1978).

Suppose that one of two stimulus sequences to be compared has been transformed by a uniform time compression or expansion. We would expect that such a transformation would produce a percept similar to that produced by stimulating each ear with uncorrelated noise, e.g.: the lower the correlation between the signals, the more spatially diffuse will be the percept (the higher the correlation, the more spatially focused). Although the effect of such transformations to one of two binaural inputs has not been tested directly, it seems clear that the system will not be capable of forming a spatially focused percept. Basically, the system would be presented with two sequences composed of sinusoid pulses of different frequency in each channel. The only time-coherent aspect of this stimulus would be the onset time for the first tone marker. The onset time for the first tone would be coherent whether or not the temporal pattern was the same or different on a trial. Thus, we would not expect the binaural comparison mechanism to be able to accomodate compressive or expansive transformations to one of the two patterns to be compared; under those conditions, performance on sequence comparison tasks should be adversely affected.

B. Experimental Plan

The present experiment combines conditions in which: (a) the sequences are presented at different intersequence delay intervals (ISIs), and (b) the second sequence is temporally compressed or expanded. Three general factors should affect performance when the delay of the sequence starting times or intersequence delay interval is manipulated: (1) mechanism, (2) masking, and (3) memory.

The putative effects of mechanism on performance at different intersequence intervals have already been mentioned. Short intersequence intervals (less than 20-ms) should allow operation of the binaural comparison mechanism; long intervals will preclude operation of the binaural comparison mechanism, but still allow operation of the pattern correlation comparator. The second factor, masking, involves energetic masking (and related interference effects) between two signals presented to the auditory system at short time separations. In the current experiment, the signals will be presented to different earphone channels and at different frequencies (in different critical bands). This mode of presentation should minimize the effects of peripheral sensory interference between the signals (see Sorkin, 1966). Finally, memory decrements related to the storage of information should increase as the ISI increases (Sorkin, 1982), but should be minimal at ISI's of less than five hundred milliseconds.

Experiment 1 tests whether sequence pattern discriminations are feasible when the sequences are presented to different auditory channels and at different frequencies. Experiment 1 also evaluates the effects of short and long sequence delays. Experiment 2 evaluates the interacting effects of intersequence delay and random temporal transformations on discrimination performance. Experiment 3 evaluates the interacting effects of temporal transformation and frequency uncertainty.

METHOD

Two groups of subjects participated in these experiments. The first group consisted of one male and three females; the second consisted of two of the original females plus two new female subjects. All subjects were undergraduate students at the University of Florida. They were paid an hourly wage plus an incentive for correct responses. Listeners had normal hearing and performed the tasks for approximately 2 h per day, 3 days per week. Listeners were seated in a double-walled acoustically insulated chamber. The stimuli were presented dichotically via TDH-39 headphones. The conditions were tested in blocks of 100 trials; typically, 8 blocks were completed in a session. Except in the uncertain duration conditions of experiment 2, all independent variables were held constant within a block of trials. Full feedback about the correct response was provided after each trial.

The subjects compared pairs of tone sequences composed of 8 sinusoidal bursts of nominal durations of either 25-ms (in experiment 1) or 30-ms (in experiment 2). After listening to each pair of sequences, the subject had to indicate whether or not the temporal pattern of tone and intertone intervals was the same or different for the two sequences. On a random half of the experimental trials, the temporal patterns were the same, e.g. $\rho_{ex} = 1.0$. On half of the trials the patterns were different, e.g. ρ_{ex} was equal to: 0.2, 0.4, or 0.6, for a block of 100 trials in experiment 1 and 0, 0.4, or 0.8 for blocks of trials in experiment 2 or 3. The tone bursts in the first sequence were at 1000 Hz and approximately 71 dBA SPL, and the tones in the second sequence were at 2300 Hz and approximately 68 dBA SPL. All tone bursts were shaped by a 4-ms linear rise and decay envelope. The first sequence was always directed to the left headphone and the second to the right. The onset (first marker tone) of the second sequence was presented at delays (ISIs) of from 0 to 2.5 seconds, relative to the onset of the first marker tone of the first sequences.

The time intervals between tones were generated by a process that enabled experimenter control of the statistics of the temporal pattern: the mean and standard deviation of the intertone interval, and the correlation, ρ_{ex}, between the patterns. The process is described in Sorkin (1990) and is summarized in Section V. The nominal mean time gap between tones was 50 ms and the nominal standard deviation of this gap was 25-ms or 20-ms; gap durations of less than 2 ms were not allowed.

EXPERIMENT 1. EFFECT OF TWO-CHANNEL PRESENTATION

The purpose of the first experiment was to examine how sequence discrimination depended on the intersequence delay interval between the starting times of the pair of sequences. In addition, we wished to extend the sequence discrimination paradigm to the case when the sequences were presented at different frequencies and to different ears.

A. Procedure

In order to compare sequence discrimination performance at short delays and when the patterns overlapped in time, the sequences were presented to different earphone channels and at different frequencies. The goal was to minimize the sensory interference between the channels at short time separations (see Sorkin, 1965). The beginning tone of the second sequence occurred either 0, 10, 20, 50, 300, 635, 725, 875, or 1375 ms after the beginning tone of the first sequence.

B. Results and Discussion

Figure 8 illustrates the effect of the delay and two-channel manipulation on performance. The four panels of the figure show the performance of four subjects; the vertical bars are the standard errors of the mean. The circles, squares, and triangle symbols show performance at the ρ_{ex}=0.2, 0.4, and 0.6 conditions, respectively. Figure 9 shows the average performance of the four subjects; the vertical bars are the average of the standard errors of the four subjects in each condition. The individual subject plots highly resemble the average data.

Performance was best at ρ_{ex}=0.2, and lowest at ρ_{ex}=0.6, as in the previous study (Sorkin, 1990). Performance at an intersequence interval of zero was high, decreasing with increasing intersequence delays. Performance was quite poor at ISIs of 50-ms and 300-ms and at ρ_{ex}=0.6. Performance increased as the delay increased from 300-ms to 875-ms.

The good performance at short delays was consistent with the operation of either a binaural cross-correlator or a temporal pattern correlator. Since the former mechanism is not available at long delays, performance at the 875-ms (and longer) conditions is consistent with that for the pattern correlator mechanism. The poorest performance was at delays of 50-ms and 300-ms, corresponding to temporal overlaps of the two sequences of 8% and 50%. To summarize: (1) either mechanism can describe performance at pattern overlaps of more than 92%, (2) the pattern correlator can describe performance at zero overlaps (when the binaural comparator cannot operate), and (3) neither mechanism is effective at overlaps of between 8 and 50%.

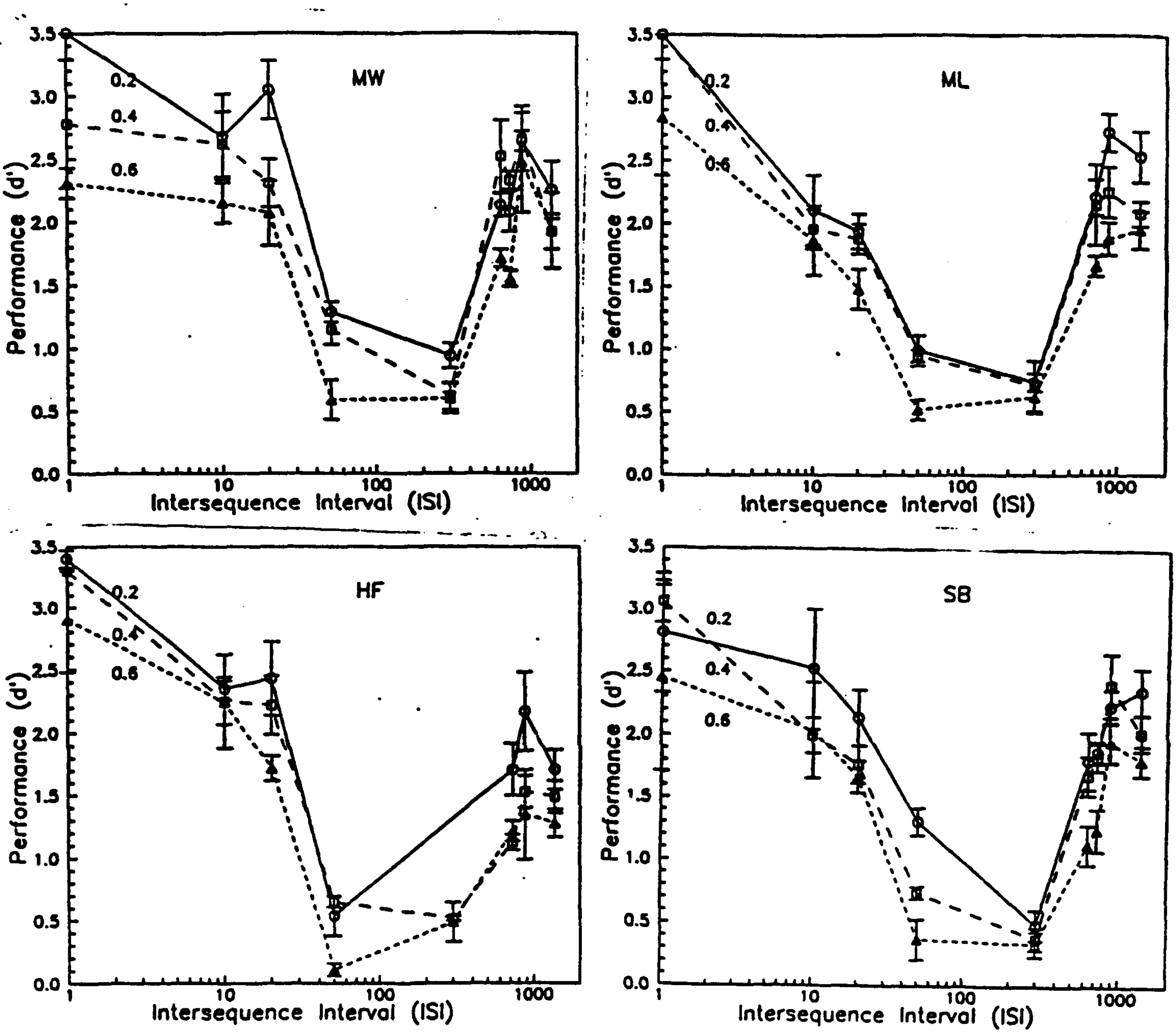

Figure 8. The performance (d') of four listeners is plotted as a function of the Intersequence Interval, the time interval between the onsets of the first marker tones in each sequence. The circle, square, and triangle symbols show, respectively, the data for the p_{ex} = 0.2, 0.4 and 0.6 conditions. The brackets show plus and minus one standard error of the mean.

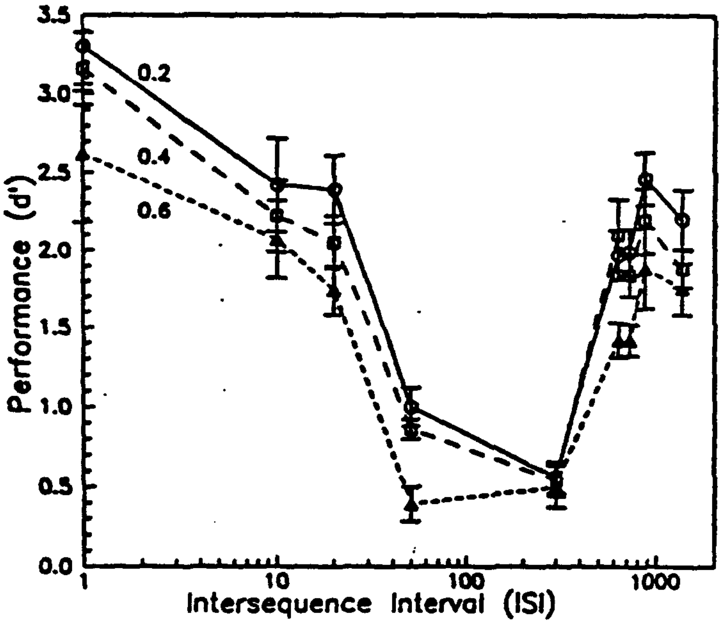

Figure 9. The average performance (d') of four listeners is plotted as a function of the Intersequence Interval, the time interval between the onsets of the first marker tones in each sequence. The circle, square, and triangle symbols show, respectively, the data for the p_{ex} = 0.2, 0.4 and 0.6 conditions. The brackets show the average standard error of the mean for the four listeners.

EXPERIMENT 2. INTERACTION OF INTERSEQUENCE DELAY AND TEMPORAL TRANSFORMATION.

The purpose of this experiment was to evaluate the pattern comparison mechanisms operating at long and short intersequence delay intervals by examining (1) the interaction between the sequence delay and time transformation manipulations at short delays, and (2) the decrease in performance at delays much longer than in the previous experiment. The results of experiment 1 indicated that either comparator mechanism could describe performance at short intersequence delays. A temporal manipulation was added that we believed would interfere with operation of one of the two putative comparison mechanisms. The manipulation was a random temporal transformation to the second of the two sequences, similar to that described in Study I. This transformation was a uniform compression or expansion of all of the times (marker tones and gaps) comprising the second sequence.

A. Procedure

Experiment 2 was similar to Experiment 1, except that an additional manipulation on the sequences was performed. This manipulation multiplied all time intervals in the second sequence by a constant, i.e. all marker tone durations and intertone gaps were expanded by a factor of 0.8, 0.9, 1.0, 1.1 or 1.2. This factor was uniformly applied to all the time intervals within a single sequence, but could vary randomly over the experimental trials. The manipulation also had the effect of modifying the standard deviation of the intertone durations. The probability of a particular one of the transformations being chosen, was 0.2. As in Experiment 1, the subject was required to indicate whether the temporal pattern of tones was the same or different, whether or not the overall tempo of the pattern had been scaled faster or slower by the time transformation. In Experiment 2 the beginning tone of the second sequence occurred either 10, 350, 900, or 2500-ms, after the beginning tone of the first sequence (and independent of the temporal transformation on the second sequence).

B. Results and Discussion

Figure 10 shows the effects of the delay and expansion manipulation on performance for the ρ_{ex}=0 conditions. The four panels of the figure show the performance of four subjects; the vertical bars are the standard errors of the mean. The circles symbols show performance under no time transformation, and the triangles show performance under random time transformations of the sequences patterns. The average data for the four subjects is shown in the three panels of figure 11. These show the averaged data for the ρ_{ex}=0, 0.4, and 0.8 conditions, respectively. The vertical bars are the average of the standard errors of the four subjects in each condition. The individual subject plots are highly similar to the average data, and the data obtained under different values of ρ_{ex} are also quite similar.

As in experiment 1, performance was best at the lowest values of ρ_{ex} and at an ISI of 350-ms. Performance at an intersequence delay of 10-ms was high, decreasing with increasing delays. Performance was quite poor at a delay of 350-ms and increased as the delay increased to 900-ms and then decreased somewhat at 2500-ms.

It is clear that the addition of the temporal manipulation caused performance to drop to the lowest levels. Performance at high ISIs however, was relatively unaffected by the manipulation. Thus, it is reasonable to conclude that the pattern correlator mechanism is much less sensitive to the time transformation manipulation, a result consistent with the previous experiment (over the current range of time compression). At short ISIs, however, the effect of the time transformation is large. The results suggest that waveform correlation is the active mechanism at short ISIs and that it is sensitive to the temporal manipulation. This conclusion is consistent with our expectations about the binaural comparator and its probable sensitivity to temporal manipulations that disturb the coherence of the patterns to be compared.

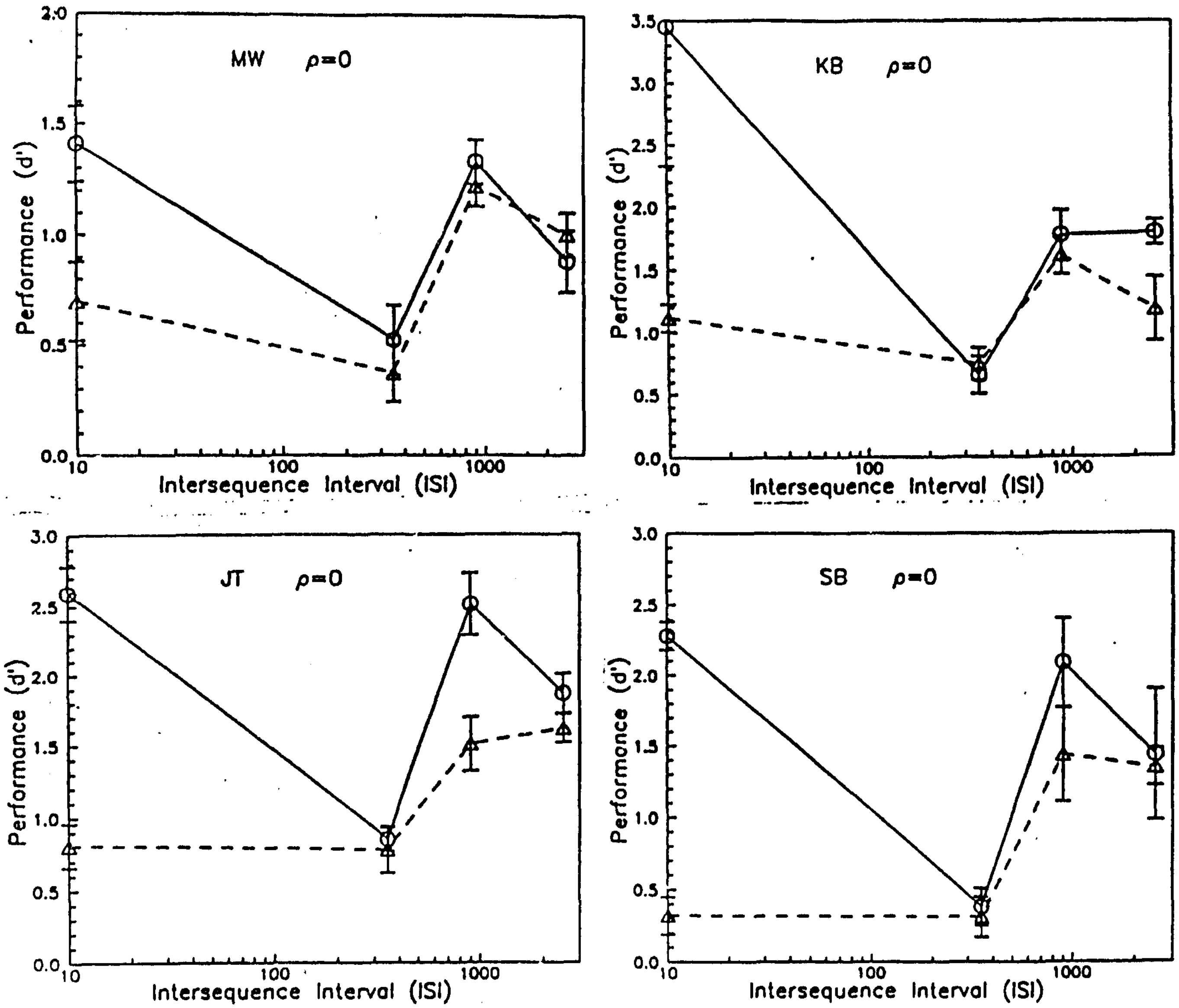

Figure 10. The performance (d') of four listeners is plotted as a function of the Intersequence Interval, the time interval between the onsets of the first marker tones in each sequence, for the p_{ex} = 0 condition. The circle and square symbols show, respectively, the data obtained in the no-transformation and time-transformation conditions. The brackets show plus and minus one standard error of the mean.

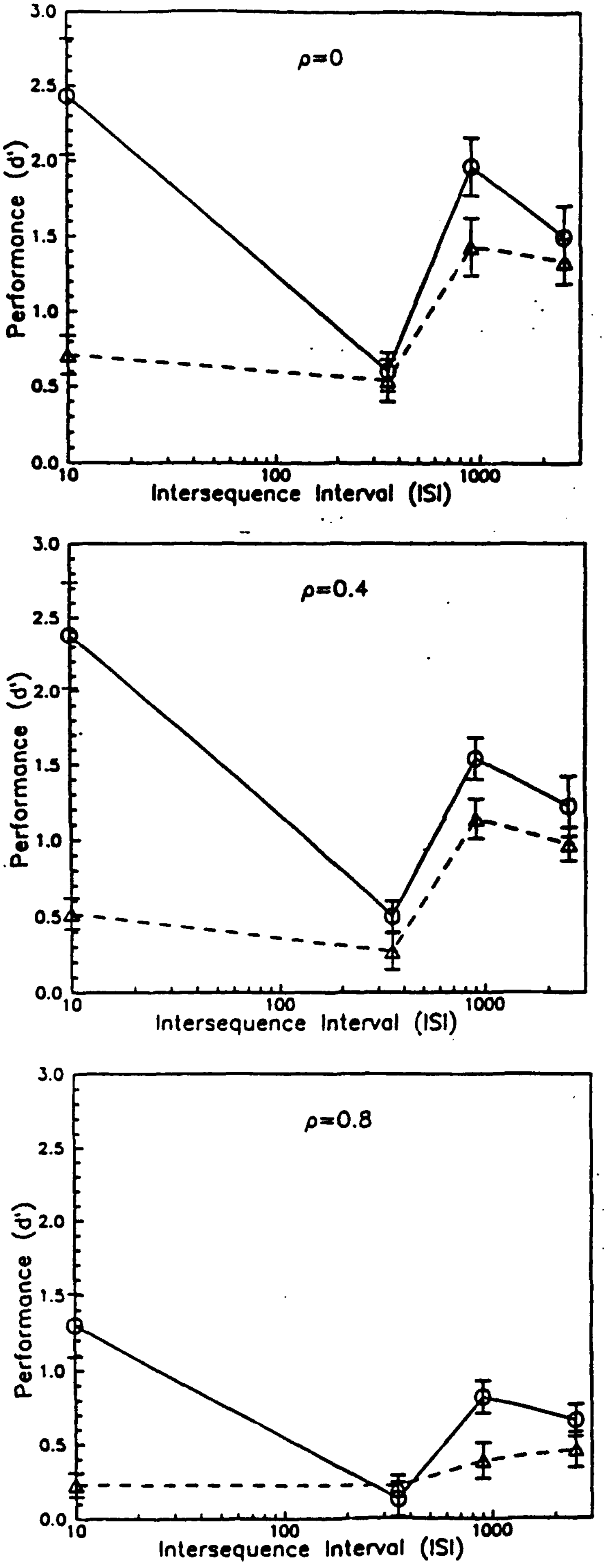

Figure 11. The average performance (d') of four listeners is plotted as a function of the Intersequence Interval, the time interval between the onsets of the first marker tones in each sequence, for the ρ_{er} = 0, 0.4 and 0.8 conditions. The circle and square symbols show, respectively, the data obtained in the no-transformation and time-transformation conditions. The brackets show the average standard error of the mean for the four listeners.

EXPERIMENT 3. INTERACTION OF TEMPORAL TRANSFORMATIONS AND FREQUENCY UNCERTAINTY.

The purpose of this experiment was to examine the possible interactions between the effects of some spectral and temporal manipulations to the stimulus patterns.

A. Procedure

The experimental procedure was the same as in Experiment 2 except that (1) only an ISI of 900-ms was employed, (2) the intertone standard deviation was 20-ms. An additional spectral manipulation, was employed: in this condition the frequency of the sinusoidal marker tones comprising each sequence was randomly varied. Instead of the sequences presented to the left and right earphone channels always being composed of 1000 Hz and 2300 Hz marker tones, respectively, the frequency of each marker tone was randomly set. Although the pattern of frequencies forming each sequence varied over trials, the particular random binary sequence was repeated in both sequence within a trial. There were four experimental conditions:

(1) Conditions same as Experiment 1, no time transformations;
(2) Conditions same as Experiment 1, with time transformations;
(3) Random (binary) pattern of marker frequencies, No time transformations; and
(4) Random (binary) pattern of marker frequencies, with time transformations.

B. Results and Discussion

Figure 12 shows the average data obtained from four subjects in the experiment; the three panels show the data for the ρ_{ex}=0, 0.4, and 0.8 conditions, respectively. The vertical bars show the average of the standard errors of the four subjects. The FXD condition represents the <u>f</u>i<u>x</u>e<u>d</u> frequency manipulation in condition 1 and 2, while the RFA (for <u>r</u>andom <u>f</u>requency <u>a</u>cross trials) indicates the frequency manipulation in conditions 3 and 4. There was little or no interaction between the spectral and temporal manipulations.

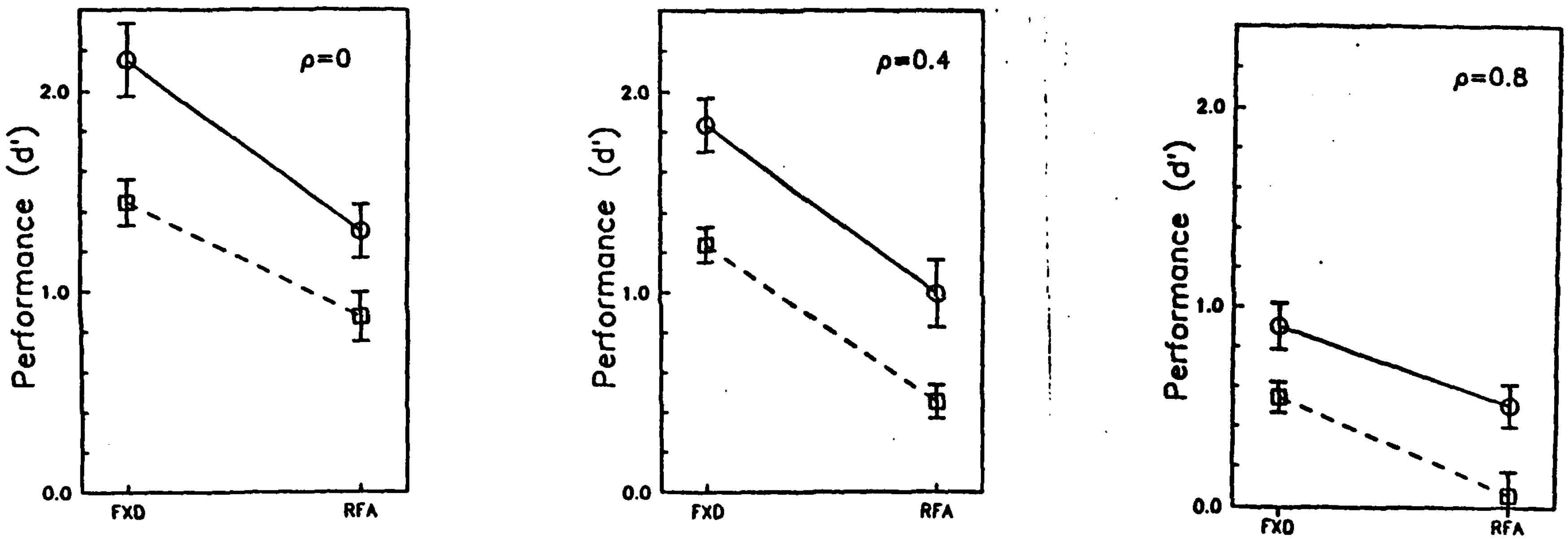

Figure 12. The average performance (d') of four listeners is plotted as a function of the spectral manipulation; FXD: no frequency variability, or RFA: random frequency across trials, in the ρ_{ex} = 0, 0.4 and 0.8 conditions. The circle and square symbols show, respectively, the data obtained in the no-transformation and time-transformation conditions. The brackets show the average standard error of the mean for the four listeners.

GENERAL DISCUSSION

The results of experiment 1 indicated that listeners could discriminate between the temporal patterns, even when the patterns contained tones of different frequency and were presented to separate ears. Performance was good when the sequences were presented either at very short or very long time delays. The results at very short delays were consistent with the predictions of both the pattern correlation model and the waveform correlation model. Since it is unlikely that the binaural correlation mechanism can function when the intersequence delay exceeds 20-ms, . . pattern correlation is the model of choice for long delay conditions.

Discrimination performance was poorest when the onset delay exceeded 20-ms and the sequences overlapped in time. Why is performance so poor when the sequences overlap in time? It is possible that the pattern correlation mechanism can function effectively only for sequentially presented stimuli; sequential processing may be a necessary condition for this mechanism to operate.

Experiment 2 added an additional condition of temporal compression or expansion, enabling differentiation of the effects of the intersequence delay interval on performance. The major effect of this temporal manipulation was in the 10-ms condition, where performance decreased greatly. This time manipulation was expected to adversely effect the performance of the (binaural) waveform correlator. Therefore, when considered together the results support a two-phase mechanism: when the intersequence onsets are less than 20-ms, the binaural correlator is the active mechanism; when the intersequence onsets are greater than 20-ms, the pattern correlator is the active mechanism.

In experiment 3, an additional condition was tested in which the tone frequencies were randomly varied within each sequence. The same pattern of tone frequencies was present in each of the pair of sequences on a trial. This spectral manipulation produced a small drop in performance. The effects of the spectral and temporal manipulations did not interact. That is, the addition of spectral uncertainty did not potentiate the effect of temporal compression. This result is consistent with the hypothesized pattern correlation mechanism, in which the listener extracts temporal information from the sequences and discards the spectral information. However, a more convincing demonstration of the independence of spectral and temporal manipulations would be to randomize the tone frequencies within, as well as across, trials. Some conditions of this type were run, but the subjects found this condition exceedingly difficult. For most of the subjects, performance was near chance; time constraints prevented running a sufficient number of trials to draw any conclusions from these conditions.

IV. REFERENCES

Abel, S. (1972a). Duration discrimination of noise and tone bursts. J. Acoust. Soc. Am., 51, 1219-1223.

Abel, S. (1972b). Discrimination of temporal gaps. J. Acoust. Soc. Am., 52, 519-524.

Allen, L. G. (1979). The perception of time. Perception & Psychophysics, 26, 340-354.

Bharucha, J. J. and Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. Perception & Psychophysics, 40, 137-141.

Bilsen, F. A. and Goldstein, J. L. (1974). Pitch of dichotically delayed noise and its possible spectral basis. J. Acoust. Soc. Am., 55, 292-296.

Bregman, A. S. (1990). Auditory Scene Analysis. The Perceptual Organization of Sound. Cambridge, Mass: MIT Press.

Bregman, A. S. and Campbell, J. L. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. Journal of Experimental Psychology, 89, 244-249.

Bregman, A. S. and Dannenbring, G. (1973). The effect of continuity on auditory stream segregation, Perception & Psychophysics, 13, 308-312.

Brunk, H. D. (1960). An Introduction to Mathematical Statistics. Boston: Ginn.

Colburn, H. S. and Durlach, N. I. 1978). Models of binaural interaction. Chapter 11 in E. C. Carterette and M. P. Friedman (Eds.) Handbook of Perception, Volume IV, New York, N.Y.: Academic Press.

Creelman, C. D. (1962). Human discrimination of auditory duration. J. Acoust. Soc. Am., 34, 582-593.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. Perception & Psychophysics, 28, 381-389.

Diehl, R. L. and Walsh, M. A. (1989). An auditory basis for the stimulus-length effect in the perception of stops and glides. J. Acoust. Soc. Am., 85, 2154-2164.

Divenyi, P. L. and Danner, W. F. (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. Perception & Psychophysics, 21, 125- 142.

Divenyi, P. L. and Sachs, R. M. (1978). Discrimination of time intervals bounded by tone bursts. Perception & Psychophysics, 24, 429-436.

Espinoza-Varas, B. and Jamieson, D. G. (1984). Integration of spectral and temporal cues separated in time and frequency. J. Acoust. Soc. Am., 76, 732-738.

Espinoza-Varas, B. and Watson, C. S. (1986). Temporal discrimination for single components of nonspeech auditory patterns. J. Acoust. Soc. Am., 80, 1685-1694.

Gabrielsson, A. (1973). Similarity ratings and dimension analyses of auditory rhythm patterns. I. Scand. J. Psychol., 14, 138-160.

Getty, D. J. (1975). Discrimination of short temporal intervals: A comparison of two models. Perception & Psychophysics, 18, 1-8.

Handel, S. (1989). Listening: An introduction to the perception of auditory events. Cambridge, Mass: MIT Press.

Halpern, A. R. and Darwin, C. (1982). Duration discrimination in a series of rhythmic events. Perception & Psychophysics, 31, 86-89.

Hirsch, I. J., Monahan, C. B., Grant, K. W., and Singh, P. G. (1990). Studies in auditory timing: 1. Simple patterns. Perception & Psychophysics, 47, 215-226.

Jeffress, L. A. and Robinson, D. E. (1962). Formulas for the coefficient of interaural correlation for noise, J. Acoust. Soc. Am., 34, 1658-1659.

Jones, M. R., Boltz, M., Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. Perception & Psychophysics, 32, 211-218.

Jones, M. R., Kidd, G., and Wetzel, R. (1981). Evidence for rhythmic attention. Journal of Experimental Psychology: Human Perception and Performance, 7, 1059-1073.

Kidd, G. R. and Watson, C. S. (1988). Detection of changes in frequency- and time-transposed auditory patterns, J. Acoust. Soc. Am., 84, S141-142.

Lim, J. S., Rabinowitz, W. M., Braida, L. D., and Durlach, N. I. (1977). Intensity perception. VIII. Loudness comparisons between different types of stimuli. J. Acoust. Soc. Am., 62, 1256-1267.

Lunney, H. W. M. (1974). Time as heard in music and speech. Nature, 249, 592.

McFadden, D. and Pasanen, E. G. (1974). High-frequency masking-leveldifferences with narrow-band noise signals. J. Acoust. Soc. Am., 56, 1226-1230.

McFadden, D. and Pasanen, E. G. (1975). Binaural beats at high frequencies. Science, 190, 394-396.

McFadden, D. and Pasanen, E. G. 1978). Binaural detection at high frequencies with time-delayed waveforms. J. Acoust. Soc. Am., 63, 1120-1231.

Miller, J. L. (1987). Rate-dependent processing in speech perception. Chapter 3 in A. W. Ellis (Ed.) Progress in the Psychology of Language, (pp. 119-157). Hillsdale, NJ:Erlbaum.

Miller, J. L. and Dexter, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. Journal of Experimental Psychology: Human Perception and Performance, 14, 369-378.

Miller, J. L. and Grosjean, F. (1981). How the components of speaking rate influence perception of phonetic segments. Journal of Experimental Psychology: Human Perception and Performance, 7, 208-215.

Miller, J. L. and Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semi-vowel. Perception & Psychophysics, 25, 457-465.

Miller, J. L. and Volaitis, L. E. (1989). Effects of speaking rate on the perceptual structure of a phonetic category. Perception & Psychophysics, 46, 505-512.

Monahan, C. B. and Hirsh, I. J. Studies in auditory timing: 2. Rhythm patterns. (1990). Perception & Psychophysics, 47, 227-242.

Monahan, C. B., Kendall, R. A., and Carterette, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. Perception & Psychophysics, 41, 576-600.

Pisoni, D. B., Carrell, T. D., and Gans, S. J. (1983). Perception of the duration of rapid spectrum changes in speech and nonspeech signals. Perception & Psychophysics, 34, 314-322.

Pollack, I. (1967). Asynchrony: The perception of temporal gaps within periodic auditory pulse patterns. J. Acoust. Soc. Am., 42, 1335-1340.

Pollack, I. (1968a). Asynchrony II: Perception of temporal gaps within periodic and jittered pulse patterns. J. Acoust. Soc. Am., 43, 74-76.

Pollack, I. (1968b). Detection and relative discrimination of auditory "jitter". J. Acoust. Soc. Am., 43, 38-315.

Pollack, I. (1968c). Discrimination of mean temporal interval within jittered auditory pulse trains. J. Acoust. Soc. Am., 43, 1107-1112.

Preusser, D. (1972). The effect of structure and rate on the recognition and description of auditory temporal patterns. Perception & Psychophysics, 11, 233-240.

Royer, F.L. and Garner, W. R. (1966). Response uncertainty and perceptual difficulty of auditory temporal patterns. Perception & Psychophysics, 1, 41-47. for rhythm study

Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. Perception & Psychophysics, 45, 291-296.

Sorkin, R. D. (1966). Temporal interference effects in auditory amplitude discrimination. Perception & Psychophysics, 1, 55-58.

Sorkin, R. D. (1982). Temporal factors in the discrimination of tonal sequences. J. Acoust. Soc. Am., 82, 1218-1226.

Sorkin, R. D. (1990). Perception of temporal patterns defined by tonal sequences. J. Acoust. Soc. Am., 87, 1695-1701.

Sorkin, R. D., Boggs, G. J., & Brady, S. L. (1982). Discrimination of Temporal Jitter in Patterned sequences of Tones. Journal of Experimental Psychology: Human Perception and Performance, 8, 46-57.

Sorkin, R. D. and Snow, M. P. Discrimination of time-expanded tonal sequences. (1987). J. Acoust. Soc. Am., 82, S40.

Sorkin, R. D. and Montgomery, D. A. Effect of time compression and expansion on the discrimination of tonal patterns. Journal of the Acoustical Society of America, (submitted).

Sturges, P. T. and Martin, J. G. (1974). Rhythmic structure in auditory temporal pattern perception and immediate memory. Journal of Experimental Psychology, 102, 377-383.

Watson, C. S. (1987). "Uncertainty, informational masking, and the capacity of immediate auditory memory," in Auditory Processing of Complex Sounds, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 267-277.

Watson, C. S., Kelly, J., and Wroton, H. W. (1976). Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty, J. Acoust. Soc. Am., 61, 1176-1186.

Watson, C. S., Wroton, H. W., Kelly and Benbasset, C. A. (1975). Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals, J. Acoust. Soc. Am., 57, 1175-1185.

Woods, D. D., Sorkin, R. D. and Boggs, G. J. (1979). Stimulus context and duration discrimination. Perception and Psychophysics, 26, 127-132.

V. DERIVATION OF THE PATTERN CORRELATION MODEL

Pattern Correlation Model

The time intervals between tones were generated by a process that enabled experimental control of the mean and standard deviation of the intertone intervals. The intervals were generated by combining the three independent, normal random variables:

$$X_a,\quad X_b,\quad X_c,\quad \text{where}\quad \mu_a = \mu_b = 0 \text{ and } \sigma_a^2 = \sigma_b^2 = \sigma_u^2$$

The random variables were combined to form the two sequences of interonset times $\{X_1\}$ and $\{X_2\}$:

$$X_1 = X_a + X_c \text{ and } X_2 = X_b + X_c \qquad \text{(A1a,b)}$$

where

$$E[X_1] = E[X_2] = \mu_c \text{ and } Var[X_1] = Var[X_2] = \sigma_u^2 + \sigma_c^2$$

To compute the correlation between the sequences $\{X_1\}$ and $\{X_2\}$:

$$\rho_{X1,X2} = [Cov(X_1,X_2)]/\sigma_{X1}\sigma_{X2} \qquad \text{(A2)}$$

$$Cov(X_1,X_2) = E[(X_1-\mu_1)(X_2-\mu_2)]$$

$$= E[(X_a+X_c)(X_b+X_c)]-\mu_c E[X_c]-\mu_c E[X_c]+\mu_c\mu_c = \sigma_c^2 \qquad \text{(A3)}$$

$$\rho_{X1,X2} = \sigma_c^2/(\sigma_u^2 + \sigma_c^2) \qquad \text{(A4)}$$

On SAME trials of the experiment, $\rho_{X1,X2}$ is set to 1.0 and

on DIFFERENT trials, $\rho_{X1,X2}$ is set to ρ_{ex}.

Suppose that the listener's response is based on the Pearson product-moment correlation, r_{12}, computed on the interonset intervals. The Fisher r to Z transformation yields a decision variable that is approximately normally distributed (Brunk, 1960):

$$Z = (1/2)\ln[(1+r_{12})/(1-r_{12})] \qquad \text{(A5)}$$

The mean and standard deviation of Z are:

$$\mu_Z \approx (1/2)\ln[(1+\rho)/(1-\rho)]+(\rho)/(2n-1) \text{ and } \sigma_Z \approx (n-3)^{-\frac{1}{2}} \qquad \text{(A6a,b)}$$

Then, d' is given by the difference between the means of the Z statistic on DIFFERENT and SAME trials, divided by the standard deviation of Z:

$$d' = [(n-3)^{\frac{1}{2}}]\left[\frac{1}{2}\ln\left(\frac{1+\rho_{SAME}}{1-\rho_{SAME}}\right) + \frac{\rho_{SAME}}{2n-1} - \frac{1}{2}\ln\left(\frac{1+\rho_{DIFF}}{1-\rho_{DIFF}}\right) - \frac{\rho_{DIFF}}{2n-1}\right] \qquad \text{(A7)}$$

We postulate an internal uncorrelated jitter, σ^2_{in}, associated with the Human listener's encoding and storage of the interonset times. The independent, normal random variables X_{in1} and X_{in2} are added, respectively, to each sequence;

$$X_1 = X_a + X_c + X_{in1} \quad \text{and} \quad X_2 = X_b + X_c + X_{in2} \qquad \text{(A8a,b)}$$

where X_a, X_c, X_{in1}, X_b, X_{in2} are all pair-wise independent and

$$\mu_{in1} = \mu_{in2} = 0 \quad \text{and} \quad \sigma^2_{in1} = \sigma^2_{in2} = \sigma^2_{in}$$

then, $E[X_1] = E[X_2] = \mu_c$

and $Var[X_1] = Var[X_2] = \sigma^2_u + \sigma^2_c + \sigma^2_{in}$

$$Cov(X_1,X_2) = E[X^2_c] - \mu^2_c = \sigma^2_c \qquad \text{(A9)}$$

$$\rho_{X1,X2} = \sigma^2_c/[(\sigma^2_u + \sigma^2_c + \sigma^2_{in})^{\frac{1}{2}}(\sigma^2_u + \sigma^2_c + \sigma^2_{in})^{\frac{1}{2}}]$$

$$= \sigma^2_c/(\sigma^2_u + \sigma^2_c + \sigma^2_{in}) \qquad \text{(A10)}$$

let $\sigma^2_{ex} = \sigma^2_c + \sigma^2_u$ (A11)

then

$$\rho_{X1,X2} = \sigma^2_c/(\sigma^2_{ex} + \sigma^2_{in}) \qquad \text{(A12)}$$

$$= (\sigma^2_c/\sigma^2_{ex})/[1+(\sigma_{in}/\sigma_{ex})^2] \qquad \text{(A13)}$$

the effective correlations will be:

$$\rho_{SAME} = 1/[1+(\sigma_{in}/\sigma_{ex})^2] \quad \text{and} \quad \rho_{DIFF} = \rho_{ex}/[1+(\sigma_{in}/\sigma_{ex})^2] \qquad \text{(A14a,b)}$$

Thus, the internal noise tends to reduce the correlation between the sequences, so that the correlation is less than 1.0 on SAME trials and less than ρ_{ex} on DIFFERENT trials.

Effect of Duration Transformations

We wish to determine the effect of multiplying each sequence, respectively, by the multiplicative factors k_1 and k_2.

That is, we let

$$X_1 = k_1X_a + k_1X_c + X_{in1} \quad \text{and} \quad X_2 = k_2X_b + k_2X_c + X_{in2} \qquad \text{(A15a,b)}$$

$$E[X_1]=k_1\mu_c, \quad E[X_2]=k_2\mu_c \qquad \text{(A16a,b)}$$

$$Var[X_1] = k_1^2\sigma^2_u + k_1^2\sigma^2_c + \sigma^2_{in}, \quad Var[X_2] = k_2^2\sigma^2_u + k_2^2\sigma^2_c + \sigma^2_{in} \qquad \text{(A17a,b)}$$

$$Cov(X_1,X_2) = E[(X_1-k_1\mu_c)(X_2-k_2\mu_c)]$$

$$= k_1k_2E[X^2_c]-k_1k_2\mu^2_c = k_1k_2\sigma^2_c \qquad \text{(A18)}$$

$$\rho_{X1,X2} = k_1 k_2 \sigma_c^2 / [(k_1^2 \sigma_u^2 + k_1^2 \sigma_c^2 + \sigma_{in}^2)^{1/2} (k_2^2 \sigma_u^2 + k_2^2 \sigma_c^2 + \sigma_{in}^2)^{1/2}] \quad \text{(A19)}$$

$$= k_1 k_2 \rho_{ex} [k_1^2 + (\sigma_{in}/\sigma_{ex})^2]^{-1/2} [k_2^2 + (\sigma_{in}/\sigma_{ex})^2]^{-1/2} \quad \text{(A20)}$$

then

$$\rho_{SAME} = k_1 k_2 [k_1^2 + (\sigma_{in}/\sigma_{ex})^2]^{-1/2} [k_2^2 + (\sigma_{in}/\sigma_{ex})^2]^{-1/2} \quad \text{(A21)}$$

$$\rho_{DIFF} = k_1 k_2 \rho_{ex} [k_1^2 + (\sigma_{in}/\sigma_{ex})^2]^{-1/2} [k_2^2 + (\sigma_{in}/\sigma_{ex})^2]^{-1/2} \quad \text{(A22)}$$

Following the same arguments, it can be shown that the addition of a constant time interval to either (or both) of the sequences, e.g. $X_1 = X_u + X_c + X_{in1} + t_1$, has no effect on ρ_{SAME} or ρ_{DIFF}. Note that if there is no internal noise, the multiplicative transformations have no effect on performance, since then $\rho_{SAME} = 1$ and $\rho_{DIFF} = \rho_{ex}$.

VI. PROJECT PERSONNEL

Elvers, G. C. Assistant in Psychology, Department of Psychology, University of Florida. Dr. Elvers worked on the project during 1988-89 and received his Ph.D. (from Purdue University) in May, 1989.

Pezzo, M. (Graduate Student, Department of Psychology, University of Florida). Mr. Pezzo worked on the project during 1988-89.

Montgomery, D. A. Graduate Student, Department of Psychology, University of Florida. Ms. Montgomery has been on the project since May, 1989.

Sorkin, R. D. Principal Investigator, Professor of Psychology, University of Florida.

Li CangPu, Visiting Scholar, Department of Psychology, University of Florida. Mr. Li has assisted in the laboratory since August, 1989.

VII. ADVANCED DEGREES

Elvers, G. C. Detection of visual signals consisting of multiple information sources: A signal detection analysis. Ph.D. Dissertation, Purdue University, May 1989.

Montgomery, D. A. Information integration under processing limitations: A weight analysis. M.A. Thesis, University of Florida (in preparation).

VIII. INTERACTIONS

Member, National Research Council Committee on Hearing, Bioacoustics, and Biomechanics (CHABA Working Group on Classification of Complex, Non-speech Sounds.

Associate Editor, International Journal on Human-Computer Interaction.

Member, Acoustical Society of America Long Range Planning Committee, 1989-.

Member, National Research Council Workshop on Digitally Processed Acoustic Signals. Mystic, Connecticut, March 1990.

Member, Research Advisory Council, Center for Applied Human Factors in Aviation, Orlando, Florida, 1990-.

Attendee, 116th Meeting of the Acoustical Society of America, 14 November, 1988. Honolulu, Hi.

Attendee, 118th Meeting of the Acoustical Society of America, 27 November, 1989. St. Louis, Mo.

Visitor, USAF School of Aviation Medicine (vision and hearing labs.), 23 February, 1990. San Antonio, Tx.

Attendee, Auditory Research Organization, 5 February, 1990. St. Petersburg, Fl.

Participant, Indiana University Conference on Human Error, 22 March, 1990. Bloomington, In.

IX. PUBLICATIONS

Elvers, G. C. and Sorkin, R. D. Detection and recognition of multiple visual signals in noise. Proceedings of the Human Factors Society 33rd Annual Meeting, 1989, 2, 1383-1387.

Makhoul, J., Crystal, T.H., Green, D. M., Hogan, D., McAulay, R.J., Pisoni, D.B., Sorkin, R.D., and Stockham, T.G.,Jr. Removal of Noise From Noise-Degraded Speech Signals, Committee on Hearing, Bioacoustics, and Biomechanics, National Research Council, National Academy Press, Washington, DC, 1989.

Sorkin, R. D. Review of M. Loeb, "Noise and Human Efficiency", American Journal of Psychology, 1988, 101, 290-293.

Sorkin, R. D. Why are people turning off our alarms? Journal of the Acoustical Society of America, 1988, 84, 1107-1108. Reprinted in Human Factors Society Bulletin, 1989, 32, 3-4.

Sorkin, R. D. Perception of temporal patterns defined by tonal sequences. Journal of the Acoustical Society of America, 1990, 87, 1695-1701. (attached to this report)

Sorkin, R. D. and Elvers, G. C. Analysis of Automated Decision Systems. Final Report of Naval Weapons Center Contract N60530-88-C-0213, June 1989.

Sorkin, R. D., Kantowitz, B. H., and Kantowitz, S. C. Likelihood Alarm Displays, Human Factors, 1988, 30, 445-459.

Sorkin, R. D., Mabry, T. R., Weldon, M., and Elvers, G. Integration of information from multiple element displays. Organizational Behavior and Human Decision Processes, 1991, (in press).

Sorkin, R. D. and Montgomery, D. A. Effect of time compression and expansion on the discrimination of tonal patterns. Journal of the Acoustical Society of America, (submitted).

Sorkin, R. D. and Montgomery, D. A. Effect of stimulus delay on the discrimination of tonal patterns. Journal of the Acoustical Society of America, (in preparation).

Sorkin, R. D. Pezzo, M. V., and Elvers, G. C. Discrimination of partially correlated temporal sequences. Journal of the Acoustical Society of America, 1989, 86, S1, S123.

Sorkin, R. D. and Robinson, D. E. Computer-aided signal detection and classification. International Journal of Human-Computer Interaction, (in preparation).

Sorkin, R. D., Wightman, F. L., Kistler, D. J., and Elvers, G. C. An exploratory study of the use of movement-correlated cues in an auditory head-up display. Human Factors, 1989, 31, 161-166.

Yost, W. A., Braida, L. D., Hartmann, W. M., Kidd, G. D. Jr., Kruskal, J. B., Pastore, R. E., Sachs, M. B., Sorkin, R. D., Warren, R. M. Classification of Complex Nonspeech Sounds, Committee on Hearing, Bioacoustics, and Biomechanics, National Research Council, National Academy Press, Washington, DC, 1989.