# 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

0001-4966/90/041695-07$00.80

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 $\mu_{\text{gap}}$, the standard deviation of the gaps, $\sigma_{\text{gap}}$, and the correlation $\rho_{\text{ex}}$ 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):

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

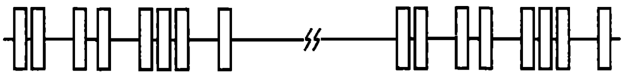
and

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

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.

(a) SAME



(b) DIFFERENT

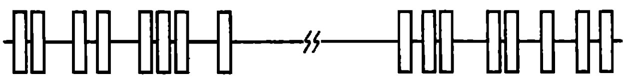


FIG. 1. The envelopes of typical tone sequences are shown for (a) same and (b) different trials.

## 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

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

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 $\rho$ is the population correlation coefficient, the mean and standard deviation of $Z$ are then given by (Brunk, 1960)

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

and

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

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 $\rho = \rho_{\text{ex}}$. 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 $\rho_{\text{ex}}$, $\sigma_{\text{gap}}$, 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 $\sigma^2_{\text{in}}$, and that this jitter is uncorrelated across the gap sequences. Adding this uncorrelated jitter $\sigma^2_{\text{in}}$ to Eqs. (1) and (2), yields

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

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

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

The magnitude of the internal temporal jitter $\sigma_{\mathrm{in}}$ 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 $Z$:

$$\begin{aligned} d' = \bigg[ & \frac{1}{2} \ln\left(\frac{1+\rho_{\mathrm{SAME}}}{1-\rho_{\mathrm{SAME}}}\right) + \frac{\rho_{\mathrm{SAME}}}{2n-1} \\ & - \frac{1}{2} \ln\left(\frac{1+\rho_{\mathrm{DIFF}}}{1-\rho_{\mathrm{DIFF}}}\right) - \frac{\rho_{\mathrm{DIFF}}}{2n-1} \bigg] \\ & \times (n-3)^{-1/2}. \end{aligned} \tag{8}$$

## 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 $\rho_{\mathrm{ex}}$ (as specified on *DIFFERENT* trials, since $\rho = 1$ on *SAME* trials) and the standard deviation of the intertone gaps $\sigma_{\mathrm{gap}}$, and to estimate the magnitude of the internal noise $\sigma_{\mathrm{in}}$.

### 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 | $\rho_{\mathrm{DIFF}}$ | Gap number | Gap mean | Gap standard deviation | Sequence duration |
|---|---|---|---|---|---|
| 1a | 0, 0.2, 0.35, 0.4<br>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 $\mu_{\mathrm{gap}}$ 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-

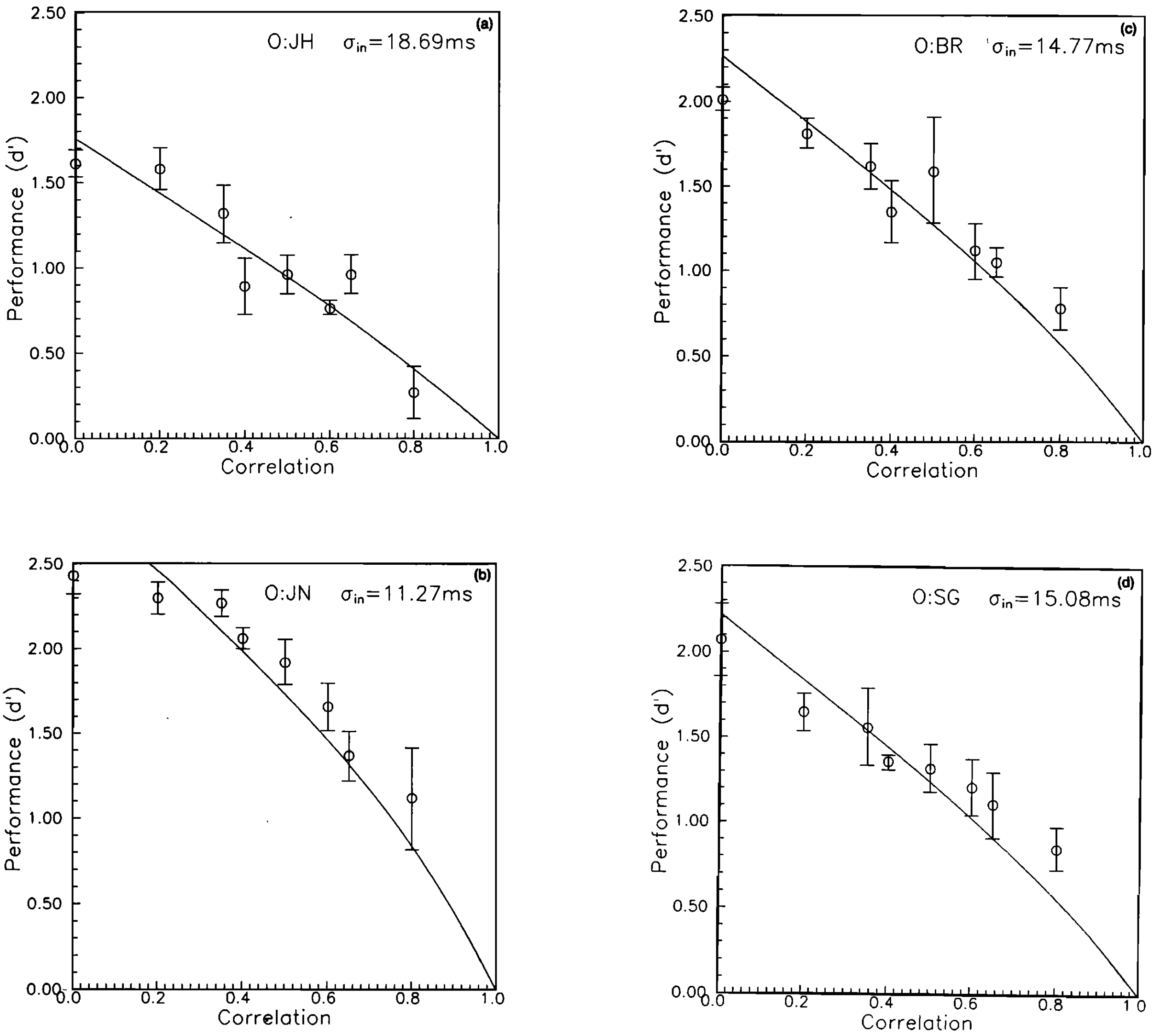


FIG. 2. Performance ($d'$) is plotted as a function of the sequence correlation, for each of four observers. The solid lines show the performance of the correlation model with the internal noise standard deviation shown (see text).

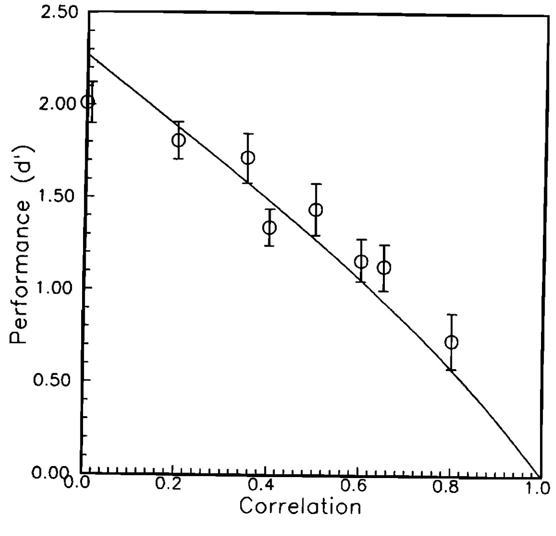


FIG. 3. The average performance of four observers ($d'$) is plotted as a function of the sequence correlation. The solid line is the prediction of the correlation model with an internal noise of 14.75 ms.

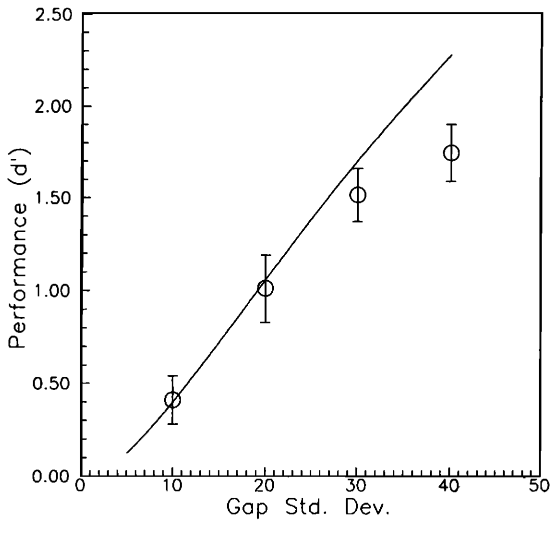


FIG. 4. The average performance of four observers ($d'$) is plotted as a function of the standard deviation of the gaps. The solid line is the prediction of the correlation model with an internal noise of 14.75 ms.

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 $\mu_{\text{gap}}$ 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 $\mu_{\text{gap}}$. However, it is reasonable to expect that a Weber's law relationship would hold, such that the magnitude of the internal jitter $\sigma_{\text{in}}$ would increase with the duration of the intervals to be judged. Such a relationship, where $\sigma_{\text{in}}$ increases in proportion to $\mu_{\text{gap}}$, 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:

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

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 (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

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

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).

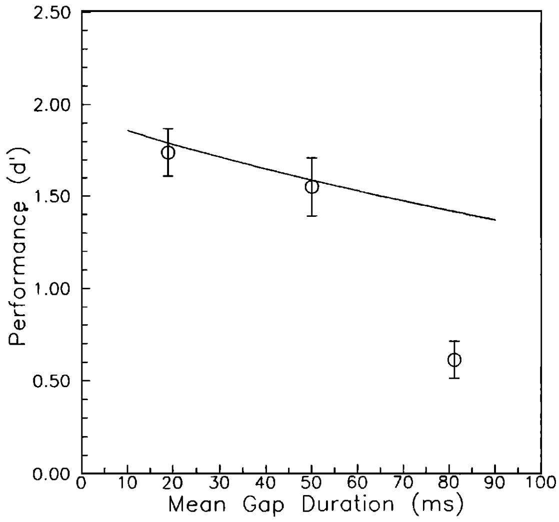


FIG. 5. The average performance of four observers ($d'$) is plotted as a function of the mean gap duration. The solid line is the prediction of the correlation model revised to incorporate the effect of mean gap (see text).

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

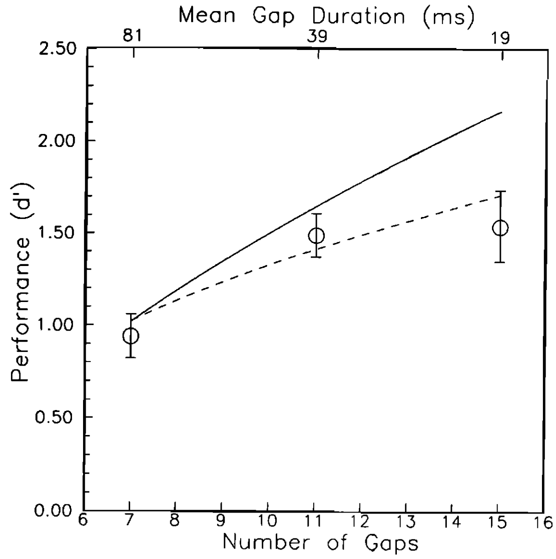


FIG. 6. The average performance of four observers ($d'$) is plotted as a function of the number of gaps (average sequence duration is fixed). The solid line is the prediction of the correlation model revised to incorporate the effect of mean gap ($\sigma_{\text{in}} = 12.25 + 0.05\mu_{\text{gap}}$).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

$$x_1 = u|x_a| + c|x_c|, \tag{A1}$$

$$x_2 = u|x_b| + c|x_c|, \tag{A2}$$

where $u$ and $c$ are constants defined by

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

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 $\mu_{\rm gap}$ and standard deviation equal to $\sigma_{\rm gap}$. 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 $\rho_{\rm ex}$, 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.