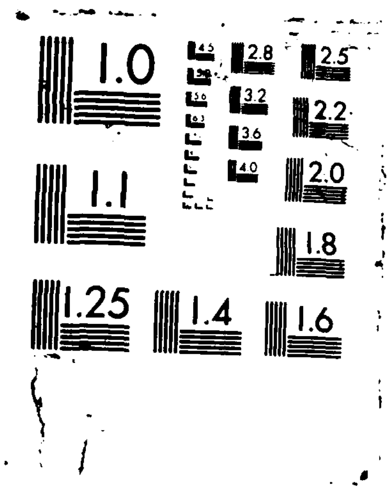
AD-A190 218 PERCEPTION OF COMPLEX AUDITORY PATTERNS(U) INDIANA UNIV AT BLOOMINGTON HEARING AND COMMUNICATION LAB C S WATSON 02 NOV 87 AFOSR-TR-87-1781 AFOSR-84-0337 1/1

UNCLASSIFIED F/G 6/4 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

SECURITY CLASSIFICATION OF THIS PAGE

AD-A190 218

REPORT DOCUMENTATION PAGE

DTIC FILE COPY

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 87-1787
6a NAME OF PERFORMING ORGANIZATION: Indiana University; 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research/NL
6c ADDRESS (City, State, and ZIP Code): Bloomington Indiana 47405	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-84-0337
8c ADDRESS (City, State, and ZIP Code): Building 410, Bolling AFB, DC 20332-6448	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 61102F; PROJECT NO 2313; TASK NO A6; WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

PERCEPTION OF COMPLEX AUDITORY PATTERNS

12 PERSONAL AUTHOR(S)
Dr Charles S. Watson

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM 1 Sep 84 TO 31 Aug 87	2 Nov 87	41

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19 ABSTRACT

This project continued and extended a series of experiments on the discrimination and identification of complex auditory patterns. The general purpose of this work is to determine the limits of human listeners' abilities to extract information from complex sounds including, but not limited to, those with temporal and spectral properties approximating speech. Experiments used criterion-controlled psychophysical methods in which listeners were trained until approaching asymptotic performance in various discrimination and identification tasks. Advances were made in the following areas: a) the spectral and temporal range of selective auditory attention; b) the time course of auditory perceptual learning; c) informational limits on pattern discrimination; d) listeners' abilities to learn to attend to multi-tone targets within longer patterns; e) individual differneces in auditory pattern discrimination abilities among listeners with normal auditory sensitivity, and f) the perception of spectrally complex sounds, including speech and non-speech sounds.

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

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

AFOSR-TR- 87-1787

Hearing and Communication Laboratory

Department of Speech and Hearing Sciences
Indiana University
Bloomington, Indiana 47405

Final Technical Report

Air Force Office of Scientific Research

AFOSR-84-0337
September 1, 1984 to August 31, 1987

Hearing and Communication Laboratory
Indiana University

October, 1987

Table of Contents

Abstract

This project continued and extended a series of experiments on the discrimination and identification of complex auditory patterns. The general purpose of this work is to determine the limits of human listeners' abilities to extract information from complex sounds including, but not limited to, those with temporal and spectral properties approximating speech. Experiments used criterion-controlled psychophysical methods in which listerers were trained until approaching asymptotic performance in various discrimination and identification tasks. Advances were made in the following areas: a) the spectral and temporal range of selective auditory attention; b) the time course of auditory perceptual learning; c) informational limits on pattern discrimination; d) listeners' abilities to learn to attend to multi-tone targets within longer patterns; e) individual differences in auditory pattern discrimination abilities among listeners with normal auditory sensitivity, and f) the perception of spectrally complex sounds, including speech and non-speech sounds. The primary significance of the overall research program is that it provides a link between auditory theories based on human listeners' abilities to detect and discriminate among isolated tones and theories concerning the perception of auditory patterns, spectrally complex sounds, and speech. Potential applications of this work include the determination of optimally detectable and discriminable signals for use in man-machine interactions, the limits of human abilities to learn nonspeech codes, and methods of identifying individuals with unusually excellent (or minimal) abilities to learn tasks requiring discrimination or identification of auditory signals.

Scientific Goals

This research continued several lines of experimentation on listener's abilities to discriminate and identify complex auditory patterns, and extended that work to include three additional topics, (a) individual differences in the ability to extract information from complex patterns, (b) the theoretical nature of the limits of processing of complex patterns, and (c) replication of the nonspeech pattern discrimination studies with real and synthetic speech stimuli. Also, a new series of studies has been conducted to determine optimal decision strategies for combined man-machine detection systems.

1. Continuation of Current Studies of Tonal Patterns

Detection and discrimination experiments were conducted to answer several questions raised by results obtained during the previous three years, and the empirical results now available have been incorporated in a first-order theory of auditory pattern perception.

a. What is the relation between the ability to detect components of complex auditory patterns and the ability to discriminate changes in the frequency, intensity, and duration of those same components?

b. What general rules describe those properties, stimulus characteristics, or relations among components, that are most salient in unexpected complex auditory patterns? In other words, what systematic auditory perceptual "sets", or pre-emphasis, characterize the processing of an unexpected sound?

c. What is the upper limit of information that a listener can extract from a single complex auditory pattern (durations between 50 and 2K msec)?

d. Can the "pre-emphasis" spectral-temporal filters implied by the results of experiments with tonal patterns predict performance with other classes of complex sounds, especially those with more complex instantaneous spectra? This new project was conducted in collaboration with D.E. Robinson, of the Department of Psychology. The discriminability of pairs of reproducible, Gaussian noise bursts have been investigated in order to determine the degree to which the general rules describing the discriminability of tonal patterns generalize to these complex waveforms. Spectral analyses of the noise samples were performed with various temporal windows to determine the temporal weighting function which best describes listeners' abilities to discriminate among these stimuli.

2. Individual Differences in Pattern Processing

A reduced eight-subtest psychoacoustic battery was developed, based on an analysis of the performance of a sample of 28 normal listeners on an earlier 22-subtest battery (Johnson, Watson, & Jensen, 1987). This new test battery has now been standardized on approximately 300 listeners (Watson, Jensen, Foyle, Leek, & Goldgar, 1982). Data from the reduced test battery will be evaluated to determine the primary dimensions of auditory discrimination abilities. Because of the unexpectedly large differences, among audiometrically normal listeners, in the ability to discriminate complex sounds, we investigated the possible significance of these differences in predicting performance on a variety of tasks. These included second-language learning and learning of non-speech auditory codes.

3. Replication of Pattern Discrimination Studies with Real and Synthetic Speech

Effects of stimulus uncertainty, prior listening experience, and specialized auditory training has been investigated with real and synthetic speech stimuli, to determine whether those variables have similar effects on speech perception to the profound ones demonstrated in work with complex, non-speech patterns (Watson, Wroton, Kelly, & Benbassat, 1975; Watson, Kelly, & Wroton, 1976; Spiegel & Watson, 1981; Watson & Kelly, 1981.)

4. Computer Assisted Detection

Using basic concepts of statistical decision theory a Contigent Criterion Model has been developed to predict optimal performance in tasks in which human and machine-based decisions about threatening conditions must be combined. Tests of the model have been conducted with human observers (Sorkin, Robinson, and Berg, 1987).

Summary of Major Results

During the 1984-1987 grant period a large number of experiments were completed by the two groups to which support was provided, the Hearing and Communication Laboratory, and the Auditory Research Laboratory, directed by C. S. Watson and D. E. Robinson, respectively. The general purpose of this work is to determine the limits of human listeners' abilities to extract information from complex sounds including, but not limited to, those with temporal and spectral properties approximating speech. Major results of these experiments include the following:

(1) A critical parameter in the discrimination of complex patterns is the proportional duration of the components subject to change, relative to total pattern durations, rather than the absolute duration of either patterns or components. A related result was erroneously interpreted in earlier experiments to mean that the number of independent pattern components is the primary variable limiting pattern discrimination.

(2) The salience of a change in physical dimension of an auditory pattern is inversely proportional to the amount of information currently encoded on that dimension.

(3) Complex acoustic patterns designated as "targets" can be identified by integrating information probabilistically distributed among such pattern features as spectral shape, temporal envelope, and degree of departure from simple harmonicity (only harmonically-related components present in the pattern).

(4)Individual audiometrically normal listeners show strong tendencies to attend more strongly to certain features, in auditory target identification tasks, and that tendency is robust, at least persisting after several days of intense training on previously ignored dimensions.

(5) Although listeners are unable to detect changes of less than 50-60 msec in the duration of components of unfamiliar patterns, minimal- uncertainty training procedures result in accurate detection of 5-8 msec changes in the same patterns.

(6) Changes in an essentially random acoustic waveform (sample of gaussian noise) are highly detectable when occurring at the end of a sample, while the same changes may be inaudible when they occur at the beginning of the sample. This heightened resolution at the end of a waveform (a) is independent of sample duration (for durations of 25-150 msec), and (b) demonstrates the generality of a similar result reported earlier for tonal patterns.

(7) Internal noise is assumed by signal detection theories to account for less-than-perfect detection and discrimination performance. A model has been developed in which the internal noise has been partitioned into peripheral and central components. That model was tested in experiments in which the external stimulus distribution were rigidly controlled.

(8) A theoretical model of man-machine decision making has been developed, to be applied to situations in which a human is monitoring a channel on which a "threat" may occur, while an automated detector monitors an independent channel that bears information about the same threatening conditions. Work to date demonstrates that

joint human- alarm system performance can significantly exceed that of either human or alarm system operating individually.

These experiments have generated several forms of useful information for operational applications. They provide a large amount of data on the level and range of performance that can be expected of highly trained human listeners whose assignments require them to detect, discriminate, or identify complex sounds. They provide "benchmarks" for the selection of unusually salient or identifiable acoustic signals. Last, they provide theoretical models of auditory discrimination and decision making which are not limited to the classes of acoustic events used in the laboratory experiments conducted to develop and test these models.

Research Accomplishments

1. Auditory processing capacity for tonal sequences

A series of experiments on the informational capacity of the auditory system for temporal patterns has now been completed. These experiments were originally conceived as a means of determining the optimal combination (for information transmission) of total-pattern and pattern-component durations. The total information in a pattern is considered to be proportional to the number of independently varying components. The patterns in each of these studies consisted of a series of 75-dB tone pulses, whose frequencies were randomly selected from the range 300-3000Hz. Successive tones in the sequences were never closer than 1/3 octave, and were gated on and off with a 2.5 msec rise-decay.

An earlier series of studies conducted in our laboratory used an adaptive-tracking procedure in which the number (n) of components in fixed-duration tonal patterns was increased or decreased from trial-to-trial, in a S/2AFC discrimination task. (S/2AFC: A paradigm in which a standard pattern is followed by two test patterns, one of which is different from the standard.) Patterns with six total durations, from 62.5 to 2000 msec, were presented in random order, while an independent adaptive-tracking history for each duration converged on the n's (number of components) required for 71% discrimination. As the numbers of components in the patterns were varied, the duration of the individual components was always the total duration/n. This procedure was repeated in seven separate experiments.

Results of these experiments suggest that when tonal patterns can be discriminated by the presence of a silent gap, or by a change in gap position, performance is determined by a critical component duration (25-50 msec, depending on the specific task). In those cases, the threshold values of n, for each total pattern duration, are approximately the total durations divided by a constant. In experiments in which discrimination requires some degree of resolution of the actual pitch contour, performance seems to reflect a fixed informational capacity for pattern discrimination, in the range of 6-9 components per pattern. No clear optimal combination of total and component duration can be seen in these data, since the same 6-9 component limit is found for a 32-fold range of total durations (62.5-2000 msec).

1.1. Isochronous vs. anisochronous patterns (Watson, Foyle)

The results of the above experiments were obtained with isochronous patterns (duration of each component = total duration/n). The relative constancy of the total information in discriminable patterns ranging from 62.5 msec to 2000 msec might be a property only of patterns which have the very salient rhythmic quality associated with isochronous temporal structure. A major difference in discriminability for isochronous and anisochronous patterns might be predicted by the results of a recent experiment reported by Sorkin (*J. Acoust. Soc. Am.* *75,* S21, 1984). To investigate that possibility, a new experiment was conducted, in which the random sequences of patterns included three levels of temporal "jitter" of the non-target component durations. In the resulting anisochronous patterns the target-component durations (the component whose frequency was incremented) were still total duration/n, while the non-target components were each randomly increased or decreased in duration by a fixed percentage of their isochronous value. The "jitter" percentages were 0%, 30%, or 50%, in separate conditions. It was found that threshold values of n were unaffected by the two levels of anisochrony, although the perceptual quality of the patterns was markedly changed by these manipulations. Sorkin's study differed from this experiment, in that he studied the effects of within-trial variation in the temporal structure of patterns, thus these results do not directly contradict his.

1.2. Capacity estimated in a true frequency-discrimination paradigm (Watson, Kidd)

In the above experiments, the dependent variable was n -- an unusual psychophysical procedure that is, to our knowledge, unique to these experiments. While we know of no theoretical reason that this paradigm would yield aberrant results compared to more traditional methods, it nevertheless seemed reasonable to attempt to estimate the pattern-discrimination "capacity" using a more traditional psychophysical approach. Another experiment was therefore conducted, in which the dependent variable in the adaptive-tracking variable was $\Delta f/f$, the proportional change in the frequency of a mid-temporal position, mid-frequency component. Threshold values (71% correct) of $\Delta f/f$ were determined for various numbers of components, for total pattern durations of 125, 500, and 1500 msec. Although there is some reduction in threshold as total pattern duration (and therefore component duration, in these isochronous patterns) is increased, that effect is extremely small compared to the changes associated with variation in n.

Taken together, the results of these experiments suggest a limit on pattern processing in terms of the total amount of information contained in tonal patterns, rather than in terms of critical values of some physical parameters. Such a processing limit is reminiscent of Miller's (1956) "magical number 7±2," and of the results of some of Pollack's (1953) experiments on the information in multi-dimensional auditory displays. It extends that earlier work to complex temporal auditory stimuli. These limits appear to be general at least for stimuli in the range of pattern durations thus far investigated (62.5-2000 msec), but only for cases in which discrimination must be based on the contents of immediate memory. When the listener has some long-term basis for focusing attention on a restricted portion of a complex pattern, then these informational limits do not yield accurate predictions of performance. When the information-processing demands are reduced, as by permitting successful use of top-down direction of attention (e.g., Spiegel and Watson, 1981) considerably greater amounts of stimulus information may be included in discriminable patterns. The predictability of the waveforms of speech (or of most music) thus affects the applicability of the limited-capacity hypothesis to such familiar and highly constrained stimuli.

1.3. Detection of level changes in multi-tone patterns (Watson, Kidd, Washburne)

The experiments on listeners' processing capacity have now been extended to the detection of changes in the level of individual tones in multi-tone patterns. Listeners' abilities to detect increments and decrements of the intensity of tones were examined with a range of sequence lengths (1 to 9) and of total pattern durations (125 to 1500 msec). We found that, in contrast to our data for the detection of changes in frequency, performance is primarily affected by individual component duration with very little influence of number of tones or of total pattern duration. This is very much like the results of our earlier experiments on the detection of gaps in multi-tone patterns. These cases have in common that it is not necessary to attend to the series of pitch changes in order to detect the change in the pattern. As a working hypothesis, it appears that "saturation" of the processing capacity for pitch changes has little or no degrading effect on listeners' abilities to detect changes in other stimulus dimensions. This result is consistent with Pollack's findings of an increase in information transmitted through the use of multi-dimensional encoding.

1.4. Proportional target-tone duration as a factor in the discriminability of tonal patterns (Watson, Kidd)

A series of experiments on listeners' abilities to extract information from patterns with varying total durations and numbers of tonal components has previously been reported [J. Acoust. Soc. Am. Suppl. 1 **73**, S44 (1983); **77**, S1 (1985)]. In those experiments listeners were tested in high-stimulus-uncertainty, same-different pattern discrimination tasks, in which the tonal patterns to be discriminated differed by changes in the frequency of one or more components. Discrimination performance in those tasks was consistent with previous measures of the frequency resolving power of the auditory system when the patterns contained one to three equal-duration components, for total pattern durations from 62.5-2000 ms. As the number of components was raised, discrimination thresholds increased by large amounts, often by factors of 10-100 for patterns with more than seven-eight components. While this result might imply an informational limit on pattern processing, it is also consistent with the hypothesis that *target tones are equally well resolved if they occupy equal proportional durations* of the patterns in which they occur. Results of a new experiment, in which the proportional durations of target tones and the number of tones per pattern were independently varied, suggest that proportional duration of the target tones is in fact the primary determinant of pattern discriminability for tonal patterns ranging from 100-1500 ms in total duration. [Abstract of paper presented at the 114th meeting of the Acoustical Society of America; Miami, Florida; November, 1987.]

2. Detection of pattern repetition in continuous tone-patterns (Kidd, Watson, Washburne)

The existence of a general processing-capacity limitation, as suggested in the previous studies, does not mean that all pattern discrimination tasks would necessarily reflect that same limit. We have therefore investigated listeners' abilities to detect the repetition of multi-tone patterns as a function of tone duration and number of tones in the pattern. In this experiment, generally modeled after that reported by Guttman and Julesz, 1963), subjects are presented with repeating or non-repeating tonal patterns using a tracking paradigm that increases or decreases the number of tones in a pattern, depending on a subject's performance. Because of the possibility that successful performance of the task might be strongly influenced by detection of the repetition of perceptually unique events within a pattern, we chose patterns designed to have few such events. In one series of tests we investigated the effects of decreasing the bandwidth, intended to reduce the likelihood of the occurrence of unique events that result from frequency-based auditory stream segregation. Preliminary data collection has been completed, utilizing 50-msec and 200-msec tones with 1/3-octave and 1-octave pattern bandwidths (centered on 1000 Hz) with 9 subjects participating in all conditions. These data showed strong effects of tone duration and bandwidth, as well as a significant interaction (due to a slightly greater effect of bandwidth at the 50 ms tone duration). The number of components for which each subject could correctly detect repetitions 70% of the time was estimated for each condition. The mean number of components for the 9 subjects for each condition is shown in Table 1. In general it can be seen that listeners are able to detect the repetition of patterns consisting of more tones with the shorter tone duration and the wider bandwidth. Interestingly, the effect of tone duration is not simply an effect of total pattern duration: subjects are able to detect the repetition of patterns with longer total durations (but fewer tones) at the 200-msec tone duration.

Table 1. Mean number of tones for 70% correct detection of repetition (total duration of detectable repeating patterns, in seconds, shown in parentheses).

	Tone Duration	
Bandwidth	50ms	200ms
1/3 Octave	62.9 (3.16)	30.7 (6.14)
1 Octave	94.1 (4.71)	35.5 (7.10)

Despite our attempts to minimize the occurrence of unique events, subjects' reports indicated that judgments were often based on the reoccurrence of particular events rather than detection of whole-pattern repetition. To further reduce the occurrence of unique events, a new version of this experiment was developed in which the sequences of pitches of consecutive tones approximated a sinusoidal series. Tones deviated randomly from strict sinusoidal variation by ± 6% and a single repeating pattern spanned three cycles. This procedure reduces the possibility of unique events by constraining adjacent tone relations while eliminating the problems of pattern-restart discontinuities and gross changes in pattern macrostructure.

Initial data collection with this new procedure revealed that unique events were still being used as a basis for repetition judgments. We are currently testing a new procedure that tracks on the deviation around the sine wave with a variable number of tones per cycle. The goal of this series of experiments is to devise a class of tonal sequences for which listeners must attend to the microstructure of an entire sequence to detect the repetition of a series of tones within the a sequence.

3. Perception of salient auditory events or figures (Watson, Kidd, Washburne)

In several studies Bregman and his colleagues (reviewed in Bregman, 1978) have described the factors associated with the emergence of auditory "streams" (sets of elements within a sequence of sounds that are more salient than their context). Similar effects have been noted in listening to repetition of the multi-tone sequences used in our experiments. A new series of experiments has been designed to more objectively measure the subpatterns (or auditory "events" or "figures") that listeners report hearing when patterns are repeated.

In an auditory figure-identification procedure (AFI), listeners work at computer terminals, where they are given one-key control over the presence or absence of each of the components of a tonal pattern (generally 10 tones). They check each tonal component by turning it on and off, to determine whether that component is a part of an auditory "figure" that emerges after the pattern has been repeated several times. When a component is identified as part of a figure, it is marked (by depressing another key), and when all components have been checked the listener can confirm his choices by a single key which turns on and off all non-marked components (ie. the "ground"). Another keystroke causes the selected subpattern and the time required to identify it to be recorded.

Results of a first experiment using the AFI procedure show excellent agreement among the figures identified by five well-trained listeners within a set of 120 patterns. In general, listeners identified figures within one frequency range (either high or low) more reliably as the range of figural components is relatively more compact and more distant from the non-figural components. The absolute frequency range of the elements that form a figure was not significantly related to its salience.

In a second experiment, the accuracy with which the figural and non-figural (ground) components are resolved was measured, using the method of adjustment described by Watson (1976). The frequency of single components was adjusted in a comparison pattern, until the listener decided that it had the same pitch as the corresponding component in a standard pattern.

In general, the adjustments of figural components are either slightly more accurate than for those that form the ground, or, in some cases, are made with the same accuracy, but require more pattern repetitions before the listener is satisfied with the match.

The primary goal of these preliminary experiments was to devise a rapid and reliable means by which listeners can identify the elements of a pattern which they perceive as a discrete auditory "figure" or "event". The AFI method is a very convenient means of a achieving that goal. In future experiments, we plan to use that method to study other factors that may be systematically related to the emergence of auditory figures or "targets" from various backgrounds.

4. Perception of multidimensional complex sounds (Watson, Kidd, Washburne)

4.1. Information integration with multidimensional complex sounds

Listeners' abilities to perceive information independently encoded in different dimensions of complex sounds were examined in experiments that required simultaneous attention to three dimensions. Stimuli consisted of sequences of 1, 3, 5, or 7 brief pulses that were generated by adding five 100-msec sinusoidal components. Each pulse had one of two values on each of the following complex dimensions: 1) harmonicity (harmonic vs inharmonic relations among the components), 2) spectral shape (linearly decreasing amplitude vs a two-peaked amplitude profile), and 3) amplitude envelope (slow vs rapid rise and decay times). Stimuli were selected such that the two values on each dimension were highly discriminable.

Two types of stimuli were generated by designating one value on each of the dimensions as the "target" value (harmonic spacing of sinusoids, the double-peaked power spectrum, and rapid rise/decay), and the other as the "non-target" value. The selection of dimensions for each component of a sequence was probabalistically determined and was adjusted to yield maximum possible (ideal) performance of 90% correct for all sequences.

Two groups of listeners were tested for 10 days. One group had four days of training in a single-dimension control experiment in which they identified target and non-target values for each of the individual dimensions while the other two dimensions varied randomly. The other group had no prior training but was tested in the single-dimension control experiment after completion of the main experiment.

Performance in the training experiment for the first group revealed very substantia' differences among listeners' abilities to attend to the individual dimensions, even after 4 days of training (240 trials per day per dimension). All listeners were able to correctly detect differences at least 70% of the time for each dimension, and close to 100% for at least one dimension.

Both groups of listeners showed an impressive ability to integrate information over pulses within sequences of up to 7 pulses, with identification performance at about 5% to 10% below that of the ideal (90%). The similarity in performance of the two groups, even at early stages of the experiment, indicates that the type of training we have used did not have a significant effect on listeners' ability to perform the information-integration task.

Performance of the second group of listeners in the single-dimension control experiment after 10 days of listening to the stimuli in the multi-pulse task was quite similar to that of the group tested prior to the multi-pulse task. There appears to be little or no effect of exposure in the integration task on discrimination ability as tested in the single-dimension control experiment.

In order to better understand listeners' attentional strategies in the integration task and how they might differ from those in the control experiment, responses to various stimulus configurations (i.e., multi-component stimuli with different numbers of target values on each dimension) were examined. The result of this analysis can be summarized as follows:

1. Although performance levels are often similar to levels that could be achieved by attending to a single dimension (approximately 80%), listeners were clearly attending to more than one dimension. Correlations between the number of stimuli with target values on a given dimension and listeners' responses were computed for each dimension and combination of dimensions. Correlations between responses and values on each individual dimension were higher than would be predicted on the basis of the correlation among the components, and correlations with the sum of all three dimensions were generally higher than with any single dimension or pair of dimensions. In other words, the listeners' decisions were in fact multi-dimensionally based.

2. Substantial individual differences were observed in the extent to which listeners attended to each of the dimensions. However, the allocation of attention suggested by the results of the single-dimension control experiment did not always agree with the apparent attentional distribution observed in the integration experiment. In some cases, the dimensions that influence listeners' responses most were not those that yielded highest performance in the control experiments (when feedback was based on a single dimension). It thus appears that a listener's ability to attend to a given dimension while others vary randomly is not a good predictor of his ability to use that same information when making decisions based on the combination of multiple dimensions.

4.2. Individual differences in the allocation of attention to specific dimensions

The existence of large individual differences in the allocation of attentional resources to various dimensions of sound sequences has recently been confirmed in another version of this experiment. Twenty-seven additional listeners were tested in a four-session "screening" protocol in which they were trained and tested in the classification of the three-dimensional target and non-target sounds. Approximately the same number of subjects displayed a preference for each of the three dimensions: Ten preferred spectral shape (the "profile" in Green's (1983) terms), nine preferred harmonicity, and eight preferred amplitude envelope. There were subjects who were skilled at processing each of the dimensions but could not seem to simultaneously process the other two, while a few subjects could reliably detect all three dimensions. The unsatisfactory generalization from these data was that there were many substantially different patterns of allocation of attention to the three dimensions and very little evidence of clusters of listeners with similar patterns.

At this point, this research has yielded two primary results concerning listeners' ability to categorize complex multidimensional sounds. One is that listeners can integrate multidimensional information in sequences of sound pulses, with little or no loss of efficiency with increasing sequence length, for sequences of one-two-seven components. The other is that they are often not very good at allocating attention to all three features (or "dimensions"), even though their absolute efficiency in this task is fairly high. In fact, comparable levels of performance are achieved with a variety of patterns of attention to the features of the stimuli. One possible interpretation of this result is that performance is limited in terms of the amount of information the listener can extract from complex sounds. The existence of small negative correlations between weightings for two of the three pairs of dimensions gives some support to this interpretation.

In new experiments we will determine (a) how well listeners can learn to process features they appeared to initially ignore, and (b) how well they can be taught to integrate information across all three dimensions. It would seem very likely that this task can ultimately be learned, given that each of the features *can* be discriminated easily. We will utilize training techniques that encourage listeners to attend to each of the individual features in the context of different stimulus configurations, as well as techniques that include the assignment of discrete labels (target identities) to each member of a large set of fairly distinct stimuli. [Recent support for this project has been provided by AFOSR; its initial phases of experimentation were supported by ONR/NMRDC.]

5. Temporal discrimination for single components of nonspeech auditory patterns (Espinoza-Varas, Watson)

(Analysis of previously collected data, and preparation of article for *J. Acoust. Soc. Am.)* The ability to discriminate increments in the duration of tonal components in ten-tone patterns was examined in three experiments employing a same-different psychophysical task. Two temporal structures (isochronous (40 msec) and random jitter (20-140 msec range)) and three levels of uncertainty (high, medium, and minimum) were examined. Superior discrimination performance was observed with increased training, decreased stimulus uncertainty, and isochronous temporal structure. These results suggest that a major determinant of the ability to discriminate the duration of of components of sequential patterns is the listener's knowledge about "what to listen for and where." Weber-Law predictions of thresholds accurately describe both minimal- (10% of single-component duration) and high-uncertainty performance (10% of total pattern duration).

6. Discriminability of complex waveforms (D. E. Robinson and S. M. Fallon)

Watson and his colleagues have provided a large amount of data concerning the discriminability of individual components within sequences of tonal patterns (Watson, Wroton, Kelly, and Benbassat, 1975; Watson, Kelly, and Wroton, 1976; Spiegel and Watson, 1981; Leek and Watson, 1984; Watson and Foyle, 1983; 1985a, 1985b; Watson and Kidd, 1987). We are now investigating the relationship between such relatively deterministic tonal sequences and essentially random waveforms. There appear to be some striking similarities between the two, apparently quite different, types of waveforms.

The research described under this heading is directed toward a better understanding of the processes by which listeners discriminate between pairs of complex auditory waveforms. The waveforms are, in all cases, samples of broad-band, white, Gaussian noise, and all experiments made use of a same-different paradigm. [Portions of the work described here have been reported in Fallon and Robinson (1985) and in Fallon and Robinson (1987).]

6.1. Effect of random variations in level

If the discrimination between pairs of noise bursts is based on a statistic such as total power, average power or energy, the discrimination should be impossible if overall level is randomized between the two bursts in the same-different paradigm. The effect of such a change was investigated at each of two durations using bursts which were either identical ("same" trials) or completely independent ("different" trials). Within a block of trials, the noise bursts were either 25- or 150-msec in duration and the level of the sample presented in one observation interval was held constant while the level of the sample presented in the other interval was randomly varied. In one experimental condition, the level of one of the samples in the pair was 3 dB greater than, 3 dB less than, or equal to the level of the other sample. The effect of a variation in level of $\pm$ 6 dB was also examined.

The data indicate that varying the level of one of the samples in a pair caused a only slight decrease in discriminability. When the bursts were 150-msec in duration, the average value of d' without variations in level was 2.98; with a $\pm$ 3 dB variation, it was 2.46; and with $\pm$ 6 dB, 2.09. For 25-msec bursts, the corresponding values of d' were 3.13, 2.79, and 2.49. Thus, although there is a slight decrease in performance with randomized levels, the samples are still quite discriminable. We conclude that the basis of the discrimination cannot be average power or energy.

6.2. Effect of temporal position of appended noise

Hanna (1984) demonstrated that samples of wide-band reproducible noise are highly discriminable over a large range of durations. We have found that discriminability can be reduced by increasing the similarity between the pairs of samples to be discriminated. During "different" trials of the same-different procedure, the second sample of the pair was generated by repeating a temporal segment of the sample presented in the first interval and combining it with a new sample of noise. The total duration of the second sample of a pair is equal to the duration of the new sample plus the duration of the repeated sample of noise. The total duration, as well as the duration and position of the new segment of noise was varied. The three total durations examined were: 150, 50, and 25 msec. The new segment of noise was either placed at the beginning, in the middle, or at the end of the repeated sample of noise.

The degree of similarity between the two samples presented during a "different" trial may be expressed in terms of the inter-pair correlation (r): the duration of the repeated sample of noise divided by the total duration of the sample. When the data are expressed in terms of correlation, the threshold value of r is independent of duration, but is highly dependent upon the position of the appended segment. Although discriminability was not affected by the total duration of the sample, the temporal position of the new segment had a large and consistent effect: segments placed at the end were more discriminable than those in the middle which were more discriminable than those at the beginning.

The effect of temporal position on discriminability also occurs with tonal sequences. Watson and his colleagues (Watson et. al., 1975, 1976) showed that discriminability increases as the location of the test tone is moved from the beginning to the end of a 450 msec tonal pattern. Hanna (1984) also determined that the discriminability of two samples of reproducible noise was dependent on the temporal positions of the repeated and appended segments. Hanna's data indicate that discriminability is best in the end condition, decreases in the beginning condition, and is worst in the middle condition. Based on the results of the present experiment as well as on the research of Watson and his colleagues, one would have predicted the middle condition to be more discriminable than the beginning condition. The discrepancy may be attributable to procedural differences such as the duration of the samples of noise or the degree of stimulus uncertainty.

The results of this experiment and the data of Watson's group indicate that the processes underlying the discriminability of sequences of tonal patterns and the discriminability of samples of reproducible noise are very similar. The just-detectable segments of "different" noise in these experiments tend to be a constant proportion of the total stimulus duration. This result is very similar to the performance described for various duration tonal patterns, in the "capacity" experiments discussed by Watson and Foyle (1985a) and by Watson and Kidd (1987). The fact that two distinctly different types of complex waveforms appear to be processed in the same manner suggests that discriminability is dependent on the more global characteristics of the complex waveform rather than on the fine structure of a specific waveform.

6.3. Effect of decorrelation: autocorrelation

This experiment investigated the discriminability of noise-samples which differed in their autocorrelation. As in the previous experiment, "same" trials were generated by repeating in the second interval the sample presented in the first interval. On "different" trials, however, the sample presented in the second interval was generated by deleting the first T-msec of the sample from the first interval and appending T-msec of independent noise to the end. In the experiment described in Sec. 6.2, new noise was appended at the beginning, middle, or end. The data from the "end" condition of that experiment are very similar to those from the present experiment. The two conditions are similar in that in each, independent noise is appended at the end of the 150-msec burst. The two conditions differ, in that, for the 'end' condition, samples in the two intervals are identical for the duration T_c, while for the autocorrelation experiment, the beginning segments differ. Since, as was pointed in Sec. 6.2, differences between samples which occur at the beginning or in the middle have only a small effect on discriminability, it is not surprising that the 'end' and the autocorrelation conditions are similar.

6.4. Effect of decorrelation: added noise

The correlation between pairs of noise samples may also be reduced by reducing the proportion of variance common to the two samples. In this experiment, "same" trials were generated by presenting identical samples of noise in both observation intervals. "Different" trials were generated by adding a new, independent, sample of noise to the sample which had been presented during the first observation interval. The relative levels of these two samples determined the Pearson product-moment correlation coefficient between the samples presented in the two intervals. The overall level of the samples in the two intervals was maintained at 50 dB SPL/Hz.

For all four of the durations examined, discriminability decreased as the correlation increased. The decrease was slight for correlations between 0.00 and 0.75, and very rapid for correlations greater than 0.75. This is to say that two samples are easily discriminable when they have less than about 50% common variance.

6.5. Effect of the temporal position of a decorrelated segment

In this experiment "different" trials were generated by decorrelating, as in Sec. 6.4, only a portion of the 150-msec waveform presented in the second interval. The decorrelated portion was located at either the beginning, the middle, or the end of the waveform. As expected from previous experiments, discriminability is highly dependent on the temporal position of the decorrelated segment. When the correlation was 0.00, threshold durations were approximately 25-, 60-, and 90-msec for segments at the end, middle and beginning. When the correlation was 0.75, these values had increased to approximately 50-, 90-, and 120-msec. The large effect of temporal position which we reported previously is still maintained as correlation is increased.

6.6. Effect of gap duration and position

In this experiment, the overall duration of the bursts of noise in a pair was 150 msec and either a 25 msec segment of new noise was appended to the end of the burst or a 50 msec sample was appended to the beginning of the burst. A silent interval or gap replaced a portion of the repeated segment either immediately following or immediately prior to the appended segment. The duration of the gap was gradually increased until only 5 msec of the repeated segment remained. Although discriminability increased as gap duration increased the presence of a brief repeated segment temporally separated from the appended segment by 90-120 msec caused a large decrement in performance. For example, when each burst in a pair consisted of a 5 msec repeated segment followed by a 120 msec gap and a 25 msec appended segment, the average P(C) is 0.72. If the 5 msec repeated segment was not present and the pair of 25 msec bursts was presented in isolation the overall P(C) increased to 0.88. It would appear, then, that interactions occuring after such a long silent interval are unlikely to be to peripheral sensory interactions, as was suggested by Hanna (1984).

7. Information integration: multiple observations and internal noise (D. E. Robinson and B. G. Berg)

The work described in this section began with two major goals. The first is to understand the processes by which humans integrate information over time or over channels. The "multiple look" problem is the basis for our initial work in this area. The basic question is, "How much additional information is gained by allowing observers more than one observation in a detection or discrimination task?" The second goal is to develop and evaluate models of "internal noise." The amount and rate of improvement in performance with an increasing number of observations will depend not only upon the amount of internal noise, but upon the level of processing at which the internal noise is added. The following section describes our attempts to describe the improvements that occur with multiple observations and to model the processes that lead to such improvements. [Portions of the work described here are reported in Robinson and Berg (1986), in Berg and Robinson (1987), Berg (1987), and Sorkin, Robinson, and Berg (1987).]

7.1. Internal noise model

Previous research has demonstrated that performance in signal-in-noise detection tasks improves as listeners are allowed more observations (Swets, et al., 1959; Swets and Birdsall, 1978). According to signal detection theory, the rate of improvement is a function of the square-root of the number of observations:

$$d'_n = (m_2 - m_1) / (v_{ext}/n + v_{int}/n)^{½} = (n)^{½} d'_1 \quad (1)$$

where:
- d'_n, d' after n-observations
- d'_1, d' for one observation
- m_1, the mean of the noise-alone distribution
- m_2, the mean of the signal-plus-noise distribution
- v_{ext}, the common variance of the N and SN distributions
- v_{int}, the variance of the internal noise.

Internal noise is assumed to be added prior to the formulation of a decision statistic. This derivation of the square-root-of-n rule assumes that the decision statistic is the mean of the n likelihood ratios (or any monotonic transformation of the likelihood ratios) obtained from the n observations. Previous research has supported this square-root-of-n prediction. However, the earlier work provided only a limited test of the model, since n never exceeded six. Our research has extended this work by using several paradigms and a greater number of observations.

7.2. Sequential presentation of n tones

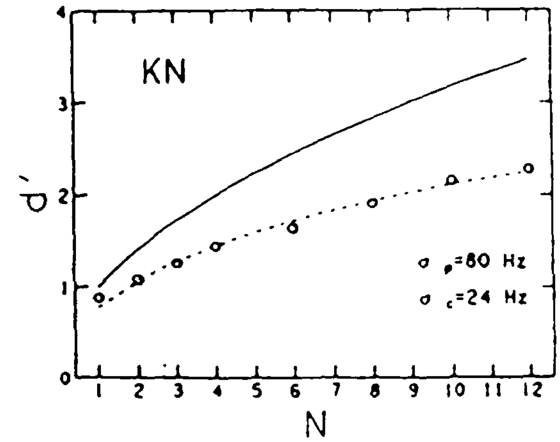
Consider the following task. There are two probability density functions on frequency, one with a mean of 1000 Hz, one with a mean of 1100 Hz, and both with a common standard deviation of 100 Hz. On each trial, n independent samples are selected from one of the distributions and presented sequentially over headphones as n, 50-msec tone bursts, separated by 50 msec silent gaps. The listener's task is to decide from which of the two distributions the n tones were sampled. Our results indicate that listeners can approach the theoretical d' for n=1, but do not follow the square-root-of-n rule, even for small n. Representative data for one subject are shown in Figure 7.2a. The solid line represents the predictions of Equation 1.

One interpretation of the model, described by Equation 1, is that some amount of internal noise is added to each observation prior to generating a decision statistic. Once the decision statistic is obtained, no additional variance is assumed. This model allows no parsimonious account of variance introduced by uncertainty of the decision criterion, changes in response bias, or memorial factors associated with the decision statistic. The model can be extended by allowing additional variance after the generation of the decision statistic. This "partitioned variance" model is represented by the equation:

$$d'_n = (m_2 - m_1) / (v_{ext}/n + v_p/n + v_c)^{1/2} \qquad (2)$$

where: v_p, the variance of the peripheral noise
v_c, the variance of the central noise.

In this model, internal noise is added at two stages: (1) at the periphery, before a decision statistic is formed and (2) centrally, after the statistic is formed. The dashed line in Figure 7.2a represents the function obtained for subject KN by a least-squares estimate of the two parameters. Similar fits were obtained for the three other subjects.



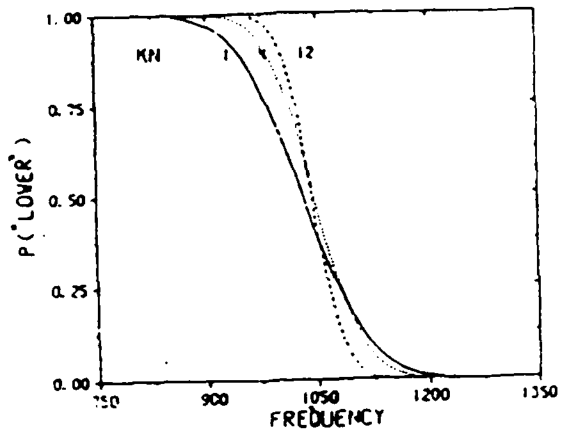


Figure 7.2a 7.2b

There is a second method of estimating the two parameters. Consider the function relating the probability of reporting "lower distribution" to the mean frequency of the sample. An ideal observer would generate a step function; when the mean frequency was less than the criterion, the ideal would report "lower", and would report "higher" when the mean exceeded the criterion. Within the model, any deviation from this step function can be attributed to internal noise. The variance of this internal noise can be estimated by fitting a normal ogive to the obtained data. Figure 7.2b shows the best fitting functions for subject KN for sample sizes of 1, 4, and 12.

The slope of the functions increase with increasing n, indicating that the total internal variance is decreasing. In this manner, estimates of the total internal variance were obtained for each sample size (n=1,2,3,4,6,8,10 and 12). Estimates of the peripheral and central variance were obtained by a least-squares fit to the equation:

$$v_{tot} = v_p/n + vc . \qquad (3)$$

A comparison of the parameter estimates obtained with Equations 2 and 3 showed remarkably good agreement for all four subjects. The partitioned variance model thus provides a reasonably good account of the data, and represents an improvement in the formal treatment of internal variance.

7.3. Simultaneous presentation of n tones

This work was conducted in collaboration with Dr. Wesley Grantham of the Bill Wilkerson Hearing and Speech Center, of Vanderbilt University. Grantham conducted an experiment similar to that described in Sec. 7.1, with the exception that the n tones were added and presented simultaneously, rather than sequentially. Data could be reasonably described by Equation 1. That is, little or no central variance was required. Preliminary conclusions seemed to indicate a fundamental difference between the processing of information presented sequentially and information presented simultaneously. However, this difference can be attributed to a procedural difference between the two studies. For technical reasons, tones were sampled without replacement for simultaneous presentation, whereas sampling was done with replacement for sequential presentations. Equation 1 assumes independent sampling and is valid for the sequential tones study, but a correction factor is required to obtain predictions when sampling is done without replacement. Obtaining fits to Grantham's data using this correction factor indicated a less than optimal growth rate in d' for all three subjects, and required the addition of central variance. A comparison of estimates of peripheral and central variance across the two studies showed relatively good agreement.

7.4. Distribution of internal noise over the tonal sequence

An important question raised by our general model is whether information from each tone in a tonal sequence is equally weighted in determining an observer's decision. We have developed a technique for assing how internal noise is distributed over the N-elements of a tonal sequence. In terms of the model as described in Eq. 2, the amount of information obtained from different tones in an N-element sequence will be reflected in the variance of the internal noise added at each temporal position. If a particular temporal position contributes little to the final decision, that position will be found to have a large amount of internal noise associated with it. If, on the other hand, a particular element contributes a great deal, that element will have less internal noise associated with it. Data from an auditory experiment were analyzed to assess how internal noise is distributed over successive temporal positions. Over many thousands of trials we store the frequency of the tones actually presented in the ith temporal position ($i = 1, 2, \ldots n$; where n is the number of tones in the sequence). We then partition these stored frequencies into bins of arbitrary width. The purpose of our analysis is to keep track of the number of trial events on which the frequency of the ith element was in each frequency bin. For each bin and each temporal position, we then compute the probability that the subject responds that the sequence came from the lower distribution. Cumulative normal distributions are then fit to the resulting ogives. The standard deviation of the best fitting normal distribution is then an estimate of the standard deviation of the total internal noise limiting performance at each display position.

Figure 7.4 shows the standard deviation of the internal noise as a function of temporal position. The parameter on the figure is the total sequence length, n. If each element in the sequence contributed equally to the final decision, the lines in Figure 7.4 would be horizontal. It is clear that the last tone in a sequence contributes more to the final decision than do tones in the middle, which contribute less than those near the beginning of the sequence.

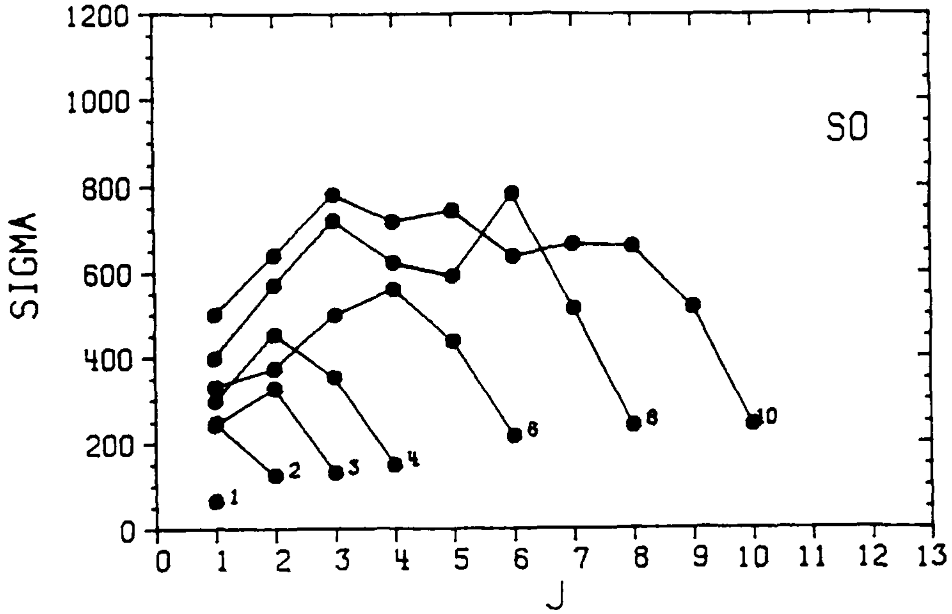


Figure 7.4

7.5. Additive or multiplicative internal noise?

A common assumption of the models discussed above is that external and internal variance are independent and additive. This assumption was tested by using a within-subjects factorial design consisting of two levels of external variance

$$(v_{ext})^{1/2} = 100 \text{ Hz or } 150 \text{ Hz}$$

and two levels of the mean frequency difference between the two distributions.

$$m_2 - m_1 = 100 \text{ Hz or } 150 \text{ Hz.}$$

For each of the four conditions, the experimental procedure was identical to the sequential tone paradigm described previously. Data obtained from four listeners indicate that estimates of internal variance are not affected by changes in the mean frequency difference for a fixed level of v_{ext}. However, estimates of internal variance increase when v_{ext} is increased. This increment in internal variance is obtained for both levels of the mean frequency difference. These data violate the assumption of additivity, and suggest that internal variance increases as a function of the external variance.

8. Computer Assisted Detection (D. E. Robinson and B. G. Berg)

In this work we consider a person-machine system consisting of an automated alarm and a human monitor. The task of the human is to monitor a noisy channel on which information about a potentially dangerous condition may appear. The alarm system monitors an independently noisy channel for information about the same threatening condition. Using basic concepts of statistical decision theory, a Contingent Criterion Model of such a person-machine system has been developed. According to the model, the human should establish two criteria for responding: one contingent on an alarm from the automated detector and one, on no-alarm. The model shows large gains in performance compared to either detector alone. The degree to which human subjects perform in the manner suggested by the model has been evaluated in two experiments: a simple auditory detection task and a scrolled letter detection task. In both experiments, the subjects were aided by a simulated alarm system. Although not reaching the performance levels possible under the model, the behavior of the subjects is well described by the model.

In our initial development of the Contingent Criterion Model, we assumed that the noise in the alarm-system channel is uncorrelated with that in the channel monitored by the human operator. Such an assumption is probably unrealistic. We have since investigated the degradation in performance which occurs with increased correlation between the two channels. The predictions of the model indicate that, although there may be a considerable performance decrement when the correlation is near unity, the model is quite robust, and system performance can exceed that of either detector alone even with correlations as high as 0.50.

More recently we have attempted to expand the Contingent Criterion Model to include signal classification (identification) as well as signal detection. This effort draws on the work of Nolte (1967) Nolte and Jaarsma (1967), Green and Birdsall (1978), and Starr, Metz, Lusted, and Goodenough (1975). Our efforts to date suggest that the performance of a system consisting of a human operator and an automatic signal classifier can be significantly improved compared to either subsystem operating alone.

There are two important observations that may be drawn from this work. First, although combined system performance (human-plus-automated detectors) was less than optimum, no effort was made to train subjects in the proper placement of their criteria. It is possible that human operators can be trained to use the available information more efficiently, and to set criteria which will lead to more nearly optimum system performance. A second observation is that system performance is dependent not only upon the behavior (sensitivity and criterion placement) of the human operator, but also upon the criterion (threshold or alarm set-point) of the automated alarm system. System performance may be improved by changing this parameter of the system. This work, portions of which have been previously reported [Sorkin and Robinson (1985), Robinson and Sorkin (1985),and Sorkin, Robinson, and Berg (1987)] was done in collaboration with Dr. Robert D. Sorkin of Purdue University. (Additional support for this work was provided by contracts with the U. S. Department of Transportation and with the U. S. Naval Weapons Center, China Lake, CA.)

9. Studies of the relation between auditory abilities measured with speech and non-speech stimuli

9.1. Temporal acuity and category boundaries for speech and non-speech stimuli

(Kewley-Port, Watson, and Foyle)

Previous research has demonstrated that discrimination functions for stop consonants differing in VOT are non-monotonic in high-uncertainty tasks such as the ABX paradigm, but may become less "categorical" as the level of task uncertainty is decreased (Repp, 1984) In particular, Sachs and Grant (1976) reported that the perception of a /ga-ka/ continuum became monotonic after considerable training in a low-uncertainty, same-different task, although that work has not been published.

A series of experiments was undertaken to replicate the Sachs and Grant results, using a bilabial VOT continuum ranging from +5 to +75 msec in 10-msec steps. Standard ID and ABX tasks showed the typical functions which demonstrate categorical perception where the best discrimination was obtained for the stimulus pair at the boundary between /ba/ and /pa/. The next experiment used was a same-different task which was also high-uncertainty (pairs randomly drawn on each trial from the /pa/-/ba/ continuum), but which had a slightly reduced memory load. In this, and on subsequent tasks, subjects were given feedback for the correct response and run for 3-5 test sessions until asymptotic performance was achieved. The high-uncertainty S-D (same-different) results were monotonically decreasing and showed the greatest improvement in performance for the short VOT stimuli.

Finally a series of minimal-uncertainty S-D tasks were employed in which only one VOT pair was presented in each block. Discrimination was best for all VOT pairs under minimal uncertainty. Functions obtained from both of the same-different tasks indicate that VOT discrimination follows Weber's Law. That is, discrimination for VOT is not different from that for other acoustic stimulus variables, such as intensity or frequency, where discrimination for a fixed increment in a stimulus is better for smaller values of the stimulus parameter.

Categorical perception of a non-speech continuum has been reported by Miller et al., (1976). Their stimuli, consisting of a short noise burst followed by a longer buzz, were non-speech analogues of VOT, stop+ vowel stimuli. In order to examine whether results from the VOT studies were special for speech continua or would generalize to the discrimination of onsets in non-speech continua, a similar series of experiments was conducted for a noise-lead-time (NLT) continuum which closely matched that of Miller et al.

The NLT stimuli were tested in identification and ABX tasks. Results showed that the categorical peak previously reported for noise-buzz analogues of stop consonants were not nearly as robust a phenomenon as for stop consonants. Apparently earlier reports of categorical perception of these nonspeech stimuli were dependent on certain special features of the testing methods employed. Discrimination of NLT in a same-different task under high-uncertainty and minimal-uncertainty replicated the monotonically decreasing functions obtained for the VOT stimuli. In fact, when the percent correct discrimination results under minimal uncertainty were converted to Weber fractions, nearly identical values were obtained (0.19 for VOT and 0.17 for NLT).

This set of studies with analogous speech and non-speech continua demonstrates that basic auditory discrimination abilities for the discrimination of temporal onsets follow familiar psychophysical laws. Furthermore, it now appears that categorical discrimination is the outcome of a more central level of auditory processing, apparently employed when listeners have extensive familiarity with the stimuli and the categories with which they are to be identified.

9.2. Psychoacoustic studies of isolated vowels (Kewley-Port, C. Watson, and Czerwinski)

Only a few studies have been conducted to determine the detectability of speech sounds, the discriminability of frequency differences in formants, and in general, listeners' abilities to "hear out" fine acoustic detail in the waveform of speech It is essential that we determine the limits of processing for spectral, temporal, and intensive properties of speech sounds in isolation if we are to gain a better understanding of their processing within the context of other speech sounds.

A series of experiments has been designed to investigate some of the psychoacoustic properties of vowels. The first step was to synthesize a set of steady-state vowels which would be identifiable by listeners in the same manner as natural vowels, i.e. with some natural confusions. A set of ten vowels was created using the Klatt synthesizer, based on steady-state formant values measured from spectrograms of vowels spoken by a female talker. Because the vowels were to be used in another experiment (see below), vowel duration was set to 40 msec., with 5-msec onset and offset ramps. A confusion matrix was obtained which showed identification performance (82% correct) for the vowel set comparable to those in the literature (e.g. Assmann et al., 1982).

Another study was undertaken to determine the effects of vowel duration on vowel identification. The nine durations used in this study were: 20, 40, 60, 90, 140, 200, 300, 685, and 1000 msec. Subjects' identification performance was fairly constant across all nine durations, although long vowels were identified somewhat more accurately between 200-685 msec., while short vowels were identified more accurately around 60-140 msec.

The threshold of detectability of three sets of 10 vowels was measured using an adaptive tracking paradigm. The vowels were equated for sound pressure at the earphones. In the first study, thresholds were obtained for the ten synthetic vowels at durations of 20, 40, 80, 160, and 320 msec. Results showed that the thresholds differed considerably across vowels. The data for the 40 ms duration are represented by the dashed line in Figure 9.3.1. In particular, the vowel /u/ had a threshold almost 20dB higher than the other vowels. The same pattern of differences in detectability across vowels was obtained at all durations.

Two new sets of natural vowels were also studied. One male and one female spoke the 10 English vowels very slowly in isolation. Using a digital waveform editor, 20, 40 and 160 ms segments were excised. Again, the vowels were equated for sound pressure at the earphones. Thresholds for each vowel, obtained for a new group of subjects, are shown in Fig. 9.2. The same pattern of thresholds was obtained for each vowel set at each duration. However, different threshold patterns are associated with the male, female and synthetic vowel sets.

Several different analyses are in progress to attempt to explain the differential detectability of equal-SPL vowels. Smoothed spectra from linear prediction analysis were used to determine the frequencies and amplitudes of the spectral energy peaks (FO and formants) for each vowel. Multiple linear regression showed that better than 90% of the variance in the threshold patterns can be predicted from the amplitudes of first three spectral peaks. Further analyses of the vowels modeled as cochlear excitation patterns will be made.

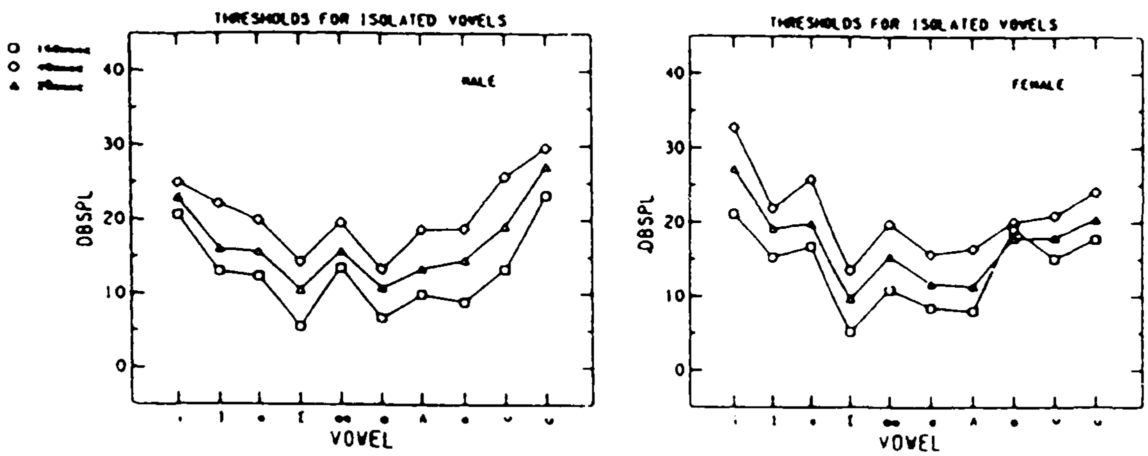


Figure 9.2

9.3. Vowel sequence studies (Kewley-Port, Watson, Hackett)

Previous research has investigated complex auditory stimuli consisting of a series of brief tones presented one after the other to form single "word-length" (about one-half-sec) tonal patterns (Watson et. al., 1976). We have now replicated some of those experiments with speech stimuli. One goal of these new experiments is to learn whether the same factors that govern the discriminability of patterns of pure tones have similar effects with spectrally complex stimuli. Vowel patterns consisting of sequences of steady-state vowels were chosen to begin these investigations.

In a series of long-term training studies, we have investigated the thresholds of detectability of single vowels in vowel patterns. These experiments were designed to be analogous to Watson and Kelly's (1981) informational masking study using ten-tone patterns. Each pattern consisted of the ten synthetic vowels described previously. In an adaptive-tracking S/2AFC detection task, a target vowel was replaced by a silent gap in two of the three pattern presentations on each trial.

9.3.1. High Uncertainty

In the first experiment we investigated the detectability of vowels under very high-uncertainty. In this condition a new vowel sequence was presented on each trial. Four subjects were tested one and one-half hours a day for 15 days. Subjects approached asymptotic performance after approximately 6 days. Thresholds differed only slightly across temporal positions, but differed considerably for different subjects and vowel types. Two important results were obtained. First, individual subjects differed greatly in the thresholds obtained for vowels in patterns, about 40dB from best to worst subject. Second, vowel thresholds followed the same pattern seen for the isolated vowels. These results are similar to those obtained for the ten-tone patterns, in terms of the time-course of learning, differential detectability of the stimuli, and unusually large individual differences among the normal-hearing listeners.

In the second study, the effect of stimulus high versus minimal stimulus uncertainty on the detectability of vowels in vowels sequences was examined. In the high-uncertainty condition a catalogue of 48 ten-vowel patterns was constructed. In the minimal-uncertainty condition, asymptotic detection thresholds were for a single pattern.

Four new subjects were recruited for the experiment, one of whom dropped out after a week, while the remaining three completed all conditions. The S/2AFC adaptive-tracking task was used again. First subjects participated in the high-uncertainty task in which one of the 48 patterns was presented randomly on each trial. Thresholds obtained after 8 one-hour sessions are shown in Figure 9.3.1 separately for each subject and each vowel tested (averaged over position in the sequence). Results were similar to the very-high uncertainty experiment in that individual differences in subject's thresholds varied greatly, in this experiment over a 25dB difference.

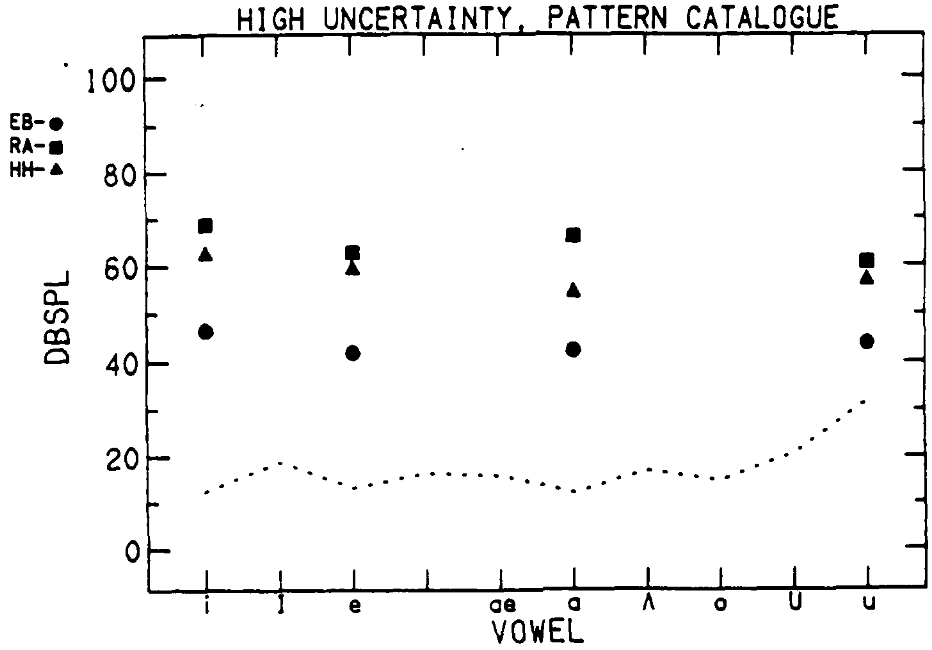


Figure 9.3.1

9.3.2. Minimal Uncertainty

Subsequently subjects participated in a minimal-uncertainty task in which thresholds were estimated for a single component in one pattern at a time. In all, three different patterns were tested. Results were analyzed in terms of the difference between the detection thresholds obtained for vowels in isolation, and those obtained in patterns under minimal uncertainty. This difference in decibels represents the amount of *informational masking* contributed by the presence of the vowels in the pattern in comparison to the detection of the vowels in isolation (see Watson and Kelly, 1981). The amount of masking obtained from these subjects under minimal uncertainty ranged from 10dB to 40dB, depending on vowel and position within the vowel sequence. Analogous research on tonal sequences (Watson and Kelly, 1981; Watson, 1987) obtained a similar pattern of results, although 40dB of masking for vowels was somewhat higher than for tones. Since there were a few methodological differences between the vowel and tonal sequence studies, some additional experiments were conducted.

Subjects in the first three minimal uncertainty experiments seemed to approach asymptotic performance after about 600 trials. Detection thresholds were therefore estimated from (approximately) the second 600 trials. However, compared to many of the long-term training studies conducted with tonal sequences (see Leek and Watson, 1984), 1200 trials does not constitute a great deal of training for tasks involving complex sounds. Furthermore, the interstimulus intervals in the S/2AFC trial structure were somewhat longer than in the tonal studies. For these reasons it was decided that two new groups of subjects would participate in long-term training studies in the minimal-uncertainty task. A shorter ISI was also used in the S/2AFC trial structure.

Two patterns previously tested under minimal uncertainty (the easiest and hardest) were tested here. Both groups os subjects showed 10 to 15dB improvements over the amounts of masking obtained in the previous experiments, after about 2000 trials. Thus the reduction in the amount of masking obtained between high-uncertainty and minimal-uncertainty conditions in sequences of sounds appears to be similar for sounds which are spectrally simple (tones) or spectrally complex (vowels). These experiments, in general, support a theory of speech perception that assumes that subtle, high-information bearing portions of speech waveforms may become salient as a result of prolonged experience rather than because of any inherent properties of the acoustical waveforms of speech.

10. Relations between auditory capabilities and phoneme perception (B. Espinoza-Varas, C. Watson, Srygler)

During the past year we examined correlations between a variety of measures of phoneme perception, and measures of discrimination ability. Measures of phoneme perception are obtained with the CUNY Nonsense Syllable Test (Dubno, J.R. and Levitt, H., *J. Acoust. Soc. Am.*, **69**, 249-261, 1981), which requires to identify VC or CV syllables presented in the quiet, at 33 dB SPL. The tests of discrimination ability are those of the battery developed by Watson et al. 1982. They include discrimination of frequency (DF), intensity (DI), and duration (DT) of pure tones; discrimination of jitter in pulse trains (RHY), discrimination of temporal order of pure tones (TOD), and detection of a tone embedded in a tonal sequence (ETT). The stimuli are presented at 75 dB SPL. In the previous progress report, results of the speech test were described which showed large differences in identifiability across syllables of a given subtest (often ranging from chance to perfect), as well as strong response biases with some phonemes. These results suggested that a measure of speech processing based on the overall performance with the entire set of phonemes (i.e., the articulation score), may be a poor indicator of the ability to identify or to discriminate some individual phonemes. Such an overall score, may not be appropriate to examine correlations with auditory capabilities. In the current year, we correlated measures of auditory capabilities with several indices of phoneme perception. Specifically, we compared correlations using: a) overall vs phoneme specific measures; b) bias-free vs. bias confounded measures of phoneme identification; and c) measures of pairwise discriminability of phonemes. The present results are based on the first three subtests of the NST only, which contain VC syllables with the same seven consonants (/p,t,k,f,θ,s,ʃ/), but with different vowels, /a/, /i/, /u/.

Figure 10a shows the identifiability of each of the seven phonemes (abscissa), using two different indices. The phonemes are ordered in terms of accuracy of identification. One index is the standard percent correct or hit probability, (triangles) and the other is P(c)max, (circles) which takes into account both hits and false alarms, and is corrected for bias. The two indexes range from about 20% (/θ, f/) to about 100% (/ʃ, s/).

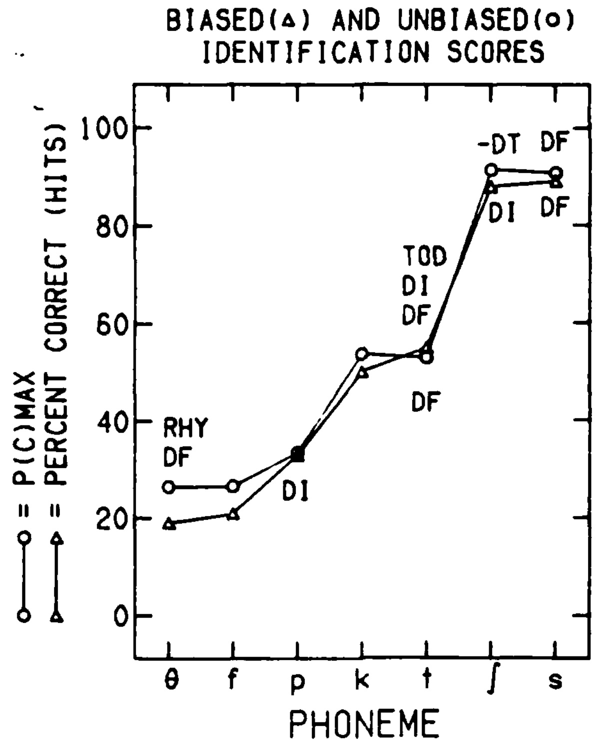


Figure 10a

The labels below the data points in Figure 10a represent the discrimination ability which showed significant Pearson correlations (p=.05) with the hit probability or percent correct for the respective phoneme. These correlations ranged from 0.3-0.5. The rate of significant correlations using the hit probability was about 10%, i.e., about twice more than the 5% to be expected by chance. The overall articulation score collapsed over all 61 phonemes of the NST, showed no significant correlations with any of the discrimination tests. The labels above the data points in Figure 10a show the Pearson correlations between the auditory discrimination measures and P(c)max for each phoneme. The rate of significant correlations for P(c) max is 17 %, that is more than three times the rate expected by chance. The correlation coefficients significant at 5% level ranged from .35-.51. Rank order correlations calculated using both hits and P(c)max yielded essentially the same rate of significant correlations (i.e. 10 and 17% respectively). Histograms of the score distributions for each of the phonemes fell in three groups: very low or very high P(C)max (/f,s/) or midrange P(C)max /k/. One obvious problem is that those phonemes that are either too difficult to identify (/f/ or /p/) or too easy (/s/ or /ʃ/) yield little or no individual differences in the scores, and thus may be lowering the correlations. Another measure of phoneme perception used for correlations was P(c)max for the discrimination between phonemes pairs. That is the discriminability between /p-t/ /t-s/ etc. Figure 10b shows P(C)max discrimination scores for each of the 21 phoneme pairs than can be formed with the set of consonants /θ,f,p,k,t,s,ʃ/, ordered from least to most discriminable. Each panel

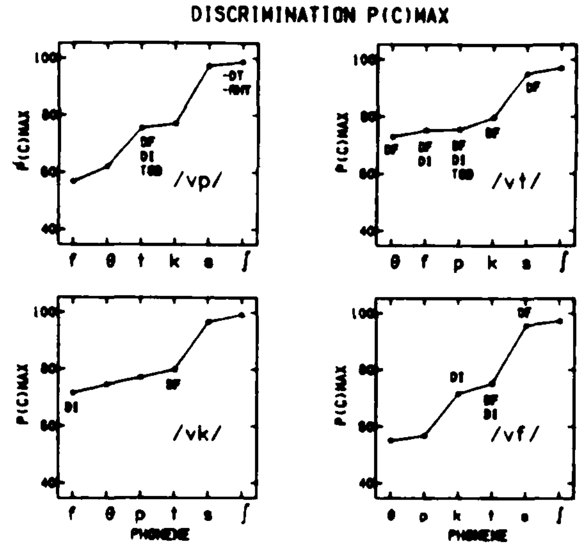


Figure 10b

shows the discriminability of the phoneme in the inset (e.g. /vF/; "v" denotes the average result for the three vowels) with any of the phonemes in the abscissa. The labels in the functions again indicate the correlations that are significant. This means for example that discrimination of "vPvT" depends on frequency, intensity and temporal order discrimination. Some pairwise discriminations yield higher rates of correlations than others (e.g. /vT/ yields about 22% significant correlations but /vK/yields only 5 %, which is no different from the chance rate). The average rate of significant correlations with the discrimination P(c)max was about 10 percent. The pairwise discriminability is generally high even though some of the phonemes have very low identification scores. For example /vF/, shows identification P(c) max of about 20% but is discriminated at better than 75 % with most other phonemes in the set. This may also be preventing the calculation of correlations: if discrimination is near perfect for most subjects, no individual differences are obtained.

Conclusions: 1) Speech measures based on the ability to process specific phonemes yield a higher rate of significant correlations with auditory capabilities than overall, whole-test measures of speech processing; 2) Relative to the number of significant correlations to be expected by chance, the number of significant correlations obtained with phoneme based measures is greater than chance, however it never exceeded 20 percent; 3) Correcting the measures of phoneme perception for response bias, or using discrimination rather than identification scores increased only slightly the rate of significant correlations obtained with the hit probability per phoneme; 4) Essentially the same results are obtained if rank-order rather than Pearson correlations, are calculated; 5) Possible factors that may prevent obtaining higher rates of significant correlations are:

a) near chance or near perfect performance with some phonemes; b) relatively high pair-wise discriminability given moderate or low identification levels; and c) redundancy and multidimensionality of the cues that differentiate phonemes.

11. The effect of sentence timing on the perception of word-initial stop consonants (G. R. Kidd)

The influence of temporal context on the perception of voice onset time (VOT) was examined in three experiments. Different versions of a 10-word precursor phrase were constructed by recording the phrase at a fast and a slow rate and then combining words from the two original phrases to produce composite phrases with various patterns of rate changes. A final (target) syllable from a 7-member /gi/-/ki/ VOT continuum constructed from natural speech was added to the end of the phrase after a variable pause duration. Subjects listened to each phrase version and judged the identity of the target syllable. In general, the VOT boundary was found to shift to shorter values with precursors that contained more fast words and shorter closure durations. Further, while the rate of the words immediately preceding the target appeared to have the greatest effect, there were also significant rate effects due to the rate of stressed words early in the precursor as well as effects due to the pattern of rate changes throughout the precursor. [Primary support from NIMH individual NRSA; AFOSR support for analysis and manuscript preparation.]

12. Individual Differences: TBAC

12.1. Intensive Training. (C. Watson, Czerwinski)

To investigate whether individual differences in specific auditory discrimination tasks are mainly the result of differences in learning rates, rather than in "hard-wired" individual differences in sensitivity or acuity, a sample of 47 college students was screened using two of the subtests of the BTN Test of Basic Auditory Capabilities (TBAC; Watson, et al., 1982a, 1982b). Three levels of ability were identified, expressed in z-score units relative to the means of the population: above average (>+2.00), average (+/-0.50), and below average (<-2.00). The subtests were the four-tone sequencing task, and the nine-tone pattern task in which the listener attempts to detect the presence of a single component, each of which include 72 S/2AFC trials, presented in approximately 7 minutes. The difficulty of both tasks is varied by using a range of durations of the tonal components. Two subjects from each performance category were trained intensively, with feedback, in an adaptive tracking version of these tasks for approximately 1200 trials on each subtest, spaced over 6 one-hour testing sessions. Although the original differences between the three performance groups were reduced, and all listeners improved, each pair of listeners was still significantly different from the others in terms performance on these temporal discrimination tasks. Clearly, more work of this kind is needed before we can draw firm conclusions about the origin of differences in performance. However, this first effort demonstrates that much more than 5-6 hours of intensive practice must be required if training is to eliminate individual differences measured in screening tests like the TBAC. The data available thus far support the hypothesis of permanent individual differences in temporal discrimination abilities.

12.2. Auditory Processing in Learning or Reading Disabled Populations (B. Watson, C. Watson, D. Goldgar)

(1). The relationship between reading performance and the ability to discriminate complex auditory stimuli has been examined in three samples including a learning-disabled group of adolescents, a control group of normal adolescents, and a group of college students. Significant correlations between the TBAC sub-scores and the reading subtests of the Woodcock Johnson Psychoeducational Battery have been found in each of the samples. The tonal-patterns task (Subtest 5) has been found to be correlated with reading abilities in each of the samples, and the temporal-order task (Subtest 6) was correlated with reading scores in the two normal groups. Additional analyses are being conducted to determine the contribution of general intelligence to these results.

(2). The hypothesized relationship between auditory processing abilities and language skills (Tallal and Piercy, 1973) was studied in a group of 87 students enrolled in first-semester foreign language courses at Indiana University. The students had taken the Modern Language Aptitude Test (MLAT) as a prerequisite to enrollment. This test, which correlates strongly with performance in learning foreign languages, contains two vocabulary acquisition tasks, and a test of grammatical knowledge. Initially 20 significant correlations were obtained between TBAC subtests and the subscores on the MLAT, of 54 combinations studied. Additional analyses suggest that this relation between language skill and auditory capabilities is largely a result correlating between both variables and I.Q. [These studies of the relation between language disorders and auditory processing were primarily supported by NIH/NINCDS.]

13. Facility Development

During the past two-and-one-half years of the present grant the experimental facilities in the Hearing and Communication Laboratory have been substantially improved. Originally HCL had a single 11/23 computer which was (and still is) heavily committed to running on-line experiments. In our 1984 proposal, because new research would greatly increase our commitments to on-line experiments, we requested funds for a new computer to assist with off-line support of laboratory activities such as stimulus generation, data analysis, program development and signal processing. While those funds were not awarded, we were able to build the needed system, a PDP 11/83, primarily through a grant which we subsequently received from NMRDC. In addition, the first year of the grant provided savings that, together with other funds, allowed us to purchase two Apollo workstations. These workstations are each powerful mini-computers which run UNIX. We participated in establishing Apollo Domain ring at Indiana together with an interdisciplinary group of investigators in Computer Sciences, Linguistics and Mathematics, all of whom had interests in speech and auditory processing. We have been able to share software applicable to the needs of this group, in particular statistical and signal processing packages, digitizing facilities, and speech recognition tools and algorithms. Several new workstations have now been added to this network by the AFOSR supported Institute for the Study of Human Capabilities. Interests between our labs and the Institute overlap considerably, and the communication with these additional investigators has further enhanced the usefulness of the Apollo system. We are fortunate to have been able to build a state-of-the-art psychoacoustic and speech laboratory over the past few years. That system now enables us to conduct a wide range of extensions of the research supported here, without need for additional apparatus, at least in the near future.

References

Assman, P.F., Neary, T.M., and Hogan, J.T. (1982). Vowel identification: Orthographic, perceptual, and acoustic aspects. *J. Acoust. Soc. Am., 71,* 975-989.

Berg, B. G. (1987). *Internal Noise in Auditory Detection Tasks.* Ph. D. Dissertation, Indiana University.

Berg, B. G. and Robinson, D. E. (1987). Multiple observations and internal noise. *Journal of the Acoustical Society,* 81, S33.

Bregman, A.S. (1978). The formation of auditory streams. In J. Requin (Ed.), *Attention and Performance, VII.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Fallon, S.M., and Robinson, D.E. (1985). The effects of changes in correlation on the discriminability of noise samples. *J. Acoust. Soc. Am., 78,* S46.

Fallon, S. M. and Robinson, D. E. (1987). The effects of decorrelation on the discriminability of noise samples. *Journal of the Acoustical Society,* 81, S33.

Green, D.M. (1983). Profile analysis, a different view of auditory intensity discrimination. *Am. Psychol., 38,* 133-142.

Green, D.M., and Birdsall, T.G. (1978). Detection and recognition. *Psych. Rev., 85,* 192-206.

Guttman, N., Julesz, B. (1963). Lower limits of auditory periodicity analysis. *J. Acoust. Soc. Am., 35,* 610.

Hanna, T.E. (1984). Discrimination of reproducible noise as a function of bandwidth and duration. *Perception and Psychophysics, 36,* 409-416.

Leek, M.R., and Watson, C.S. (1984). Learning to detect auditory pattern components.

J. Acoust. Soc. Am., *76*, 1037-1044.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psych. Rev.*, *63*, 81-97.

Miller, J.D., Wier, C.C., Pastore, R.E., Kelly, W.M., and Dooling, R.M. (1976). Discrimination and labeling of noise-buzz sequences with varying noise lead times: An example of categorical perception. *J. Acoust. Soc. Am.*, *60*, 410-417.

Nolte, L. W. (1967). Theory of signal detectability: Adaptive optimum receiver design. *Journal of the Acoustical Society*, 42, 773-777.

Nolte, L. W. and Jaarsma, D. (1967). More on the detection of M orthogonal signals, *Journal of the Acoustical Society*, 41, 497-505.

Pollack, I. (1953). The informational content of elementary auditory displays. *J. Acoust. Soc. Am.*, *25*, 765-769.

Repp, B.H. (1984). Categorical perception: Issues, methods, findings. In N.J. Lass (Ed.), *Speech and Language: Advances in Basic Research and Practice, Vol 10.* New York: Academic Press, 243-335.

Robinson, D.E., and Berg, B.G. (1986). A partitioned variance model for multiple observations. Paper presented at the Nineteenth Annual Meeting of the Society for Mathematical Psychology, Cambridge, MA.

Robinson, D.E., and Sorkin, R.D. (1985). A contingent criterion model of computer assisted detection. In R. Eberts and C.G. Eberts (Eds.), *Trends in Ergonomics/Human Factors, Vol.II*, North-Holland: Amsterdam.

Sachs, R.M., and Grant, K.W. (1976). Stimulus correlates in the perception of voice onset time (VOT): II. Discrimination of speech with high and low stimulus uncertainty. *J. Acoust. Soc. Am.*, *60*, S91.

Sorkin, R.D., and Robinson, D.E. (1985). Alerted Monitors: Human Operators Aided by Automated Detectors, DOT/OST/P-34/85/021, U. S. Department of Transportation.

Sorkin, R. D., Robinson, D. E., and Berg, B. G. (1987). A detection theory method for the analysis of visual and auditory displays. *Proceedings of the Human Factors Society*, In press.

Spiegel, M.F., and Watson, C.S. (1981). Factors in the discrimination of tonal patterns. III. Selective attention and the level of target tones. *J. Acoust. Soc. Am.*, *69*, 223-230.

Starr, S.J., Metz, C.E., Lusted, L.B., and Goodenough, D.J. (1975). Visual detection and localization of radiographic images. *Radiology, 116*, 533-538.

Swets, J.A., and Birdsall, T.G. (1978). Repeated observation of an uncertain signal. *Perception and Psychophysics, 23*, 269-274.

Swets, J.A., Shipley, E.F., McKey, M.J., and Green, D.M. (1959). Multiple observations of signals in noise. *J. Acoust. Soc. Am., 31*, 514-521.

Tallal, P., and Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature (London), 241*, 468-469.

Watson, C.S. (1987). Uncertainty, informational masking, and the capacity of immediate auditory memory. In W.A. Yost and C.S. Watson (Eds.) *Auditory Processing of Complex Sounds*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Watson, C.S. (1976). Factors in the discrimination of word- length auditory patterns. In S.K. Hirsh, D.H. Eldredge, I.J. Hirsh, and S.R. Silverman (Eds.), *Hearing and Davis: Essays Honoring Halowell Davis*. St. Louis: Washington University.

Watson, C.S., and Foyle, D.C. (1983). Temporal and capacity limitations of auditory memory. *J. Acoust. Soc. Am., 73,* S44(A).

Watson, C.S., and Foyle, D.C. (1985a). Central factors in the discrimination and identification of complex sounds. *J. Acoust. Soc. Am., 78,* 375-380.

Watson, C.S., and Foyle, D.C. (1985b). Capacity limitations for the discrimination of isochronous and anisochronous tonal patterns. *J. Acoust. Soc. Am., 78,* S46.

Watson, C.S., Jensen, J.K., Foyle, D.C., Leek, M.R., and Goldgar, D. (1982). Performance of 146 normal adult listeners on a battery of audit discrimination tasks. *J. Acoust. Soc. Am., 71, Suppl. 1,* S73.

Watson, C.S., Johnson, D.M., Lehman, J.R., Kelly, W.J., and Jensen, J.K. (1982). An auditory discrimination test battery. *J. Acoust. Soc. Am., Supp. 1, 71,* S73.

Watson, C.S., and Kelly, W.J. (1981). The role of stimulus uncertainty in the discrimination of auditory patterns. In D.J. Getty and J.H. Howard (Eds.), *Visual and auditory patterns*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Watson, C.S., Kelly, W.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., 60,* 1176-1186.

Watson, C. S. and Kidd, G. (1987).Proportional target-tone duration as a factor in the discriminability of tonal patterns. *Journal of the Acoustical Society*, 82, S40.

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

Publications

Berg, B. G. and D. E. Robinson (1987) "Multiple observations and internal noise." *J. Acoust. Soc. Am.*, 81, S33.

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

Espinoza-Varas, B., D.G. Jamieson and J. Wahn (1984) "Perception of critical-band adjusted vowel continua by sensorineural hearing impaired listeners." *J. Acoust. Soc. Am.*, 76, S80.

Espinoza-Varas, B., and C.S. Watson.(1985) "Basic auditory capabilities and resolving power for phonemes." *J. Acoust. Soc. Am.*, 78, S47.

Espinoza-Varas, B. (1986) "Levels of representation of phonemes and bandwidth of spectral and temporal integration." In: M.E.H. Schouten (Ed.), *The Psychophysics of Speech Perception*. M. Nijhoff, The Netherlands, 80-90.

Espinoza-Varas, B. (1986) "Involvement of the critical band in identification, perceived distance, and discrimination of vowels." In: M.E.H. Schouten (Ed.), *The Psychophysics of Speech Perception*. M. Nijhoff, The Netherlands, 306-313.

Espinoza-Varas, B., and C.S. Watson (1986) "Temporal processing abilities of hearing impaired listeners." *J. Acoust. Soc. Am.*, 80, S12.

Espinoza-Varas, B., C.S. Watson, and D.A. Geddes (1986) "Correlations between auditory capabilities and measures of phoneme perception." *J. Acoust. Soc. Am.*, 79, S23.

Espinoza-Varas, B., and C.S. Watson, (1987) "Temporal discrimination for single components of nonspeech auditory patterns." *J. Acoust. Soc. Am.*, 80, 1685-1694.

Fallon, S.M. and D.E. Robinson (1987) "The effects of decorrelation on the discriminability of noise samples." *Journal of the Acoustical Society*, 81, S33.

Foyle, D.C. and C.S. Watson (1984) "Stimulus-based versus performance-based measurement of auditory backward recognition masking." *Perception & Psychophysics*, *36*, 515-522.

Gilkey, R.H., D.E. Robinson, and T.E. Hanna (1985) "Effects of masker waveform on signal-to-masker phase relations on diotic and dichotic masking by reproducible noise," *J. Acoust. Soc. Am.*, *78*, 1207-1219.

Gilkey, R.H. and D.E. Robinson (1986) "Models of auditory masking: A molecular psychophysical approach," *J. Acoust. Soc. Am.*, *79*, 1499-1510.

Hanna, T.E. and D.E. Robinson (1985) "Phase effects for a sine wave masked by reproducible noise," *J. Acoust. Soc. Am.*, 77, 1129-1140.

Jamieson, D.G., C.W. Ponton, and B. Espinoza-Varas (1985) "Reduction of formant bandwidth improves vowel identification with sensorineural impairment." *J. Acoust. Soc. Am.*, 77, S8.

Jamieson, D.G., E. Slawinska, M.F. Cheesman, and B. Espinoza-Varas (1984) "Timing perturbations with complex auditory stimuli." *Ann. N.Y. Acad. Sci.*, 423, 96-102.

Johnson, D.M., C.S. Watson, and W.J. Kelly (1984) "Performance differences among the intervals in forced- choice tasks." *Perception & Psychophysics, 35,* 553-557.

Johnson, D.M., C.S. Watson and J.K. Jensen (1987). "Individual differences in auditory capabilities. I." *J. Acoust. Soc. Am.*, 81, 427-438.

Kelly, W.J. and C.S. Watson. (1985) "Stimulus-based limitations on the discrimination between different temporal orders of tones." *J. Acoust. Soc. Am., 79,* 1934-1938.

Kewley-Port,D. (1986) "Converging approaches toward establishing invariant acoustic correlates of stop consonants." In Perkell, J. and D. Klatt (Eds.) *Invariance and Variability in Speech Processes,* Lawerence Erlbaum: New Jersey, 193-197.

Leek, M.R. and C.S. Watson (1984) "Learning to detect auditory pattern components." *J. Acoust. Soc. Am., 76,* 1037-1044.

Robinson, D.E. and R.D. Sorkin (1985) "A contingent criterion model of computer assisted detection," In *Trends in Ergonomics/Human Factors,* Vol. II, R. Eberts and C.B. Eberts (Eds.), North-Holland: Amsterdam, 75-82.

Sorkin, R.D. and D.E. Robinson (1985) "Alerted Monitors: Human Operators Aided by Automated Detectors," DOT/OST/P-34/85 /021, U.S. Dept. of Transportation. (Final Report: Contract # DTRS-5683-C-0047, not refereed).

Spiegel, M.F. and C.S. Watson (1984) "Performance on frequency-discrimination tasks by musicians and nonmusicians." *J. Acoust. Soc. Am., 76,* 1690-1695.

Yost, W.A. and C.S. Watson (1987) *Auditory Processing of Complex Sounds* , Lawerence Erlbaum: New Jersey.

Watson, C.S. and D.C. Foyle (1985) "Central factors in the discrimination and identification of complex sounds." *J. Acoust. Soc. Am., 78,* 375-380.

Watson, C.S., T.R. Dolan, J.E. Hind, I.J. Hirsh, D. McFadden, and W. A. Yost (1985) "Introduction: Workshop on basic research in hearing." *J. Acoust. Soc. Am., 78,* 295-298.

Watson, C.S., (1987) "Uncertainty, informational masking, and the capacity of immediate auditory memory." In W.A. Yost and C.S. Watson (Eds.) *Auditory Processing of Complex Sounds* , Lawerence Erlbaum: New Jersey, 267-277.

Manuscripts In Progress

Berg, B.G. and N.J. Castellan, Jr. "Adaptive rating scales in sequential decisions." *Behavior Research Methods, Instruments, and Computers*, (in press).

Espinoza-Varas, B. (1986) "Perception of critical bandwidth changes in format frequency," *J. Acoust. Soc. Am.* (submitted)

Espinoza-Varas, B. and C.S. Watson (1987) "Perception of complex auditory patterns by humans." In: S.H. Hulse and R.J. Dooling (Eds.) *The Comparative Psychology of Complex Acoustic Percepttion* (in press).

Kewley-Port, D., C.S. Watson and D.C. Foyle. "Auditory temporal acuity in relation to category boundaries; speech and non-speech stimuli," *J. Acoust. Soc. Am.* (in press).

Leek, M. and C.S. Watson. "Learning a vocabulary of nonspeech patterns." (submitted to Perception and Psychophysics).

Sorkin, R. D., Robinson, D. E., and Berg, B. G. (1987). "A detection theory method for the analysis of visual and auditory displays." *Proceedings of the Human Factors Society* ,in press.

Watson, C. S., Foyle, D. C., & Kidd, G. R. "Limited processing capacity for auditory pattern discrimination." (in preparation).

Oral Presentations

Berg, B.G. (1986) "Adaptive rating scales in sequential decisions." presented at the Society for Computers in Psychology meeting, New Orleans, LA, November, 1986

Espinoza-Varas, B. and Watson, C.S. (1985) "Basic auditory capabilities and resolving power for phonemes" *J. Acoust. Soc. Am.*, *78*, S47, presented at the 110th meeting of the Acoustical Society of America, Nashville, TN, November, 1985.

Espinoza-Varas, B., C.S. Watson, and D.A. Geddes (1986) "Correlations between auditory cpabilities and measures of phoneme perception." *J. Acoust. Soc. Am.*, *79*, S23, presented at the 111th meeting of The Acoustical Society of America, Cleveland, OH, May, 1986.

Espinoza-Varas, B., and C.S. Watson, (1987) "Informational limits in speech processing by normal and hearing impaired listeners." *J. Acoust. Soc. Am.*, *82*, R12, presented at the 114th meeting of The Acoustical Society of America, Miami, FL, 1987.

Fallon, S.M. and D.E. Robinson (1985). "The effects of changes in correlation on the discriminability of noise samples," *J. Acoust. Soc. Am.*, *78*, s46(A), presented at the 110th meeting of the Acoustical Socniety of America, Nashville, TN, November, 1985.

Jamieson, D.G., C.W. Ponton, & B. Espinoza-Varas (1985) "Reduction of formant bandwidth improves vowel identification with sensorineural impairment." *J. Acoust. Soc. Am.*, *77*, S8, presented at the 109th meeting of the Acoustical Society of America, Austin, TX, April, 1985.

Kewley-Port, D. (1985) Invited lecture on speech synthesis presented in "The Seminar on Basic Hardware and Software Issues in the Speech and Hearing Sciences Laboratory." *Asha 27*, 146, presented at the 1985 Annual Convention of the American Speech-Language-Hearing Association, Washington, D.C., November, 1985.

Kewley-Port, D. and C.S. Watson (1985) "Stimulus uncertainty in the discrimination of category boundaries." *Asha 27*, 186, presented at the 1985 Annual Convention of the American Speech-Language-Hearing Association, Washington, D.C., November, 1985.

Kewley-Port, D., C.S. Watson, and M. Czerwinski (1986) "Information Masking in Vowel Sequences." *J. Acoust. Soc. Am.*, *79*, S66, presented at the 111th meeting of The Acoustical Society of America, Cleveland, OH, May, 1986.

Kidd, G.R. (1985) "The effect of sentence timing on the perception of word-initial stop consonants." *J. Acoust. Soc. Am.*, *78*, S21, presented at the 110th meeting of the Acoustical Society of America, Nashville, TN, November, 1985.

Kidd, G.R. and C.S. Watson (1987) "Perception of multidimensional complex sounds." *J. Acoust. Soc. Am.*, *81*, S33, presented at the 113th meeting of the Acoustical Society of America, Indianapolis, IN, May, 1987.

Robinson, D.E. and B.G. Berg (1986). "A partitioned variance model of computer assisted detection," Ergonomics/Human Factors Conference, Purdue University, West Lafayette, IN.

Robinson, D.E. and R.D. Sorkin (1985). "A contingent criterion model of computer assisted detection," Ergonomics/Human Factors Conference, Purdue University, West Lafayette, IN.

Watson, C.S., Kewley-Port, D., & Foyle, D.C. (1985). "Temporal acuity for speech and nonspeech sounds: The role of stimulus uncertainty." *J. Acoust. Soc. Am., 78,* S27, presented at the 109th meeting of the Acoustical Society of America, Austin, TX, April, 1985.

Watson, C.S. and D.C. Foyle (1985) "Capacity limitations for discrimination of isochronous and anisochronous tonal patterns," presented at the 110th meeting of the Acoustical Society of America, Nashville, TN, November, 1985.

Watson, C.S. and D. Kewley-Port (1985) "Stimulus uncertainty in complex auditory pattern discrimination." *Asha 27,* 186, presented at the 1985 Annual Convention of the American Speech-Language-Hearing Association, Washington, D.C., November, 1985.

Watson, C.S.,and B. Espinoza-Varas, (1987) "Rate and number as limiting factors in the perception of sequences of syllables." *J. Acoust. Soc. Am., 82,* R11, presented at the 114th meeting of The Acoustical Society of America, Miami, FL, 1987.

Watson, C.S. and G. Kidd (1987) "Proportional target-tone duraton as a factor in the discriminability of tonal patterns." *J. Acoust. Soc. Am., 82,* S40, presented at the 114th meeting of the Acoustical Society of America, Miami, FL, November, 1987.

Personnel

Name	*Position Title*	*Department*
Charles S. Watson, Ph.D.	Professor & Chair	Speech & Hearing Sci.
Donald E. Robinson, Ph.D.	Professor	Psychology
Diane Kewley-Port, Ph.D.	Asst. Prof.	Speech & Hearing Sci.
Gary R. Kidd, Ph.D.	Assistant Scientist	Speech & Hearing Sci.
David C. Foyle, Ph.D.*	Assistant Scientist	Speech & Hearing Sci.
Blas Espinoza-Varas, Ph.D.	Visiting Research Associate	Speech & Hearing Sci.

* current address: Naval Weapons Research, China Lake, CA.

Advanced Degrees Awarded

David A. Geddes, M.A.; May, 1986; *Estimation of Hearing Loss in Elementary School Children using Acoustic-Reflex Thresholds.*

Bruce G. Berg, Ph.D.; August, 1987; *Internal Noise in Auditory Detection Tasks.*

Christine Kapke, M.A.; August, 1987

END
DATE
FILMED
4-88
DTIC