WADD TECHNICAL REPORT 60-249

RESEARCH ON PSYCHOLOGICAL PARAMETERS OF SOUND

Robert W. Peters

Mississippi Southern College

FEBRUARY 1960

Contract No. AF 33(616)-3644
Project No. 7231
Task No. 71786

AEROSPACE MEDICAL LABORATORY
WRIGHT AIR DEVELOPMENT DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

1,000 — July 1960 — 41-1201

FOREWORD

This study was initiated by the Aerospace Medical Laboratory, Wright Air Development Division, Wright-Patterson Air Force Base, Ohio, under Air Force contract No. AF 33(616)-3644, in support of Project No. 7231, Task No. 71786, "Research on Psychological Parameters of Sound." Jack E. Steele, Maj, USAF (MC) of the Bioacoustics Branch served as contract monitor.

Basic research, compilation, and analysis of data were performed at Mississippi Southern College, Acoustic Laboratory, by Dr. Robert W. Peters, Project Director. Dr. H. Paul Kelley, Coordinator of Measurement and Statistical Services, University of Texas, and Ens. Tony Morton, USN, of the U. S. Naval School of Aviation Medicine, Pensacola, Florida, gave advice on multidimensional scaling procedures. Dr. M. S. Edwards and Mr. L. M. Odell of the Digital Computation Branch, Aeronautical Research Laboratory, WADD, programmed the multidimensional scaling procedure for computer analysis and processed the data for the two large multidimensional studies. Major Steele was particularly helpful in his encouragement and served as liaison between the author and the Digital Computation Branch. These contributions are gratefully acknowledged.

ABSTRACT

Experimental approaches to the determination of psychological parameters for complex sounds are discussed and relevant studies are summarized. The results of a series of experiments which were conducted treating the problem of parameters for complex sounds are included in this report. The studies involved language of auditory experience, scaling of auditory stimulus sets on equal-interval scales, paired-comparisons scaling, scaling by direct magnitude estimations, observer generated scales, multidimensional scaling of groups of sounds, and changes in auditory perception under constant stimulation. The results of the various studies indicated that the most promising approach for the study of parameters of complex sounds was the multidimensional scaling model. In a pure tone study this scaling model isolated the expected dimensions, pitch and loudness. In the scaling of complex sounds, for several stimulus sets evaluated, between three and six dimensions appeared to emerge.

PUBLICATION REVIEW

This report has been reviewed.

FOR THE COMMANDER:

Walter F. Grether

WALTER F. GRETHER
Director of Operations
Aerospace Medical Laboratory

TABLE OF CONTENTS

EXPERIMENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The basic concepts of this program of research were that dimensions of auditory experience account for man's ability to differentiate one auditory stimulus from another; that these dimensions relate in a meaningful manner to the capacities of the perceiving organism, to the properties of the stimulus, to previous auditory experience, to the presently occurring set of stimuli, and that these dimensions are time variant and are not independent of the total functioning of the organism.

PURPOSE

The purpose of the program of research summarized in this report was to investigate and evaluate experimental approaches to the determination of psychological dimensions of steady-state complex sounds.

PROBLEM

The understanding of man's reaction to sound requires knowledge of the properties of the auditory stimulus to which he may react. It is first necessary to determine the dimensions of auditory experience in order to relate these dimensions to the physical properties of sound. The study of psychological dimensions of complex sounds is a twofold problem. There is first the task of isolating and defining the dimensions, and second, the task of measuring the dimensions. A fortunate experimental approach, or model, would be one which both defines and measures the dimensions. The test of the utility of such a model would be the extent to which it successfully allowed prediction of responses to sounds. The construction of predictive models which relate to observables in nature or in behavior is the goal of all science.

FACTORS IN THE STUDY OF PSYCHOLOGICAL DIMENSIONS OF SOUND

Although in actual practice it is not possible to separate these factors, the study of psychological dimensions of sound may be considered with reference to the measurement model, the observables in human behavior, and the auditory stimulus. In order to discover laws which govern auditory perception or make predictions about it, it is necessary to consider each of these three factors as they relate to each other. This might involve noting systematic variations in the observables of behavior as the result of the altering of the auditory stimuli in a known manner. It might involve studying the transformations which the auditory stimulus may undergo and still result in constancy of auditory experience. Another possibility is that of holding stimulus and observer constant and deriving a model which predicts changes in auditory experience in time.

MEASUREMENT MODEL

The model should involve the assignment of numbers to the property of an auditory event in a meaningful fashion. It is assumed that there is an isomorphic relationship between the numbers which are assigned and the gradations of the property. For example, if we order sounds on a psychological continuum with regard to the property of pitch, we are assuming that this is a property which may be related to a straight line geometric function. We might further assume that the distances between the numbers assigned to the property are some form of equalization units. Finally we might relate points on the continuum as distances from a common orgin. If we achieve these three objectives, we have utilized the features of the number system to measure

---

Manuscript submitted by the author December 1959 for publication as a WADD Technical Report.

the property of pitch. We then have a predictive measurement model regarding pitch because the relations between the numbers are indicative of the relations between auditory events with regard to pitch. In actual scaling we may be satisfied with models which only define order, or only define order and distance between numbers, with regard to the property of pitch, and do not define orgin.

OBSERVABLES IN HUMAN BEHAVIOR

The determination and measurement of observables in human behavior which relate to auditory perception is a basic task in the study of psychological dimensions of sound. Perceptual scales may be based on direct judgments by an observer or inferred from his responses. The judgments an individual is able to make about auditory stimuli is one kind of measurable behavior. Examples of this kind of behavior are judgments of the presence or absence of an auditory experience, judgments of order, equality, difference, and magnitude about stimuli with respect to a specific property of the stimuli. An observer's ability to identify stimuli, his ability to localize the source of a sound are other measurable aspects of behavior. The measurement of reaction time and the performance of tasks in response to auditory stimuli are also observables in behavior which may be related to auditory perception. Non-verbal reponses such as changes in skin resistance, blood pressure, and nerve potentials are physiological factors which may be considered in the measurement of observables which relate to auditory stimuli.

THE AUDITORY STIMULUS

Sound may be completely defined in terms of wave form or spectrum. Wave form considers amplitude and time; spectrum considers frequency, amplitude, and phase. In the measurement of sound, for the purpose of relating psychological to physical parameters of sound, it is necessary to use compromise methods which are not pure wave form or spectrum analyses. This is partly due to the limitations of sound measuring devices and partly due to the influence of hypotheses as to how the human auditory mechanism responds to sound. Frequency, amplitude, duration, phase, frequency and amplitude change in time, rate of frequency and amplitude change, band width, spectrum, and various combinations of these parameters are physical measures of sound which are frequently used in studies which relate physical parameters to auditory experience.

APPROACHES TO THE STUDY OF PSYCHOLOGICAL DIMENSIONS OF SOUND

The study of psychological dimensions of sound may be approached with emphasis on either the measurement model, the human observer, or the auditory stimulus. In this section, studies relevant to the problem of dimensions of auditory experience are discussed from the point of reference of either the model, the observer, or the stimulus. This grouping of studies is largely arbitrary for convenience of discussion since many of the studies could be treated equally well from any of the three points of reference. Studies which relate to psychological scaling, information processing, and signal detection are discussed relative to the model since the criteria as to the nature of the observer's responses are established by the measurement model. Auditory abilities, semantic descriptions of sounds, and changes in auditory experience in time are related to the observer. Studies of synthetic speech and music are discussed with regard to the auditory stimulus.

THE MEASUREMENT MODEL

Studies on auditory perception where the emphasis was on the measurement model are treated in this section.

Unidimensional Scaling

The traditional approach in the study of psychological attributes of sound has been unidimensional in nature. An attribute has been defined as an experience which could be ordered on a magnitude scale where the numbers of the scale have a meaningful relationship to the experience of the individual. In the psychophysics of sound, the task for the observer in the development of these attribute scales has been either direct magnitude judgments or subjective estimates of scale values. Fractionation, equisection, and direct magnitude estimations are variations of the direct judgment techniques. In the fractionation task it is assumed that an observer is capable of reporting the magnitude of a ratio between two stimuli. This may involve asking the observer to indicate a prescribed ratio, as for example, one-half or one-fourth magnitude of an attribute, or may require the observer to state the ratio which exists between two stimuli. In the equisection methods the observer is asked to either bisect, trisect, or in some other manner, divide the distance between stimuli into equal sense distances. In direct magnitude estimation, the observer is required to provide a numerical value for a stimulus with respect to an attribute of the stimulus. A standard stimulus may or may not be provided as a reference for the observer in making his estimate of magnitude. In the subjective estimate method, an observer is required to place a stimulus at some point on an equal-interval scale with respect to a particular attribute. A number of attributes for both pure tones and complex sounds have been scaled by using one of these techniques or a combination of these techniques. Loudness and pitch are two attributes which have been extensively studied, especially the loudness and pitch of pure tones. The two physical parameters of pure tones, amplitude and frequency, can be varied independently and because of this, the attributes of pure tones have been considerably easier to study than have complex sounds where it has not been possible to independently vary each physical parameter.

Loudness. Loudness is defined as an attribute of auditory experience where sounds may be ordered on a scale which extends from soft to loud. It is an attribute of particular interest in this discussion because of the extensive work which has been done in the attempts to scale loudness, especially the scaling of complex sounds with respect to loudness. There has also been a very real interest in loudness because of the problems of noise and the need to devise a measuring device which would relate physical measures of the amplitude of sounds to the subjective experience of loudness. The first serious attempt to establish a scale of loudness was made by Fletcher and Munson (18). They inferred half-loudness values by measuring the intensities of sounds heard by one and by both ears and established equal-loudness contours by determining the intensity levels at various frequencies necessary to sound as loud as a 1000 cps tone which was used as a standard at various intensity levels. The resulting equal-loudness contours was called a phon scale of loudness. Stevens (66) proposed the name sone for the loudness scale and defined a sone as the loudness of a 1000 cps tone 40 db above threshold. The adequacy of the phon or sone scale to predict loudness is diminished for sounds of increased complexity and considerable variability is found to occur between the direct judgments of loudness by observers and the predicted loudness from either of the two scales. This discrepancy has resulted in numerous re-evaluations of loudness scaling. Pollack (50) was concerned with the problem of whether or not loudness could be established as a separate and distinct aspect of noise. He required observers to match a narrow band of noise to a standard noise signal in terms of loudness, volume, density, and force. He found that experienced observers differ considerably in their judgments of these attributes. Garner (19) studied loudness scaling with reference to the techniques of fractionation and bisection and questioned whether or not comparable scales could be obtained using these two methods. Garner (22) subsequently proposed what he termed a true ratio scale of loudness. This scale was based on two loudness functions, equisection and fractionation judgments. The advantages of using discriminability criterion in the development of a loudness scale was discussed by Garner (23). He advocated a loudness scale based on discriminability or judgmental dispersion measures. He indicated that a scale based on these measures would have less observer variability, be less influenced by experimental conditions, and less affected by total stimulus set than would a scale based on direct judgments by observers. An equal discriminability scale based on dispersions of absolute judgments has been shown to relate in a similar manner to intensity as does a scale which is based on cumulative difference limen for intensity (20). Stevens, (67) in a comprehensive review of the data on the measurement of loudness, summarized the techniques of scaling and the results of numerous studies. He proposed a formula for the calculation of the loudness of complex sounds based on a consideration of over-all sound level and the spectrum of the sound. The notion of the sone scale determined through direct magnitude estimation, equisection, and ratio determination is retained as the method to relate loudness to physical measures of amplitude of sound. He suggests that magnitude and fractionation methods may be used to generate a ratio scale and that these results may be used to predict the results which would be obtained by the bisection method. However, he noted that bisection alone would not be suitable to establish a ratio scale. Stevens (68) also presented a formula for the calculation of the loudness of complex noise based on the loudness in sones of bands of noise of the total noise. Garner (24) presented a computational procedure based on a loudness scale determined by equisection for predicting the loudness of complex sounds. The two basic assumptions were that the loudness of components add in proportion to the sum of their squares and that a constant amount of loudness should be subtracted from the total loudness due to the effects of masking. It may be noted from this discussion on loudness that the development of a meaningful scale of loudness, one which successfully predicts loudness from physical measures, becomes increasingly difficult as one deals with those complex sounds which are more realistic in terms of the sounds of the environment than are pure tones.

Pitch. Pitch is defined as an attribute of auditory experience in terms of which sounds may be ordered on a scale which extends from low to high. Pitch correlates primarily with the frequency of a sound and also relates to the amplitude and wave form of a sound. Since pitch relates primarily to frequency, a scale relating pitch to frequency serves as a good approximation of the pitch function. Stevens, Volkmann, and Newman (70) devised a subjective scale for the measurement of pitch based on judgments of half-values of pitch at various frequencies. This pitch scale showed good agreement with integration of difference limen for pitch. This original scale was modified by Stevens and Volkmann (69) by the use of fractionation and bisection procedures. For pure tone stimuli pitch seems to be a reasonably stable attribute and the mel scale of pitch which resulted from the two experiments cited above is an excellent example of good psychophysical scaling. However, there appears to be considerable variation between individuals in pitch judgments of pure tones (47) and the task of judging pitch becomes extremely difficult for complex sounds (33,41). Since most of the attempts to scale complex sounds for pitch have been unsuccessful, investigators have suggested that it may be multidimensional in nature and have proposed two dimensions to account for pitch perception (11,42).

<u>Volume</u>, <u>Brightness</u>, <u>and</u> <u>Density</u>. In addition to loudness and pitch, a number of other psychological attributes have been proposed for both pure tones and complex sounds. Some of the terms applied to describe attributes are timbre, volume, density, brightness, fullness, roughness, rhythm, vocality, and harshness. With respect to pure tones, volume, brightness, and density have been studied. Halverson (29) studied difference limen for volume as a function of amplitude of sound and Rich (55) measured difference limen for volume as a function of frequency. The criterion of these early investigators as to whether or not volume was an attribute apart from loudness was that if the difference limen for the two were different, then loudness and volume could be considered as separate attributes. Gundlach (26) concluded that there was not sufficient evidence to indicate that volume was a separate attribute and in a later study, concerned with the effect of phase upon perceived attributes of sound, found that brightness and volume were equivocal attributes (27). Banister (3) and also Zoll (83) in discussions of tonal attributes, questioned the existence of the attribute volume. Stevens (63,64,65) required that his observers judge equality rather than difference limens and plotted equal-volume, equal-density, and equal-brightness contours for various tones using a 1000 cps tone as a standard. Boring and Stevens (4) later concluded that since brightness appeared to vary directly with intensity and frequency of tones in the same manner as density, the two were not separate attributes and density was the appropriate term. Thomas (72) determined equal-volume contours for a wide range of intensities and frequencies and interpreted his results as indicating that volume is an independent attribute and not a combination of loudness and pitch judgments. Thomas (73) also determined equal-loudness and equal-volume contours for bands of white noise. The two sets of contours appeared to be different and this difference was cited as support for a separate volumic attribute. Timbre has been discussed as another attribute of complex sounds (17,44) by a number of investigators, particularily those concerned with music and speech. It appears to be an ill-defined attribute much in the same category as the term quality as applied to voice. In clinical speech, nasality (59) and harshness (60) have been scaled to the extent that there is good agreement among observers with respect to their ratings of these attributes. Threshold values for auditory flutter (51) and for trill (45) have been determined and for flutter have been related to annoyance factors (10). Lichte (40) concluded that, for complex tones, there were probably three attributes in addition to pitch and loudness. He identified these attributes as brightness, roughness, and fullness.

Multidimensional Scaling

Multidimensional scaling differs from traditional scaling methods in that judgments of similarity between stimuli may be utilized instead of judgments on a given continuum and, in that the dimensions and scale values are determined from the data instead of the dimensions being specified by the experimenter. This approach has advantage in stimulus domains where the stimuli dimensions are unknown, as in the area of complex sounds. In traditional methods of scaling, as was noted in the previous section, the experimenter determines the dimensions on which the observer is to make his judgments. The observer needs to know what is meant by such terms as loudness, brightness, volume, and density. In some cases, a dimension specified by the experimenter may be in reality complex, and in other cases, the dimension may not be relevant. The inadequacy of the traditional approach is apparent. Although multidimensional scaling has not been applied in the area of auditory perception, it has been applied in various other fields to determine the nature and number of psychological dimensions. In the area of color perception, Richardson (56) and Messick (43) found good agreement between the results of multidimensional scaling and the Munsell color system (48). Messick also used the multidimensional method of successive intervals to evaluate attitudes toward war, capital punishment, and treatment of criminals. He found that the three attitudes could be represented in two dimensions, a war and punishment dimension. Attneave (2) differed size and shape as well as color of objects and found that the scaling method revealed the appropriate number of dimensions. Klingberg (38) studied the mutual friendliness of seven great powers before World War II and noted that a three-dimensional system could account for mutual international distances. The multidimensional model and the studies discussed above are based on psychological distance between stimuli which is determined from observers' judgments of similarity of stimuli. Shepard (58) described a stochastic model which may be used to represent stimuli or responses in psychological space on the basis of overt errors made during learning rather than depending upon direct judgments of similarity between stimuli. Multidimensional scaling would appear to hold promise for scaling in the area of auditory perception, especially for the perception of complex sounds where independent control of physical parameters is difficult.

Information Theory

Information theory is another model which may be of value in studying the dimensions of auditory perception. Specifically, the concept of channel capacity, the measure of covariance between the input and output of a system, may be utilized to infer the psychological dimensions of stimuli sets. In information theory, the unit of measure is a bit and a bit is defined as the amount of information needed to make a decision between two equally likely alternatives. We may regard the observer as the communication channel and, if we require him to make absolute judgments about stimuli, measure the extent to which his responses match the input stimuli. If channel capacity is found to be a fairly stable measure with respect to dimensionality of stimuli, we might be able to infer from the observer's responses information as to the dimensions in domains where the dimensions are unknown. Channel capacity for unidimensional stimuli seems to be stable. It has been generally found that about 2.5 bits of information are transmitted for unidimensional stimuli. In terms of absolute judgments, this is about six stimuli identified correctly. Pollack (52) measured the amount of information transmitted by observers who made absolute judgments of pitch. He found that when there were five or more tones there were a number of confusions. Garner (21) obtained similar results for absolute judgments of loudness. Pollack (53) had observers judge both loudness and pitch of pure tones and found that 3.1 bits of information were transmitted. This is short of the amount of information which would be expected if the bits for pitch and loudness were directly additive. In another study of multidimensional stimuli, Klemmer and Frick (37) required observers to make absolute judgments of the position of a dot in a square. The channel capacity for this task was 4.6 bits which may be compared to the channel capacity of approximately 3.25 for a similar unidimensional task of judging positions on a linear scale(28). Miller and Nicely (46) analyzed the perceptual confusions among 16 English consonants within an information theory model. The consonants were heard by listeners under various conditions of noise and frequency distortions. The listeners' responses were tabulated in confusion matrices and the patterns of confusions were analyzed to reveal dimensions which distinguished the different consonants. This was done by determining the sub-matrices necessary to account for the information transmitted by the confusion matrices. Five dimensions appeared adequate to explain the confusions among the 16 consonants. These dimensions were judged to be voicing, nasality, affrication, duration, and place of articulation. The basic notion of this approach, that dimensions of auditory experience can be inferred from the nature of observers' responses, would seem to be promising and have application in other auditory areas other than speech.

Information Processing and Signal Detection Models

In the preceding section a model, information theory, was described which could be used to predict auditory behavior. This model was based on the notion of binary decisions. In a sense, much of the work in the area of audition, particularly the work which related physiological and neurological data on sensory mechanisms of hearing to auditory perception, has been concerned with model building and were called theories of hearing. A number of theoretical models to explain certain aspects of auditory perception have been developed. The models attempt to describe the way the auditory mechanism should operate and then experiments are performed to test the adequacy of the model. Two models which have been recently developed merit discussion as possible approaches to the study of psychological dimensions of auditory experience. The first of these models deals with the storage and filter action which is performed on incoming auditory stimuli by the observer (5). The model is based on the concept that since the perceptual system has a limited capacity, a selective operation is performed on all inputs to the system. This selective operation takes the form of selecting all inputs having some characteristic in common and depends upon such factors as the magnitude of the stimulus, the time of arrival of the stimulus, the absence of recent inputs to the system, and the information which has been stored from previous auditory events. Broadbent cites as evidence in support of the model that simultaneous stimuli do interfere with each other, the first of two nearly simultaneous stimuli has an advantage in passing through the perceptual channel while the second is at a disadvantage, intense stimuli are more likely to be responded to than are weaker stimuli, and that a channel which has been previously busy is less likely to pass a stimuli than a channel which has been inactive. There are several powerful concepts in this approach which are useful in the study of auditory theory. These concepts have been admirably expanded by Broadbent (6) and are: that the organism has a limited capacity to respond to stimuli, perception is in part determined by the information stored from previous auditory events, and that the study of the effect of one perception upon another is a possible method to understand the functioning of the sensory process. If an adequate model is to be developed to predict auditory perception, these variables need to be considered as part of the model. The second model is concerned with the problem of sensory detection (71). This is a probability model in which the responses of an observer may be compared to what would be expected of a perfect detection device responding to the same signals. Two concepts of this theory are called the detection index and the efficiency index. The detection index is the specification of physical parameters in terms of the ratio of signal energy over noise power per unit band width necessary for the ideal detection device to match the performance of the receiver under study. Efficiency is defined as the ratio of the signal energy required by the ideal receiver as compared to the signal energy required by the receiver under study when the performance of the two is the same. In the application of this model to the detection of auditory signals, Veniar (77,78,79) studied the detection of a signal in noise as a function of signal ensemble size and ensemble frequency range. She found that observer performance decreased with increasing ensemble size but not with increasing ensemble range, indicating that to describe the observer as a narrow-band scanning device was not adequate to explain detection phenomena. Egan (16) applied the model to the study of voice communication channels. He studied the probabilities associated with the responses to voice messages in terms of identification of the message and the decision that the receiver made about his response. He applied this information to the prediction of the reception of messages in noise and developed predictive techniques whereby estimates could be made about message reception. It appears that increased use of probability models is warranted in the study of auditory perception in that as we are better able to account for the variability in observer responses, our ability to predict responses to auditory stimuli is increased. By way of comparison of this kind of model with the traditional scaling models, in the traditional model the systematic variation in responses is attributed to the stimuli rather than to the observer.

THE OBSERVER

In this section auditory abilities, semantic descriptions of sounds, and measurable changes in pitch and loudness as a function of time of stimulation are discussed.

Auditory Abilities

One approach to the study of psychological dimensions of sound, where the emphasis is on the observer, is that of isolating primary auditory abilities. A major argument for this approach is that the generally noted attributes, pitch, loudness, volume, and density, under close examination do not appear to be unitary dimensions. Harris (32) found that loudness discrimination and the ability to resist masking were not pure auditory traits, but rather, appeared to be factorially complex. Pitch, also, appeared to be a binary trait rather than unitary in nature. Harris proposes to sample small segments of the auditory area with factorial studies and, by using tag tests, eventually isolate the primary auditory abilities. The study of auditory abilities was pioneered by Seashore (57) in his study of musical abilities and he developed a series of tests in this area. These tests were assumed to measure basic musical abilities. Factorial analyses have been applied to auditory areas other than music. Karlin (36) subjected the results of 33 tests covering such items as pitch, loudness, auditory analysis, age, and intelligence for 200 subjects to multiple factor analysis. The analysis indicated that eight factors were necessary to account for the correlations between the tests. He identified seven of these factors as pitch-quality discrimination, loudness discrimination, auditory integral for perceptual mass, auditory memory span, auditory synthesis and analysis, speed of closure, auditory span formation, memory span, memory or incidental closure, and one unidentifiable residual factor. Hanley (30) employed factor analysis in the study of speech perception. Eight factors were noted and identified as verbal facility, detectability for tones, Seashore battery factor, voice memory, resistance to signal distortion, unpleasantness, and synthesis. Solomon (61) approached the problem of psychological dimensions of auditory experience by first obtaining descriptive adjectives for complex sounds. The sounds were passive sonar signals. Fifty descriptive adjectives which were obtained from sonar operators were rated on a seven-point scale where the adjective and its opposite defined the ends of the scale. The intercorrelations between scales were subjected to factor analysis. Seven factors were isolated and these were identified as continua of heavy-light, beautiful-ugly, clear-hazy, mild-intense, loose-tight, familiar-strange, and colorful-colorless. In a subsequent study, Solomon (62) attempted to relate these factors to physical measures of sound. The physical measures were octave band analyses of the sounds. Although some relations were found between the energy concentrations and psychological factors, subjective experience of heavy correlated with low frequencies and the experience of light with high frequencies, the relating of spectrum of the sounds to psychological dimensions was generally not fruitful. In a factor analysis of 12 physical measures of speech (1) four factors were isolated. These were identified as time, power, power variability, and pitch variability. Although a factor analysis approach to the study of psychological dimensions of auditory experience has merit, there are two major weaknesses in the method. The first is the considerable problem of factoring intercorrelations between variables to meaningful dimensions. The second is the problem of selection of tests or measures to be included for the factorial evaluation. Unless all meaningful measures are included, any subset of these measures will only reveal dimensions of the subset and not dimensions of the total area of auditory perception. Harris (32) has anticipated this problem by advocating the use of tag tests from previous analyses in subsequent analyses.

Loudness and Pitch Shifts in Time

Observers' responses with time as the variable and a constant stimulus are the concern in studies of loudness and pitch changes in time. A considerable amount of work has been done in measuring the changes in loudness which occur under conditions of constant auditory stimulation. A lesser amount of work has been concerned with pitch changes in time. The measurable change in loudness which occurs in time has been given various names depending upon the measure used by the investigator and the aspect of the phenomena where his interest was focused. Adaptation, perstimulatory fatigue, post-stimulatory fatigue, and temporary threshold shift are some of the terms applied to this measurable change in loudness with time. Wever and Truman (82) noted that auditory threshold declined under constant stimulation and that this decline in threshold appeared to level off after about two minutes. Other investigators have defined the course of change as a function of characteristics of the stimuli and the duration of the exposure (13,15,35). Although the time for the fatigue to reach an asymptote is apparently different for pure tones as compared to noise, three and one-half minutes as compared to seven minutes, and also differs with frequency and amplitude of the signal, the general function of a predicted change in loudness is well established. Predictive formulae relating fatiguing stimuli to recovery patterns for the auditory threshold have been determined (80). Changes have also been noted in pitch under constant stimulation. Pitch discrimination has been found to decrease as a function of constant stimulation (76,34,31). Pitch shifts, both upward and downward, have been noted under prolonged stimulation with pure tones (8,81). The fact that measurable changes may be noted in two attributes of auditory experience, both during stimulation and following stimulation, is important in the study of psychological dimensions of sound. Under constant stimulation, any systematic variation in response may be attributed to the perceiving organism and a predictive model developed to explain these variations. A second possibility is that under constant stimulation there may be changes in all of the meaningful dimensions of auditory experience as well as in pitch and loudness. The measurement of these changes might well define the dimensions of the auditory process.

THE AUDITORY STIMULI

In the preceding sections the model and the observer have been considered in approaches to the study of psychological dimensions of sound. In this section the emphasis is on the auditory stimuli. This includes the topics of synthetic speech and music and the transformation of auditory stimuli into other modes of sensory presentation.

Speech Synthesis

The basic problem in synthetic speech, music, and the transformation of auditory stimuli into other sensory modes is that of extracting the information carrying parameters from the speech, music, or auditory stimuli and converting these parameters into dimensions of the new information conveying medium. With regard to speech, we might transform it into visible units and then transform the visible units back into speech. This is precisely the approach used in the pattern playback machine (12). A source which generates 50 harmonic tones 120 cps apart is controlled by light reflected from paintings on a moving tape. The approach is that of converting a time-frequency analysis of speech back into speech. Other speech synthesis devices first code speech parameters and these parameters in turn are used to control the synthesizer (14,7). Extraction of pitch frequency, two formant amplitudes, two formant frequencies, central frequency, and average amplitude has proved adequate to code speech for synthesis. The work with synthetic speech has not yet resulted in an unique coordinate system of physical parameters which is adequate to reproduce all of the auditory elements of speech. One of the most interesting aspects of the work with synthetic speech has been the discovery that acoustic units will be identified differently depending upon the units which follow or precede them. This precludes establishing a one-tone relationship between phomenes and acoustic units and means that the statistical nature of the language will need to be considered in establishing simulated speech.

Music Synthesis

Music synthesis is comparable to speech synthesis with the exception that the task is probably simpler for music since music is not as complex as speech. Olson and Belar (49) described an instrument for the production of music. The physical parameters of frequency, intensity, duration, growth and decay, frequency glide, wave form, low-frequency modulation, and irregular deviation were given as the parameters which must be controlled by a synthesizer in order to produce music. The psychological parameters of pitch, loudness, portamento, timbre, and vibrato were related to the above controllable physical aspects. Clark (9) considered similar physical parameters in his proposed musical instrument and these parameters were growth and decay, frequency, and harmonics.

Transformation of Auditory Stimuli into other Sensory Modes

Attempts have been made to transform speech into a visible, readable form (54). The device, the visible speech apparatus, operates on the principle of spectrum analysis in time. Since speech as portrayed in this manner is readable, it appears that the essential parameters of speech are successfully transformed into a visual presentation. Guelke and Huyssen (25) report on a device for transforming speech into tactile sensations. Frequency was defined by positions on the fingers. The device performed a frequency analysis on the speech sounds and transmitted this information by means of vibrators to the fingers. It was found that vowels were relatively easy to identify and that consonants were not easily identified. It is difficult to generalize about the usefulness of the data on physical measures of sound. One of the complications is that the perception of language is not based on physical parameters alone but is also based on the characteristics of the language itself. Any study of the physical parameters necessary to reproduce speech, music, or speech through tactile sensory input with the notion that successful transformation defines the parameters of auditory perception is complicated by the characteristics of speech and music. The study of the physical parameters becomes confused with these characteristics. Language is highly redundant and for this reason a minimum of information in transformed parameters will carry the information of the language. With regard to the perception of language, it is not certain whether it is the absolute properties of the phomenes or the difference between phomenes which contributes to their identification. Perhaps the most important contribution that linguistics can advance to the area of auditory perception is by studies of the way in which languages are learned. In this way we might gain some knowledge as to how the human goes about structuring auditory events, particularily auditory events as complex as language.

EXPERIMENTS

The purpose of the program of research was to isolate and measure psychological dimensions of sound. With respect to the observer, two approaches were possible. The dimensions could be based on direct judgments by observers about sounds or inferred from the responses of observers to sounds. Both of these techniques were used. With respect to the measurement model, traditional scaling methods and recently developed scaling techniques were used. The program of research was ambitious in that it was hoped that a set of unique psychological dimensions for steady-state complex sounds could be isolated and scaled and the independence of the dimensions defended both in terms of the nature of the observers' responses and the mathematical model which was used to evaluate their responses.

STIMULI, OBSERVERS AND APPARATUS

Information relative to stimuli, observers and apparatus is noted in this section in order to avoid unnecessary repetition of these factors in the discussion of each of the experiments. Pure tones, multitones, white noise, speech, and music were the sounds used as stimuli. The multitones were prepared by recording back and forth between two magnetic tape machines, adding another pure tone to those which had already been recorded on each re-recording. Unless otherwise noted in the discussion of an experiment, the stimuli were heard by the observers at a comfortable listening level of approximately 80 db (re .0002 dyne/cm^2).

The observers were male and female college students. Although auditory thresholds were not obtained on these subjects for all of the experiments, in cases where thresholds were not determined, individuals who indicated that they had or suspected that they had depressed thresholds were not used as subjects.

The following equipment was available for use in programming or presenting stimulus materials: magnetic tape recorders, Ampex Models 350, 350-3, 600, and 601; console, eight channel Altec-Lansing Model 1560A; oscillator, General Radio Model 1302A; filter, Spencer-Kennedy Model 302; noise generator, H. H. Scott Model 811A; voltmeter, Hewlitt-Packard Model 400AB; attenuators, Hewlitt-Packard Model 350A; speakers, Altec-Lansing Model 604C and Ampex amplifier-speaker Model 620; electronic switch, Grason-Stadler Model 829S51; interval timer, Grason-Stadler Model 472; oscillograph, Dumont Model 401R; microphone, Altec-Lansing Model 21C; and headset circuits with matched earphones, either PDR-3 or PDR-8 receivers. The experiments were conducted in an isolated and sound-treated room. There were nine listening positions in the room.

The experiments are summarized under the main headings of language of auditory experience, absolute judgments on equal-interval scales, scaling by paired-comparisons, scaling by direct magnitude estimations, observer generated scales, differentiation of auditory patterns, inference of dimensions from channel capacity of observers, effect of complexity of stimuli on learning, multidimensional scaling, and changes in auditory perception under constant stimulation.

THE LANGUAGE OF AUDITORY EXPERIENCE

The descriptive terms which individuals apply to auditory experience may be revealing of the person's structuring of auditory events. There is no reason to believe that this structuring is not coherent, meaningful, and along efficient dimensions. The following studies were concerned with obtaining descriptive terms which might serve to define the extremes of equal-interval scales for subsequent studies.

Descriptive Terms Applied to Environmental Sounds

The task for the subjects was to prepare an exhaustive list of descriptive terms which characterized the sounds of their environment. The subjects were given instructions as to their task and were allowed several days to prepare the list of terms. Thirty students served as subjects.

The terms obtained from the subjects were tabulated according to frequency of occurrence. Some of the most frequently occurring words were whistle, buzz, bang, cry, rattle, and ring. Other frequently occurring words were those which described active events or objects, as for example, grinding and horn. The words which are usually listed in the literature on sound as describing attributes of auditory experience, high, low, soft, and loud, were listed by only about one-third of the subjects.

Descriptive Terms Applied to Complex Sounds

The subjects' task was to listen to a sound 10 times and give a different descriptive term, if possible, which would characterize the sound each time the sound was heard. There were 20 sounds and the duration of each sound was 15 seconds. Each sound was successively presented 10 times before another sound was presented. The 20 sounds were programmed on magnetic tape for presentation to the subjects and were heard by the subjects under a free-field condition through a loudspeaker. The 20 sounds were taken from a group of sounds which had been produced by an electric organ. Two subjects participated in this task.

Descriptive terms were tabulated according to frequency of occurrence and the words which occurred most frequently were loud, clear, high, low, bright, dull, hiss, buzz, rich, thin, and click. A few of the words which appeared were descriptive of actions and objects as were the words of the previous study.

Descriptive Terms of Similarity and Difference for Paired Sounds

Five multitones were programmed in all possible pair combinations and presented free field for listeners to make judgments as to how the two sounds of each pair were different and how they were similar. The subjects heard each pair of sounds for a total of three minutes during which time they made their judgments of similarity and difference. Each sound of the pair was of three seconds duration with a one second interval separating the two sounds.

Five pure tones of the frequencies of 500, 1000, 2000, 4000, and 8000 cps were programmed in the same manner in order that observers could make similar judgments regarding pure tone pairs as were made for the multitones. Twenty-seven subjects participated in the task concerned with multitones and 14 subjects in the pure tone part of the study.

The descriptive terms provided by the observers as to similarities and differences between the sounds were tabulated for both sets of stimuli. For the complex sounds, the terms which described similarities and differences between the sounds of each pair related to loudness, pitch, smoothness, and clearness. In addition to these four groupings of descriptive terms, words such as dense, diffuse, and full were also noted. Only two sets of terms, those which related to loudness and pitch, were obtained for the judgments relative to pure tones.

Associative Terms for Complex Sounds

The subject's task was to listen to 100 multitones and either name the thing of which the sound reminded him, indicate that it was a sound that reminded him of something but that he couldn't name the thing, or no, it was not a sound which could be related to anything. Each sound was of 10 seconds duration. The sounds were presented free field. Nine students participated as subjects.

The responses which were obtained by asking for associative descriptions were much the same as for the observers who were asked to describe the sounds of their environment. That is, the sounds were described in terms of events and objects, as for example, whistles, bells, drill, and grinding.

At the time these several studies were conducted, the results seemed highly disappointing. It was hoped that a set of terms could be obtained which, in the observers' own language, defined polar dimensions of auditory experience. While words such as volume, density, and brightness were not necessarily expected to appear, it was assumed that descriptive words such as roughness, harshness, and mellowness might appear. Reinspection of these data, which were collected early in the program of research, indicates factors of interest. The descriptive terms which were given by observers were relative to objects, movement of objects, and events which are characterized by particular sounds. The most frequently occurring descriptions, whistle, clang, ring, and cry, were sounds which are important to the individual. Solomon, (61) in his recent semantic approach to auditory perception, noted that several of the factors which he isolated were listener orientated: pleasant, safe, gentle, and happy. These results may indicate that individuals extract meaningful information from sounds as it is important for their well being (39). This is all by way of arguing that there is probably a hierarchy of dimensions in auditory perception and that this hierarchy and the structuring of auditory events are dependent upon past auditory experiences and those which are likely to occur.

ABSOLUTE JUDGMENTS ON EQUAL-INTERVAL SCALES

Several sets of auditory stimuli were scaled using this measurement model. The stimuli were bands of white noise, sounds produced by band instruments, and multitones. The stimuli were scaled with respect to arbitrarily selected attributes of sound.

Scaling of High-Pass and Low-Pass White Noise Samples

The task for the subjects was to rate 20 sounds with regard to eight characteristics. The ratings for each characteristic were made on a nine-point scale. The eight characteristics were descriptive terms which had appeared in the subjects' responses concerning environmental sounds and differences and similarities between pairs of sounds. These characteristics were clarity, fullness, brightness, loudness, highness, roaring, and buzzing. The 20 sounds were produced by selective high- and low-pass filtering of ASA white noise at the cut-off frequencies of 100, 200, 400, 500, 600, 700, 800, 900, and 1000 cps for both filtering conditions.

Before each of the eight rating tasks, the subjects were told the characteristic and how they were to judge the following group of sounds. For example, "You will rate the following group of sounds for clarity. Remember that <u>one</u> on the scale represents least clarity and <u>nine</u> represents greatest clarity." General instructions to the subjects as to the use of the rating scale preceded the specific instructions for each characteristic. The order of presentation of the 20 sounds was randomly varied for each of the eight successive rating tasks. Twenty-two college age males served as experimental subjects. The subjects heard the sounds through PDR-8 receivers.

Listener ratings for each of the 20 sounds with respect to the eight characteristics were tabulated and median (S) scale values and semi-interquartile (Q) values were determined according to the method outlined by Thurstone and Chave (75). Product moment (r) correlations were computed between the eight characteristics using scale values as criterion measures. These correlations are shown in the following table.

Table I.

Correlations Between Eight Selected Characteristics for 20 White Noise Stimuli

	Fullness	Brightness	Loudness	Highness	Roaring	Buzzing	Ringing
Clarity	.40	-.48	-.34	-.39	.58	-.58	-.25
Fullness		-.29	-.96	-.95	.73	-.91	-.95
Brightness			.94	.99	-.79	.91	.90
Loudness				.97	-.77	.89	.94
Highness					-.76	.92	.90
Roaring						-.75	-.74
Buzzing							.90

The Q values which indicated dispersion of judgments were as follows:

Clarity	Fullness	Brightness	Loudness	Highness	Roaring	Buzzing	Ringing
2.40	1.39	1.44	1.06	1.15	2.62	1.33	1.29

The high correlations betwen characteristics would seem to indicate that at most the judgments were on the basis of not more than two dimensions. This is probably a result of the limited variation which was present among the white noise samples. The Q values give some indication of agreement among subjects and it may be noted by the low Q values for loudness and pitch (highness) that the observers agreed fairly well on these characteristics.

Scaling of White Noise Samples of Differential Band Widths

The task for the subjects was to make repeated ratings of six white noise samples on a nine-point scale. The samples were rated with respect to six characteristics. These characteristics were relative loudness, density, volume, pitch, brightness, and fullness. Each sample was rated four times with respect to each characteristic. The six white noise samples differed in band width. The band widths were 100, 500, 1000, 2000, 4000, and 8000 cps with the center of the band at 4000 cps. A limiter amplifier, Altec-Lansing, Model A3220, was used to maintain relatively constant amplitudes for the different noise samples. The stimuli were heard through PDR-8 receivers. The subjects were 20 college-age males.

The criterion measure for the evaluation was the mean value of each subject's four ratings of each sample with respect to each characteristic. A triple-classification analysis of variance was employed to evaluate differences among characteristics and noise samples. The results indicated significant differences in mean values among characteristics and among band widths (Characteristics, F = 5.33, required for 5 % level, 2.19; band widths, F = 52.57, required for 5 % level, 2.19). The interaction between characteristics and band widths was also significant (F = 10.59, required for 5 % level, 1.60). This would indicate that mean values for band widths did not follow the same trend for each characteristic. The mean values for the six band widths of noise are shown in Table II. A difference of 0.63 between mean values, columns and rows, indicates a significant difference between means at the 5 % level.

Table II.

Mean Values of Observers' Ratings of Six White Noise Samples With Respect to Six Selected Characteristics of Sound

	Band Width (cps)					
Characteristic	100	500	1000	2000	4000	8000
Loudness	3.21	3.21	3.56	3.81	4.51	6.09
Density	4.08	3.98	4.06	4.40	5.56	6.85
Volume	3.90	3.86	3.83	4.30	5.00	6.95
Pitch	5.63	5.49	5.58	5.76	5.54	4.28
Brightness	5.26	5.03	5.16	5.46	5.21	5.15
Fullness	4.24	3.71	3.94	4.49	5.11	7.01

With increased band width of the white noise samples, ratings for all characteristics, except brightness, significantly changed. Pitch was significantly decreased for the band width of 8000 cps as compared to the other band widths. Mean rating values for the remaining characteristics increased with increased band width. The mean values for loudness, density, volume, and fullness follow similar patterns with increased band width.

Scaling of Band Instrument Sounds

Forty-four sound samples were recorded from 16 different band instruments. The subjects' task was to rate each of these sounds on a nine-point scale for volume, pitch, density, loudness, clarity, fullness, and brightness. Each sound was presented for a duration of five seconds with an interval of 10 seconds between the onset of each sound. Nineteen students served as judges. The sounds were presented free field.

Scale (S) and semi-interquartile range (Q) values were determined according to the method given by Thurstone and Chave (75). Product moment correlations (r) were determined between characteristics. These correlations are shown in Table III.

Table III.

Correlations Between Seven Selected Characteristics for 44 Band Instrument Sounds

	Pitch	Density	Loudness	Clarity	Fullness	Brightness
Volume	.40	-.41	.68	.24	-.29	.16
Pitch		-.91	.05	.14	-.83	.95
Density			.02	-.15	.86	-.93
Loudness				-.04	-.02	.07
Clarity					-.26	.16
Fullness						-.81

Semi-interquartile range (Q) values were:

Volume	Pitch	Density	Loudness	Clarity	Fullness	Brightness
2.52	1.18	2.48	1.82	2.22	2.36	1.44

With respect to the particular sounds used in this study, only three independent characteristics seemed to emerge, pitch, loudness, and clarity. Volume related directly to loudness and density. Fullness and brightness did not seem to be independent of pitch. The variability among listeners with respect to the ratings was large as indicated by the size of the Q values.

Scaling and Factor Analysis of Multitones

Subjects were required to rate 32 complex sounds, multitones, on a nine-point scale for loudness, referring to the over-all softness or loudness of a sound; loudness coherence, referring in a negative sense to the ease with which a judgment of loudness may be made and in a positive sense to the dispersion of the loudness of a sound; pitch, referring to the over-all highness or lowness of a sound; pitch coherence, again referring in a negative sense to the ease with which a judgment of pitch may be made; coherence, referring to the over-all unity or non-unity of a sound; numerosity, a related factor to coherence, referring to the number of distinguishable components of a complex sound; clarity, referring to a dimension where muffled or highly damped represents one extreme and ringing or undamped represents the other extreme; roughness, referring to a continuum where smoothness represents the opposite extreme; and fullness, referring to a dimension which runs from thinness to fullness.

Thirty-six students served as subjects. The subjects heard the sounds free field and were instructed as to how the ratings were to be made in terms of the above explanations of the arbitrarily selected properties. Practice sounds preceded the test sounds before each of the nine series of ratings. Inter-correlations (r) based on scale values were determined and are shown in Table IV.

Table IV.

Correlations Between Nine Selected Characteristics for 32 Complex Sounds

	Loudness Coherence	Pitch	Pitch Coherence	Over-All Coherence	Numerosity	Clarity	Roughness	Fullness
Loudness	.89	-.22	.76	.62	.70	.09	.54	.82
Loudness Coherence		-.34	.86	.77	.82	-.08	.63	.87
Pitch			-.25	-.11	-.17	.83	-.12	-.61
Pitch Coherence				.91	.96	-.07	.82	.72
Over-All Coherence					.96	-.16	.85	.68
Numerosity						-.02	.82	.66
Clarity							-.18	-.36
Roughness								.50

The inter-correlations between characteristics were factored by the centroid method (74). The resultant factor matrix is shown in Table V.

Most of the variance in factor loadings can be accounted for by the first three factors, 70, 20, and 7 % respectively for factors I, II, and III. Inspection of the factor matrix indicates that the variables loudness, coherence, numerosity, roughness, and fullness characterize the first factor. In making judgments of coherence, the observers were instructed that the low end of the scale represented most coherence and the high end of the scale represented least coherence. This accounts for the high positive correlations between coherence judgments and the numerosity and fullness judgments. Factor I is probably bi-polar, although this is not indicated by the loadings, with coherence defining one end of the scale and numerosity and roughness defining the other end of the scale. Factor II would appear to relate to pitch and clarity and

Table V.

Factor Matrix for Nine Selected Characteristics of Sound

Variables	Factors I	II	III	IV	V
Loudness	.760	.191	.216	-.073	.029
Loudness Coherence	.894	.142	.285	.305	.178
Pitch	-.450	.678	-.335	.152	.032
Pitch Coherence	.908	.342	.076	.042	-.005
Over-All Coherence	.853	.324	-.167	-.096	.045
Numerosity	.862	.404	.076	.144	.027
Clarity	-.272	.620	.022	.251	.112
Roughness	.751	.319	-.059	-.018	.008
Fullness	.878	-.189	.501	-.039	.005

factor III likewise may relate to pitch as well as to fullness.

The results of the several studies based on discriminability judgments by observers do not provide for clear-cut isolation of psychological dimensions of sound. It would appear that this type of unidimensional scaling model is not adequate to deal with the kind of psychological space involved in the perception of complex sounds. However, in this series of studies the use of another discriminability model to scale sounds was attempted. This was the method of paired comparisons.

SCALING BY THE TECHNIQUE OF PAIRED COMPARISONS

The observers' task was to judge which member of a pair of sounds was louder, higher, brighter, and smoother than the other member of the pair. The stimuli were 10 multitones which were programmed into all pair combinations. The sounds were presented through PDR-3 receivers. The observers were instructed as to their task and were allowed to practice on a series of sound pairs before they made the judgments on the test pairs of sounds. The experiment was conducted in four sessions, one for loudness, one for pitch, one for brightness, and one for smoothness judgments. Ten observers participated in this experiment.

Scale (S) values were determined for each of the 10 stimuli with respect to the four attributes of loudness, pitch, brightness, and smoothness. These values are noted in Table VI.

It may be seen, by inspection of the values in Table VI, that the ranges of scale values for smoothness and brightness were less than for loudness and pitch. Scale values followed a different order for loudness, pitch, and smoothness. The order of scale values was almost identical for pitch and brightness. The results of this study appeared promising, especially in terms of the ease of the task for the observers and as a method of determining the range of a property for stimuli where the physical dimensions are not known. However, because of the limited number of sounds which can be handled in this model, the computational time involved, and the observer judgment time involved, it does not appear to be a method which can be used to an advantage in the scaling of various groups of auditory stimuli.

Table VI.

Scale Values for 10 Stimuli on Four Selected Continua

Stimuli	Loudness	Pitch	Brightness	Smoothness
	Scale Values			
1	0.00	0.75	0.00	5.29
2	1.45	0.00	0.91	5.68
3	1.90	2.06	0.91	5.15
4	2.89	2.77	2.29	4.53
5	5.04	3.03	2.09	3.43
6	5.31	2.42	2.74	2.82
7	5.04	2.74	2.20	2.02
8	5.04	4.86	3.38	0.37
9	6.14	4.86	4.13	0.40
10	5.95	5.62	4.21	0.00

SCALING BY DIRECT MAGNITUDE ESTIMATION

The task for the observers was to make direct magnitude estimations of properties of complex sounds. The sounds consisted of 18 multitones. The selected properties for which the observers were to make magnitude estimations were loudness, loudness coherence, pitch, pitch coherence, over-all coherence, numerosity, clarity, roughness, and fullness. The observers were instructed as to the nature of the property which they were to judge and were told to use any numbering system which they felt was appropriate for the sounds which they were judging. Estimations of magnitude were made for the same 18 sounds for each of the nine properties. Each observer made four estimations for each sound for each property. The sounds were presented in random order and were heard by the observers for a duration of 15 seconds for each estimation. A standard sound was not provided for the observers in making their estimations. At the completion of the experimental session, each observer was asked to note the range of the numbering system which he had used in assigning numbers to the series of sounds. The observer's assigned values were converted to a hundred-point scale and the values on this new scale were recorded for each observer. Forty-five observers participated in the experiment and since each observer made judgments with respect to only one property, this resulted in five subjects estimating the magnitude of each property for the 18 sounds. The mean value of the four judgments made by each observer for each sound became the criterion value for analysis of the data. Product moment (r) correlations were determined between the nine properties using as values for each of the sounds the mean scores based on the judgments of the five observers who made estimations with respect to each property. The correlation matrix is shown in Table VII.

Table VII.

Correlations Between Nine Selected Properties for 18 Complex Sounds

	Loudness Coherence	Pitch	Pitch Coherence	Over-All Coherence	Numerosity	Clarity	Roughness	Fullness
Loudness	.65	.79	-.11	.86	-.59	.88	.27	-.84
Loudness Coherence		.47	.30	.42	-.34	.61	.45	-.47
Pitch			-.10	.72	-.76	.89	-.04	-.91
Pitch Coherence				-.20	.13	-.15	.33	.21
Over-All Coherence					-.43	.73	.29	-.73
Numerosity						-.78	.41	.68
Clarity							-.08	-.92
Roughness								.07

Inspection of the values in the correlation matrix would seem to indicate that, for these stimuli, pitch and loudness varied in a similar manner. Values for coherence and roughness seemed to be different, however, from each other and from the other values. Inspection of the values for each observer indicated that there was a considerable amount of variation, both in a single individual's successive judgments of the same stimuli, and between different observers. The estimation of magnitude appeared to be a difficult task for observers. The extremely variable responses which were obtained both between individuals and within an individual's responses would appear to limit the usefulness of this scaling procedure in the area of auditory perception of complex sounds.

OBSERVER GENERATED SCALES

Several attempts were made to obtain the dimensions and the values on the dimensions directly from the observers. In one experiment the task for the observers was to place six complex sounds, multitones, into as many orders as possible and describe their basis for ordering the sounds. By means of selector switches the observer could alternate among the various sounds and listen to samples of each sound for durations of his own choice. Ten observers participated in this task. The approach was not successful in that there was little agreement among observers and most of the individuals did not go beyond pitch and loudness in selecting their own dimensions.

In a similar manner, subjects were asked to place each of 150 sounds in categories of their own choosing. The sounds were on individual magnetic tape loops. Even with unlimited time to complete the task, the subjects did not go beyond pitch and loudness in selecting categories for the 150 sounds.

DIFFERENTIATION OF AUDITORY PATTERNS

In the studies which have been discussed up to this point, the concern has been on observer judgments of amounts of assumed properties of sound and on the measurement model within which these judgments have been evaluated. The study reported on in this section was concerned with the effect of changing the amplitude of one component tone of a multitone series upon listeners' judgments of sameness or difference of a sound. The objective was to relate the physical changes in the sound to the observer's report of a change in the sound and attempt to determine the psychological property of the sound which had changed, resulting in the observer's response of a difference in the sound.

Five experienced listeners heard stimuli composed of either two, three, four, five or six pure tones. The stimuli were altered by progressively increasing or decreasing the amplitude of one of the component tones. Each listener's task was to signal when he perceived that the sound was no longer the same as the original sound. The listeners were instructed not to respond to only a change in loudness of the sound. The tones used were 700, 800, 900, 1000, 1100 and 1200 cps. The initial auditory stimuli for each judgment were all possible two-, three-, four-, five-, and six-tone combinations of these tones with equal amplitudes for all tones. Each listener's score was the mean of five trials in decibels of change required for him to perceive a change in the stimulus. Mean values and standard deviations were determined on the basis of the responses of the five listeners. The overall mean value for all tones in all combinations with respect to the amount of increase in amplitude needed to result in a response of perceptual difference was 2.7 db with a mean standard deviation of 0.97 db. The overall mean for a decreased level of a tone was 3.8 db with a mean standard deviation of 1.29 db. Other than the consistent difference between values for increased and decreased levels of component tones, inspection of the data did not reveal consistent trends which relate frequency separation between tones and number of tones to the listeners' judgments of a change in auditory experience. Several factors may relate to the negative results of the study including the fact that the component tones were not widely separated in frequency and that the measure of amplitude of one component is probably not a good measure of physical parameters which are changed when one tone of a group of tones is altered in amplitude.

CHANNEL CAPACITY

Channel capacity is defined in communication theory framework as the upper limit of the extent to which an observer is able to match his responses to the stimuli to which he is responding. If it is assumed that an individual's ability to make absolute judgments about stimuli relates to the psychological properties of the stimuli, then from measures of channel capacity it should be possible to make inferences about the dimensions of auditory experience.

Channel Capacity as a Function of Discriminable Units of Pitch and Loudness

The purpose of this experiment was to evaluate information transmitted as a function of difference limen separation and difference limen range of components for pitch and loudness singly and combined. Six auditory conditions were involved in the study. These were pitch, approximately 110 DL range; loudness, approximately 110 DL range; pitch, approximately 55 DL range; loudness, approximately 55 DL range; the first two conditions combined; and the third and fourth conditions combined. The pure tone stimuli for each of the six conditions are shown in Table VIII.

Table VIII.

Pure Tone Stimuli for Six Conditions of Difference Limen Range of Loudness and Pitch

Conditions											
I		II		III		IV		V		VI	
cps	db	cps	db	cps	db	cps	db	cps	db	cps	db
1000	85	1000	70	1125	85	1000	77	1000	70	1125	77
1125	85	1000	77	1187	85	1000	81	1125	77	1187	81
1250	85	1000	85	1250	85	1000	85	1250	85	1250	85
1375	85	1000	92	1313	85	1000	89	1375	92	1313	89
1500	85	1000	100	1375	85	1000	92	1500	100	1375	92

Four duration times of the components of the display system were also included as part of the experimental conditions. These were 1.0, 0.5, 0.25, and 0.125 seconds duration.

Twenty-five observers served as subjects. The observer's task was to select from five sequentially presented stimuli one which corresponded to a stimulus which he heard immediately preceding the stimuli sequence. Each observer made five judgments with regard to each stimulus for each condition.

The criterion measure for each observer was the number of correct identifications for each display condition for each display time. The data were treated with analysis of variance instead of determining the usual communication theory measure, bits of information, because the interest was in whether or not significantly more information was transmitted by stimuli as frequency and amplitude range was increased. The results indicated that a significantly greater amount of information is transmitted as frequency and amplitude range between stimuli is increased ($F = 24.55$, required for 5 % level, 3.35), that roughly the same amount of information is transmitted by pitch and loudness for similar DL separation of components, and that the amount of information transmitted is not independent of the duration of the stimuli ($F = 4.56$, required for 5 % level, 4.28). The mean values of correct stimuli identifications for the six stimuli conditions and the four duration conditions are shown in Table IX. A difference of 0.53 between mean values, columns and rows, indicates a significant difference between means at the 5 % level.

Table IX.

Mean Values of Correct Identifications for Six Stimuli Conditions at Four Durations

	Conditions					
Duration(Seconds)	I	II	III	IV	V	VI
1.0	3.16	3.04	2.68	2.24	3.68	3.32
0.5	3.04	2.96	3.00	2.04	3.52	2.76
0.25	3.00	2.60	2.04	1.88	3.44	2.80
0.125	3.16	3.32	2.36	2.32	3.24	3.04

That a predictable relationship exists between information transmitted by stimuli and dimensions and scale values of stimuli, as was indicated in this study, is of value in studying dimensions of sound because while it is probably not possible to directly infer dimensions from amount of information carried, it should be possible to test the adequacy of a proposed set of inclusive dimensions and scale values by measuring the amount of information transmitted by the stimuli. That is, if the dimensions and the stimuli values on the dimensions are correctly guessed, the predicted information transmitted should correspond to the actual information transmitted.

Confusion Matrices

One way in which the channel capacity concept may be used to infer psychological dimensions for a set of auditory stimuli is by the use of confusion matrices. In a confusion matrix the results of observers' absolute judgments for a set of input stimuli are noted in terms of the correspondence between the stimulus which was given and the stimulus which the observer selected as being given. By noting the error responses of the observer and combining and regrouping the stimuli to minimize error responses, inferences may be made as to the dimensions of the set of stimuli. This was the technique used by Miller and Nicely (46) to determine the distinguishing features for a set of consonants. They induced error responses on the part of the observers by submitting the consonants to noise and frequency distortion, and studied the pattern of errors which resulted, noting that five distinctive features (dimensions) appeared adequate to explain observer confusions in judging consonants.

An attempt was made to use this approach to determine a set of psychological dimensions for complex sounds. The stimuli were 20 multitones. An observer was required to learn an identifying number associated with each sound. The sounds were then presented under various conditions of noise and high- and low-pass frequency filtering. The experiment was not successful. The observer, even after several weeks of daily learning trials, was not able to correctly identify the stimuli to a satisfactory criterion. The error responses, also, did not seem to fall in patterns but rather, under the noise and distortion, randomized. Although the negative results may have been due in part to the particular observer who participated, the results would seem to indicate that this approach is better suited for evaluating stimuli systems which are already known to individuals, such as language, than in application to stimuli systems unknown to observers because of the difficulty observers have in learning new sets of complex auditory stimuli.

EFFECT OF COMPLEXITY OF STIMULI UPON LEARNING

In this experiment three groups of stimuli of differing physical complexity were prepared. The first group consisted of tones of the frequencies of 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 cps. For the second group of sounds, the first multiple of each tone was added to each tone to produce a set of 10 multitones. For the third set of stimuli, the first and second multiples of each of the original tones were added to each tone to produce the 10 stimuli. The subject's task was to learn a word associated with each of the sounds. Learning was by the anticipation method. Three groups, each composed of 27 persons, served as subjects. Amount of learning was measured on the last 15 of 20 trials.

The criterion measure for each subject was the number of correct anticipations on the 15 trials. The significance of the difference between the means among the three groups of subjects was evaluated by analysis of variance. The subjects who learned words paired with pure tones did significantly better than did the other two groups of subjects ($F = 3.42$, required for 5 % level, 3.23), indicating that with increased complexity of stimuli, learning is more difficult. These results may have been due in part, however, to increased similarity between complex stimuli which resulted from adding first and then first and second multiples to each of the original tones. The multiples in some instances corresponded to frequencies of components of other stimuli.

It was noted previously that information transmitted by stimuli should be predictable if the dimensions and stimuli values on dimensions are known for a set of stimuli. Learning as it relates to the complexity of stimuli is part of the same problem. While it may not be possible to predict dimensions and scale values for stimuli from measures of learning, it should be possible to anticipate learning by knowing properties of stimuli and values of the stimuli with respect to the properties. The adequacy of a set of dimensions may thus be tested by measuring learning. The knowledge that is necessary, however, is the general relationship between learning and dimensions of stimuli. The possibility exists that individuals only extract that information which is necessary in terms of dimensions for learning and may discard dimensions which are not necessary to learn a particular set of stimuli. This structuring of auditory perception by individuals may confound an attempt to relate dimensions to learning.

MULTIDIMENSIONAL SCALING

Several studies were conducted which involved multidimensional scaling. The stimuli for these studies were pure tones, complex sounds, and speech sounds. These experiments are discussed in greater detail than the other studies summarized in this report because the results of these studies are more directly concerned with the basic purpose of this program of research than are the other studies and because time wise, a larger proportion of the effort was spent upon multidimensional scaling than with the other experimental techniques.

Multidimensional Scaling of Pure Tone Stimuli*

The purpose of this investigation was to apply a multidimensional scaling model to an auditory area of known dimensionality, that of pure tones, in order that this model might be evaluated for use in auditory areas where the dimensions are not well known.

Three basic steps were involved in the multidimensional scaling procedure used in this study (43). First, comparative distances in similarity were obtained between all pairs of pure tone stimuli; second, an estimated additive constant was used to convert the comparative distances into absolute distances; and third, the dimensionality of the psychological space necessary to account for the absolute distances among the pure tone stimuli were determined.

The stimuli were 16 pure tones, four frequencies at four sound pressure levels. The frequencies were 500, 1000, 2000, and 3000 cps each at the sound pressure levels of 70, 80, 90, and 100 db (re .0002 dyne/cm^2). The stimuli are identified by number in Table X.

Table X.

Pure Tone Stimuli Identified by Stimulus Number

Stimulus Number	Frequency (cps)	Sound Pressure Level (db)
1	500	70
2	1000	70
3	2000	70
4	3000	70
5	500	80
6	1000	80
7	2000	80
8	3000	80
9	500	90
10	1000	90
11	2000	90
12	3000	90
13	500	100
14	1000	100
15	2000	100
16	3000	100

The 16 pure tones were programmed in all possible pair combinations. This resulted in 120 pairs. The duration of each tone was one second. A one-second interval separated the two tones of each pair. An identifying carrier number preceded each stimulus pair with an interval of seven and a half seconds between the onset of each carrier number. The presentation order of the pairs of tones was determined by random selection.

Thirty-nine male and female college students served as observers.

* This study is discussed more fully in WADC Technical Report 59-201, Multidimensional scaling approach to the determination of basic psychological parameters for pure tones.

The observers' task was to judge each pair of sounds for similarity on a nine-point scale. Number one on the scale represented extreme similarity and number nine represented extreme dissimilarity. The directions for making the judgments were tape recorded and also printed on the first page of the test booklet so that the observers both heard and read the directions prior to each experimental session. Twenty practice pairs of tones preceded the test stimuli. After the observers had judged the practice stimuli, the directions were briefly repeated.

The observers heard the stimuli through headsets and took part in the experiment in groups of nine or less. The comparative distances between stimuli were determined on the basis of the observer's judgments.

The raw data were tabulated in a 120 (stimulus pairs) x 9 (scale categories) table where the cell values, f_{ig}, indicated the frequency that the i^{th} pair of stimuli were placed in the g^{th} category. The frequencies were cumulated, converted to proportions, and from the proportions were determined normal deviate values, z_{ig}. The deviate values were weighted according to the function, Z^2/pq, where Z was the ordinate of the normal curve which corresponded to the proportion, p, and $q = (1-p)$. A table of successive differences in deviate values was constructed and the differences themselves were weighted according to the formula, $W_1W_2/W_1 \not+ W_2$, where W_1 referred to the weight applied to one deviate value and W_2 referred to the weight applied to the other deviate value. Weighted averages of successive differences were determined and scale values, t_g, were computed. This scale was used as the ordinate in determining graphically the scale values and discriminal dispersions for each stimulus pair. The resulting scale values for each pair were converted to positive values by setting the smallest scale value to zero. These values, s_{jk}, represented comparative interpoint distances.

Matrices, S^2_{jk}, squared relative distances, A, E, and H were constructed. The elements, a_{jk}, e_{jk}, and h_{jk} of matrices A, E, and H were:

$$a_{jK} = 1/2 \left(1/n \sum_{K}^{n} S^2_{jK} + 1/n \sum_{j}^{n} S^2_{jK} - S^2_{jK} - 1/n^2 \sum_{j}^{n}\sum_{K}^{n} S^2_{jK} \right)$$

$$e_{jK} = 1/n \sum_{K}^{n} S_{jK} + 1/n \sum_{j}^{n} S_{jK} - S_{jK} - 1/n^2 \sum_{j}^{n}\sum_{K}^{n} S_{jK}$$

$$h_{jK} = 1/n \quad \text{for } j \neq K \quad \text{and} \quad h_{jj} = (1 - 1/n)$$

A B* matrix was determined ($B^* = A \not+ cE \not+ 1/2c^2H$) using as the additive constant, $c = 3.75$.

The B* matrix was solved for eigen values and eigen vectors. Five non-zero roots and their corresponding vectors were retained and the coefficients for the sum of the latent roots of the B* matrix were computed using the following equation:

$$\sum \beta = X_i' A X_i + c \sum_{i}^{p} X_i' E X_i + 1/2 \; pc^2$$

In the above equation p = the number of roots and c is unknown. The coefficients for the sum of the diagonals of the B* matrix were determined. The equation for the sum of the diagonal elements was:

$$\sum_{j}^{n} b^*_{jj} = \tfrac{1}{2} n \sum_{jk}^{nn} S^2_{jk} + \tfrac{c}{n} \sum_{jk}^{nn} S_{jk} + \tfrac{1}{2} (n-1) c^2$$

The two resulting equations were:

$$\sum \beta = 44.6125 + 14.7534 \; c + 2.5 \; c^2$$

$$\sum_{j}^{n} b^*_{jj} = 35.7818 + 29.6663 + 7.5 \; c^2$$

These equations were set to equal each other and the resulting quadratic was solved for two values of c: $c = -5.8147$ and $c = 0.8437$. The root which gave the larger $\Sigma\beta$ was $c = 0.8437$. Because c was a small positive number, the analysis was continued with c set for $c = 0$.

A factor matrix was computed using the general matrix factoring solution, $(BX) K^{-1} = F$, where X is the matrix of X_i vectors, and K is the matrix where $X'BX = K'K$. The factor matrix was subjected to orthogonal rotation to achieve simple structure and meaningful dimensions. The final rotated matrix of projections is shown in Table XI. The projections of stimuli for the final rotated matrix for factors I, II, and V are shown in Figures 1 and 2.

The results would seem to indicate that two factors, pitch and loudness, are fairly well defined. A third factor, not as well defined, seems to relate to pitch. Factor I appears to be a pitch dimension with the frequencies of 500, 1000, and 2000 cps producing a continuum for the several sound pressure levels. The exception for this factor is that the 3000 cps stimuli occupy approximately the same scale positions as the 2000 cps tones. The major loadings for factor II are accounted for by the four 3000 cps stimuli and with respect to these stimuli appears to relate to pitch. Factors III and IV do not have sufficient loadings to be considered legitimate dimensions. Factor V appears to be a loudness dimension which is best defined by the scale positions for the 1000, 2000, and 3000 cps stimuli and not well defined for the 500 cps stimuli.

It would appear on the basis of the results of this study that the multidimensional scaling model was successful in isolating the basic psychological parameters, pitch and loudness, for pure tones and could therefore be of value as a model for exploring auditory areas where the dimensions are not well known.

Table XI.

Final Rotated Matrix of Projections for Pure Tone Stimuli

			Factors		
Items	I	II	III	IV	V
1	-1.543199	-0.424865	-0.164403	0.321357	0.028359
2	-0.241817	-0.417206	-0.280011	-0.423572	-1.192284
3	0.510181	0.062482	1.394003	-0.406371	-0.156507
4	0.417622	1.748473	0.123048	0.236991	-0.435328
5	-1.667362	0.310291	0.173105	-0.128389	-0.392301
6	0.152980	-0.642132	-0.525826	-0.132984	-0.830362
7	0.459168	-0.125967	0.536876	1.592291	0.069693
8	0.669558	1.262105	0.016588	0.092865	0.479725
9	-1.695652	-0.546389	-0.296510	-0.278850	0.261841
10	0.290617	-1.479993	-0.546379	-0.335696	0.003222
11	0.968834	-0.218335	0.539932	0.243135	0.045479
12	0.750189	0.993678	0.013601	0.415003	0.839024
13	-1.381028	-0.860953	1.564339	-0.749929	0.104332
14	0.290271	-0.471981	-0.608213	-1.854985	-0.316614
15	1.047424	-0.259593	1.080559	-0.231475	0.508725
16	0.725420	1.375894	-0.428802	0.181480	1.123717

Multidimensional Scaling of Complex Sounds

The same multidimensional scaling model was applied to a set of complex sounds. The stimuli were 32 multitones. These 32 tones were selected from 100 multitones which had been rated by observers for loudness, pitch, volume, clarity, roughness, density, fullness, and brightness on a seven-point scale. Twenty observers made judgments with respect to each attribute. Four sounds were chosen, on the basis of these judgments, which represented low, medium-low, medium-high, and high scale values for each attribute. The 32 sounds were selected in this manner in order that after the multidimensional scaling had been completed, it could be determined whether or not these attributes could be related to the dimensions revealed by the multidimensional model.

The original 100 multitones were prepared with the restriction that each sound have at least two and not more than 10 pure tone components. The number of components was randomly determined. The components for each sound were also chosen by random selection with any frequency between 20 and 10,000 cps having an equal chance for inclusion.

The 32 complex sounds were programmed into all possible pair combinations, resulting in 496 pairs. The duration of each sound was approximately two seconds with a brief interval separating the sounds of each pair. An identifying carrier number preceded each pair of sounds with an interval of approximately five seconds between the onset of each carrier number. The order of presentation of the sound pairs was randomly determined. The 32 stimuli are identified by stimulus number in Table XII.

Forty-six male and female college students served as observers in making judgments of similarity-dissimilarity with regard to the stimuli pairs.

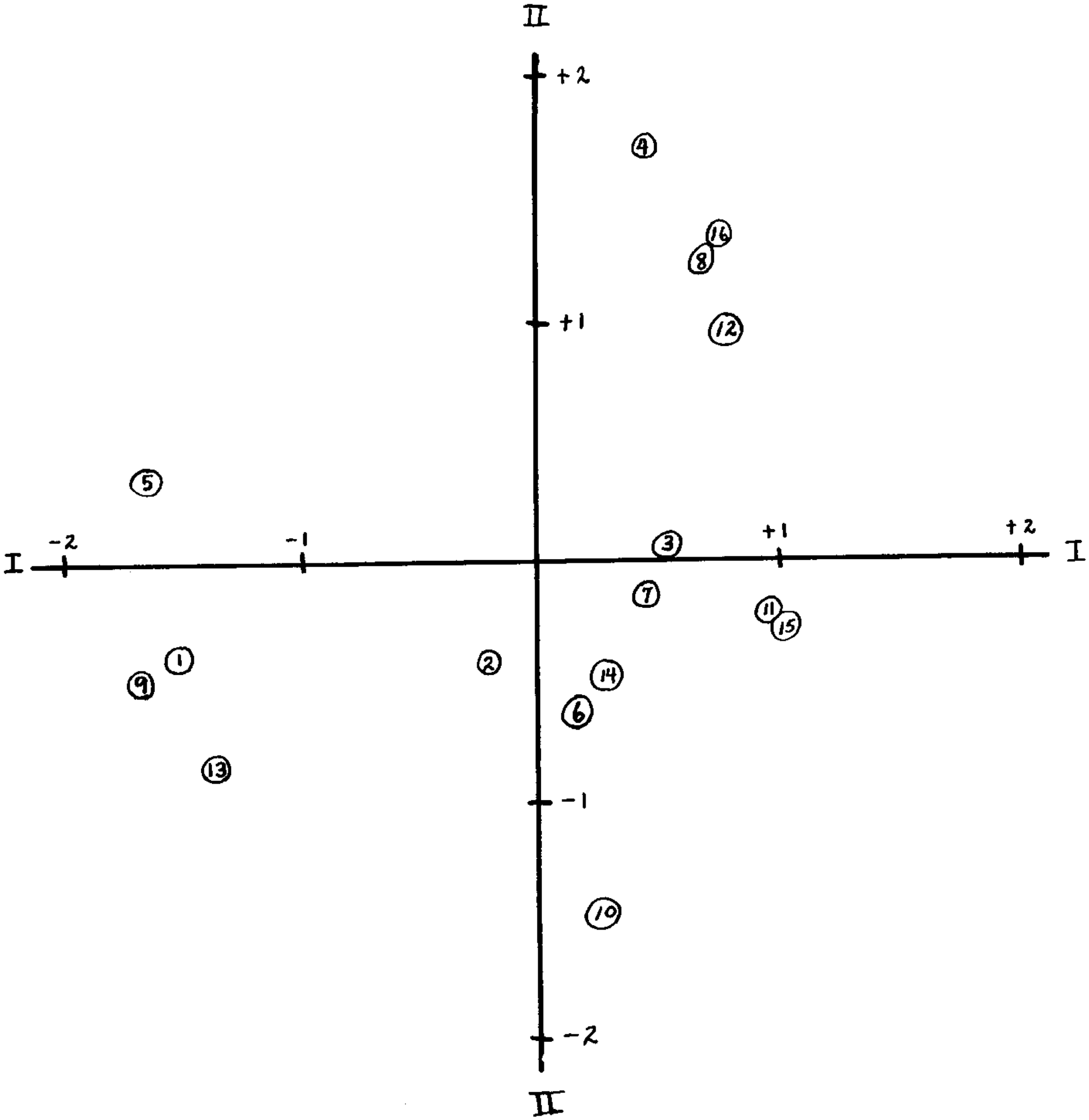

Figure 1. Factor I versus Factor II, Fotated Factor Matrix, for 16 Pure Tone Stimuli

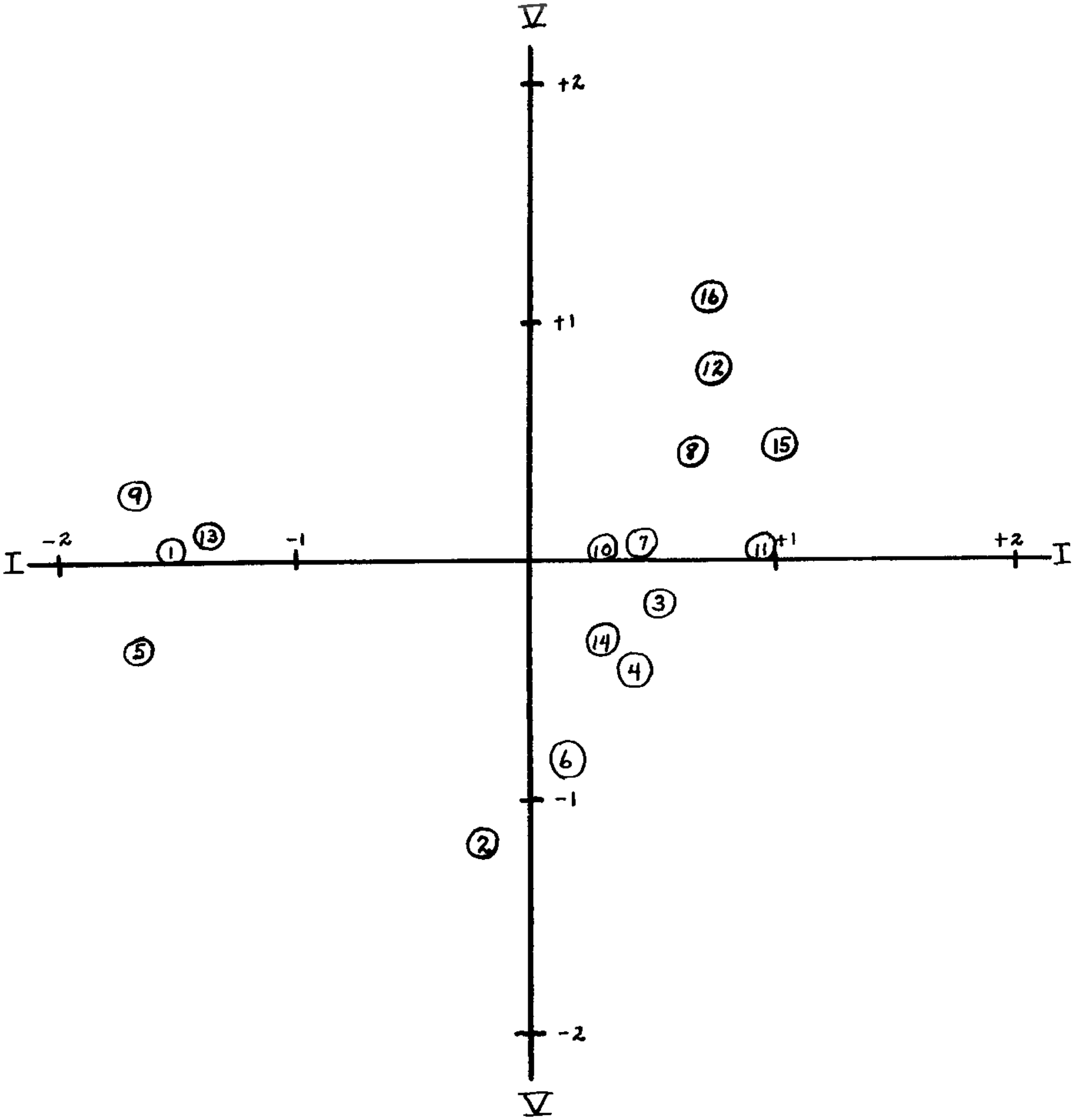

Figure 2. Factor I versus Factor V, Rotated Factor Matrix, for 16 Pure Tone Stimuli

Table XII.

Complex Sound Stimuli Identified by Stimulus Number

Stimulus Number	Approximate Frequency of Components (cps)
1	636-813-8199-8601
2	7736-9882
3	6376-6424-8005-616-5696-4128-3340-1609
4	9769-9829-5020-9943-7430-5717-5007-5798-4457
5	6652-6476-151-5509-6126-461
6	3110-5920
7	7092-2264-260-7510-472-4011-1375-3977-1784
8	5593-3469-4467-6085
9	7081-6953-6012
10	5078-5188-9830-7051-5178-603
11	4759-2310-3383-3972
12	1602-4723-8941-4953-7391-2924
13	111-4990
14	2142-8908-5390-3784-4165-3930-9209-8269-9572
15	7404-6422-1201-2240-8440-924
16	3110-8013-5484-1018-5219-3168-2078-1152
17	2974-1410-149-9705-8524-2939-4395-7077-3115-6802
18	9130-4778-4930-5245-8203-1376-279
19	4062-5121-606-8100-8205
20	7009-4229
21	99-6229-2911-5575-9265-4787-2660-5554-9093-6722
22	9919-659-8829
23	8408-183-2818-5583-5591-2014-346-7016
24	202-9103-5015-9950-5298-9928-135-320
25	6369-6053-1468-8509
26	3191-6612
27	9387-8896-4407-5380-8999-2423-1207-2735-7829
28	9204-4887
29	1954-8112-6743-9921-8782-4259-2200-6504
30	561-8139-8200-427-1123-8827-3128-3952
31	163-3785-1219-1499-3372-5849-8405-762
32	2506-2072-1949-9304-7373-8224-4470

The observers judged each pair of sounds for similarity on a nine-point scale. The same as in the previous study, number one on the scale represented extreme similarity and number nine represented extreme dissimilarity, with other points on the scale representing equal distances in similarity-dissimilarity. The observers judged practice pairs of stimuli before judging the test stimuli. The sounds were heard by the listeners through earphones.

The observers judgments were tabulated in a 496 (stimulus pairs) x 9 (scale categories) table in which the cell values indicated the frequency that a stimulus pair was placed in a particular category. Two analyses were made with respect to the data. One analysis followed the procedure as outlinedin the preceding section with regard to multidimensional scaling of pure tone stimuli, with the computations being done for the most part on a desk calculator. The eigen values and eigen vectors for the B* matrix, however, were determined through use of a computer. Essentially the same method that was described in the preceding pure tone study for obtaining comparative distances between stimuli, using a categorical judgment model and determining dimensions and scale values using a multidimensional scaling model, was programmed on an IBM 1103A computer. The second analysis of the data was done on this computer. Since the results of the two analyses, the desk calculator and computer, were comparable, only the results of the latter are reported in this paper.

The input data for computer processing was the 496 (stimulus pairs) x 9 (scale categories) matrix in which the cell values were the frequency with which individuals selected a scale category number to indicate their judgment of similarity-dissimilarity for each stimulus pair. Comparative distances between stimuli were determined, and using the largest comparative distance element as an estimated additive constant, c = 3.224589, a B* matrix was obtained. The B* matrix was solved for eigen values and eigen vectors. The eigen values were: 61.78, 33.95, 28.56, 21.74, 19.68, 18.36, 17.48, 16.69, 15.49, 14.97, 13.46, 13.04, 12.37, 11.69, 11.17, 10.34, 9.51, 8.65, 8.52, 7.55, 7.41, 6.73, 5.91, 5.08, 4.72, 4.05, 3.38, 2.18, 2.01, -0.000004,20, -1.16, and -2.71. Twenty-seven eigen values and their corresponding vectors were retained. Two more successive estimations were made of the additive constant and two more successive determinations were obtained providing two new B* matrices. The third approximation to the constant, c = 3.1923681, was in close agreement with the first estimation of the constant. The final set of eigen values also corresponded closely to the values listed above. The factor matrix was determined using the third B* matrix, and, in order to determine the adequacy of the factorization, the sum of squares of the B* matrix was compared to the sum of squares of the factor matrix. The two sums of squares were 9246.9395 and 9229.6790, indicating good factorization.

The matrix of projections, for 16 factors, for the 32 complex sounds, is shown in Table XIII. The projections of the stimuli for factors I - VI are given graphically in Figures 3 - 7. Inspection of the eigen values and of the factor loadings for the 32 stimuli would indicate that possibly three, and not more than six, factors should be considered as legitimate dimensions. Two methods were used in an attempt to identify and name the dimensions. In the first method the stimuli projections were evaluated with respect to the unidimensional scaling of the sounds which was completed prior to the multidimensional scaling. On the basis of unidimensional scaling, each successive four stimuli represented scale points from least to most for eight arbitrarily selected attributes. Stimuli 1 - 4 represented loudness; 5 - 8 represented pitch; 9 - 12, volume; 13 - 16, clarity; 17 - 20, brightness; 21 - 24, density; 25 - 28, fullness; and roughness, stimuli 29 - 32. In relating these unidimensional results to the multidimensional results, factor I, multidimensional scaling, appears to be a density, brightness, and pitch dimension. This may be noted in Figure 3 where stimuli 21, 22, 23, and 24, stimuli 20, 19, 18, and 17, and stimuli 8, 7, 6, and 5, representing density, brightness, and pitch respectively, each form progressions on factor I. Factor II for the multidimensional scaling appears to relate to the unidimensional scaling attribute of clarity as indicated by the stimuli items 13, 14, 15, and 16. Other than these instances which have been noted, stimuli projections for the multidimensional scaling factors did not appear to relate to the scale values for stimuli on the attribute scales which were obtained in the unidimensional procedure. In programming the stimuli for presentation to the observers, an attempt was made to maintain relatively equal loudness for all sounds. This is probably the reason that a factor did not appear in the multidimensional scaling which corresponded to the unidimensional scaling attribute of loudness.

Table XIII.

Matrix of Projections for 32 Complex Sounds

Items	Factors I	II	III	IV	V	VI	VII	VIII
1	-0.76769	-1.41142	0.92316	1.21469	0.24767	-0.60573	-0.55854	-0.93763
2	-1.80448	0.60840	0.81120	0.24674	0.44200	0.03834	0.11490	0.28208
3	-0.34479	-1.44550	-0.52751	0.34407	0.43519	0.20537	0.06837	1.63445
4	-2.08777	0.45293	0.10846	-1.21121	-0.58268	-0.80200	0.87086	-0.50404
5	2.55705	1.65483	-0.15429	0.34734	0.41690	0.51491	0.52361	0.32155
6	2.32283	-0.59543	1.40473	-0.89424	-1.02982	-0.20937	0.03000	-0.20790
7	0.58711	-1.02472	-0.59522	-0.98036	0.28111	0.76438	-1.49670	0.53033
8	-1.38567	-0.36360	-0.90980	-0.10880	-0.61668	-0.61399	0.76286	-0.10388
9	-1.50663	-0.68556	-0.17633	0.47015	-0.41148	-0.77456	-0.63103	0.82874
10	0.84824	-0.51526	0.27073	1.80251	-0.52030	-0.79800	0.01792	0.54439
11	-1.14036	-0.48093	-0.73806	-0.86234	-0.23179	-0.37243	-0.20679	0.22349
12	-0.56970	-0.53683	-1.05178	-0.61342	-1.61826	-0.37504	-0.85656	-0.02129
13	1.29975	2.77669	-0.09526	-0.73328	0.05281	-0.15947	0.86318	1.33449
14	0.21446	-0.04599	-1.51658	0.50264	-0.96406	1.01729	0.09391	-0.64664
15	-0.12763	-1.12391	-0.08608	-0.20736	0.93280	-1.19095	-0.49618	-0.45059
16	-0.72154	-1.22246	-0.64069	-0.73392	1.18399	-0.73544	0.35790	0.57430
17	2.12956	-0.38028	-0.54507	-1.54210	0.07809	-0.88370	0.96337	-1.01248
18	0.87509	-0.11975	-1.06536	0.76010	0.51931	-0.27909	0.84862	0.14503
19	0.16852	-0.65487	-0.56835	1.51532	-0.88890	0.12360	1.55134	-0.04597
20	-1.56141	0.77708	1.06855	-1.06519	1.23393	0.45807	0.29002	0.13102
21	-1.11846	1.02865	-0.30036	0.04641	-0.92612	0.67428	-0.41385	-1.17224
22	-0.17472	-0.51629	2.45632	1.04564	0.25431	1.17657	0.58851	-0.25863
23	1.01594	0.69331	-0.75437	-0.46965	-0.12237	1.54859	0.21328	-1.40429
24	2.78058	1.94956	-0.22129	0.62235	-0.18016	-1.55162	-1.39468	0.21555
25	-1.16532	0.14210	-0.17664	1.04537	0.91909	-0.21661	-1.06639	-0.90754
26	-1.74645	1.48999	-0.91495	-0.10608	-0.00020	0.24552	-0.70331	-0.10101
27	-0.76995	-0.74588	-0.45218	-0.45950	-0.43485	-0.01174	0.23223	-0.14454
28	1.32332	-1.20438	2.24959	-0.69485	-0.29166	0.01977	-0.25966	0.20746
29	-1.75932	0.76886	0.24521	0.58338	1.12257	0.03315	-0.25548	-0.06751
30	0.08430	-1.03640	-0.45870	-0.34872	-0.07305	1.72693	-1.16629	1.42946
31	1.33349	-0.10176	-1.40862	0.03704	1.87611	0.46283	0.15516	-0.86271
32	-1.22466	0.40785	0.89935	0.44665	-1.10417	0.47882	0.04990	0.44651

Table XIII (cont.)

Matrix of Projections for 32 Complex Sounds

Items	Factors IX	X	XI	XII	XIII	XIV	XV	XVI
1	0.12024	0.20932	-0.49333	-0.83825	-0.25419	0.34956	-0.74177	1.22981
2	-0.77452	-0.29159	1.37775	0.60022	0.63756	0.50579	0.71235	-0.36811
3	0.52319	-0.32601	-0.75516	-0.38360	-0.21874	-0.06213	-0.19985	-1.10418
4	0.89819	0.50391	-0.30264	0.04783	0.23089	-0.05692	0.31685	-0.32878
5	-0.15834	-0.83147	-0.32817	0.48158	0.29763	-0.74326	0.87796	1.38144
6	-0.05717	-0.61889	-0.30030	1.15754	0.08382	0.48472	-0.25133	-0.66244
7	1.22807	-0.07443	0.88284	-0.05223	-0.55814	-0.32570	1.03719	0.59183
8	0.47269	-0.31242	-0.23380	-0.14178	1.49974	-0.76141	-0.92942	0.37924
9	-0.29735	0.52303	-0.81383	0.15494	-0.38960	0.52581	0.27061	0.09147
10	0.46555	-1.04496	0.46891	-0.08867	-1.04597	0.28207	-0.46713	-0.53443
11	-0.19317	0.70244	-0.82644	-0.02973	-0.48714	-1.05875	-0.18622	-0.08871
12	-0.50074	-0.13383	1.48189	-0.51829	-0.48030	-0.18788	-0.51759	0.55643
13	-0.89206	1.33384	0.29552	-0.90598	-0.61551	0.58358	-0.66551	0.10576
14	-0.07877	0.54254	0.38353	-0.55233	1.01909	1.00082	0.72293	-0.25969
15	-0.78814	1.18663	0.27537	0.20534	-0.43710	-0.04483	1.15415	-0.62024
16	-0.39036	-0.55994	-0.14545	0.35354	0.07719	1.44773	-0.37270	1.20480
17	0.29356	-0.31052	0.04808	-1.04767	-0.26593	-0.65705	0.52097	0.04544
18	1.52141	0.82437	0.58730	1.15914	0.12661	0.58001	-0.22303	0.36381
19	-0.63588	-0.14765	0.30709	0.06305	0.13814	-0.49003	1.00051	0.01055
20	-0.44491	-0.93372	0.87473	0.34914	0.25946	-0.50717	-0.67865	0.06597
21	-0.42639	-1.04693	-0.33363	-0.98321	0.12152	0.77032	-0.17785	-0.12706
22	-0.59221	0.89328	-0.02834	-0.53515	-0.54870	-0.30345	0.05514	0.26821
23	0.52590	0.37481	0.09769	1.17914	-1.21351	0.24437	-0.88533	-0.45298
24	0.00060	-0.40139	-0.10262	1.32351	0.57022	-0.45568	-0.07570	-0.61106
25	-0.75987	0.65058	-0.07767	0.93296	0.53475	-0.66730	-0.60335	-0.10349
26	1.09268	-0.14472	-1.47357	0.03813	-0.11369	0.71440	0.88338	-0.24802
27	-1.42018	-1.24491	-0.69422	0.76809	-0.86924	-0.06467	0.30416	-0.01321
28	0.27876	0.49820	0.24545	-0.09108	1.25555	0.58577	-0.18706	-0.34510
29	1.16069	-0.92504	0.81290	-1.00977	-0.40142	-0.30331	-0.06598	-0.56767
30	-0.47724	-0.01235	-0.49226	-0.26018	0.64532	-0.28051	-0.27671	-0.21030
31	-0.28569	0.16314	-0.62441	-0.67402	0.40039	-0.09191	-0.15631	-0.56148
32	0.59132	0.68720	-0.11314	0.48892	0.00166	-1.01169	-0.19497	0.41610

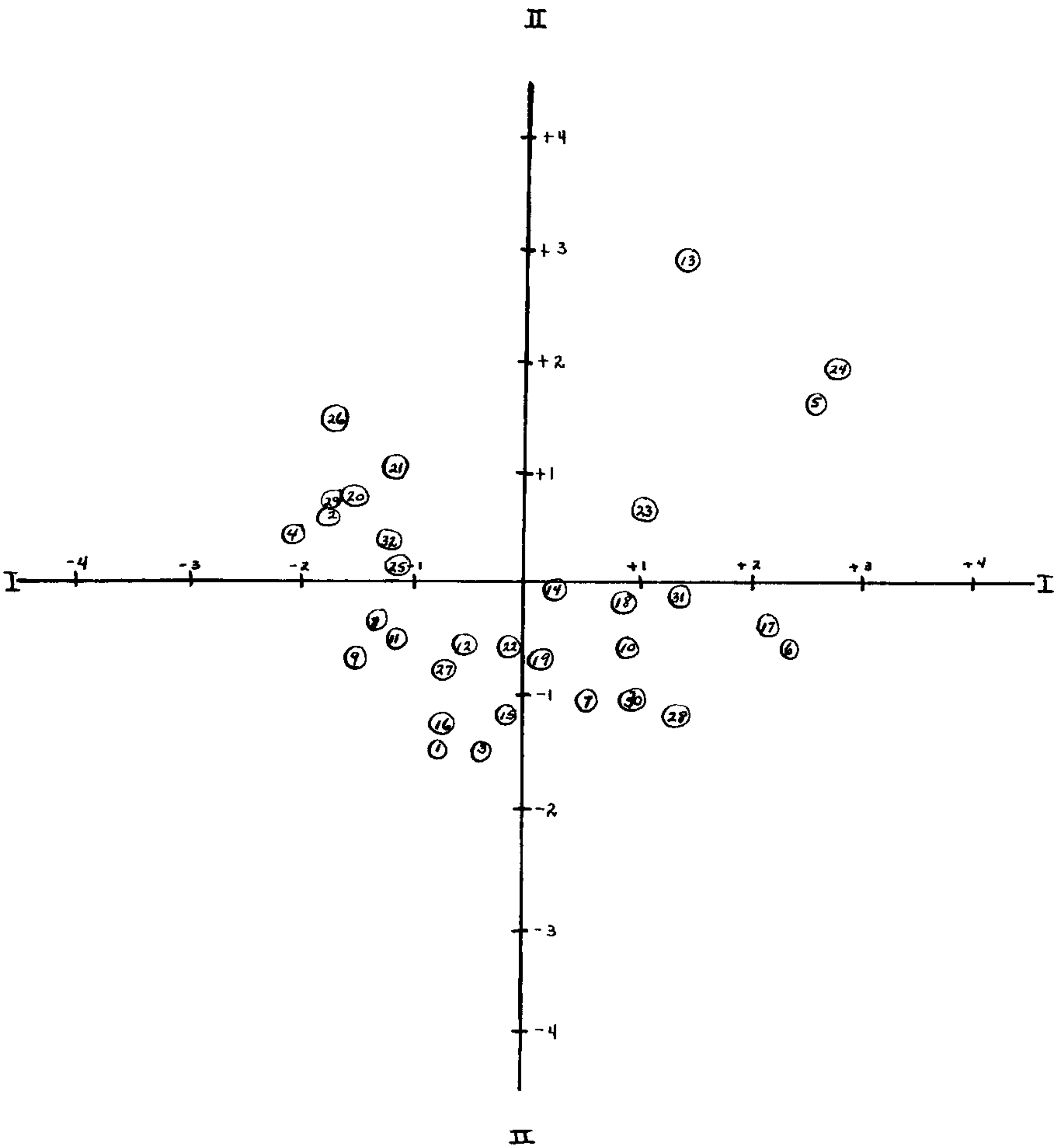

Figure 3. Factor I versus Factor II for 32 Complex Sounds

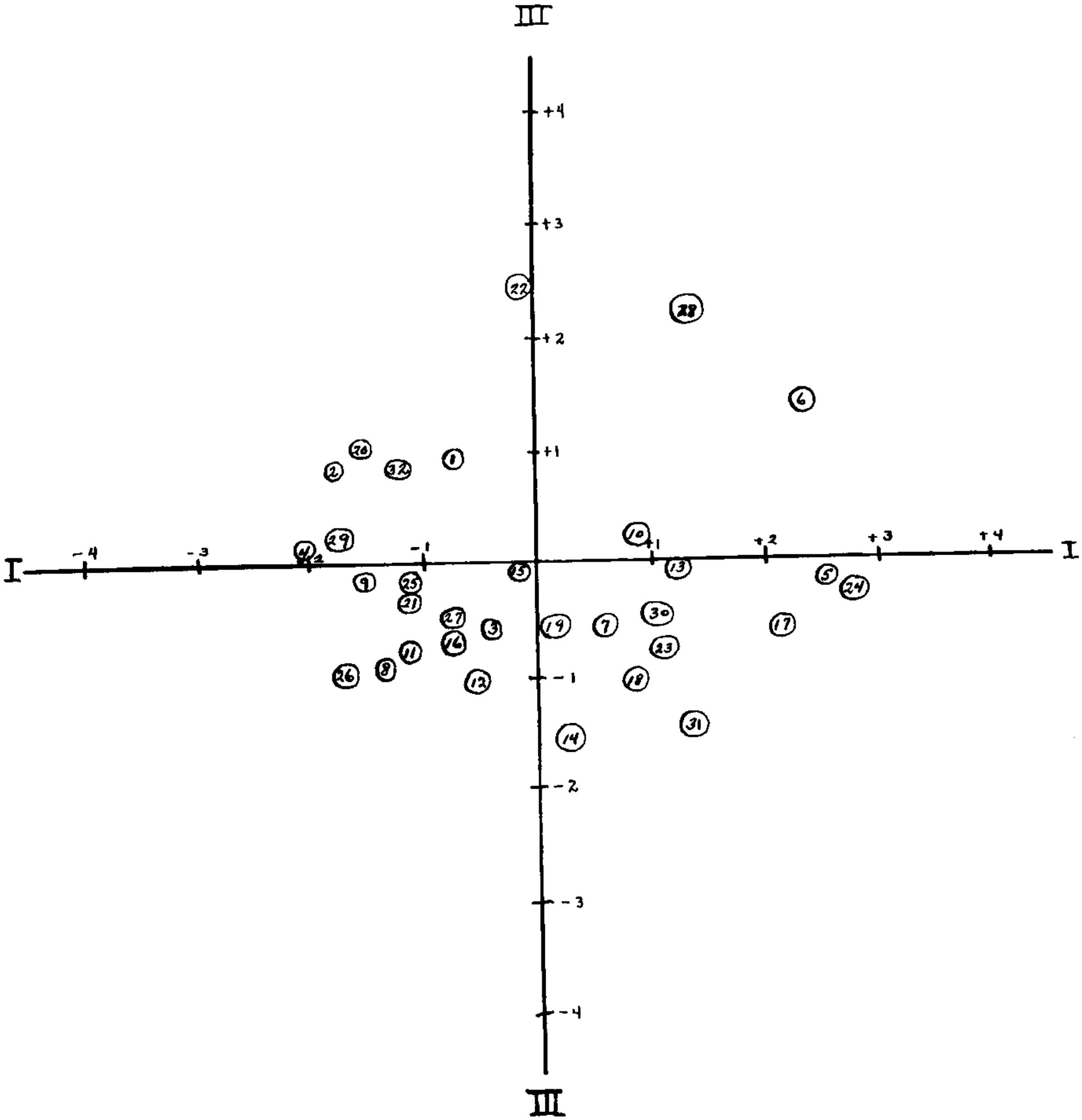

Figure 4. Factor I versus Factor III for 32 Complex Sounds

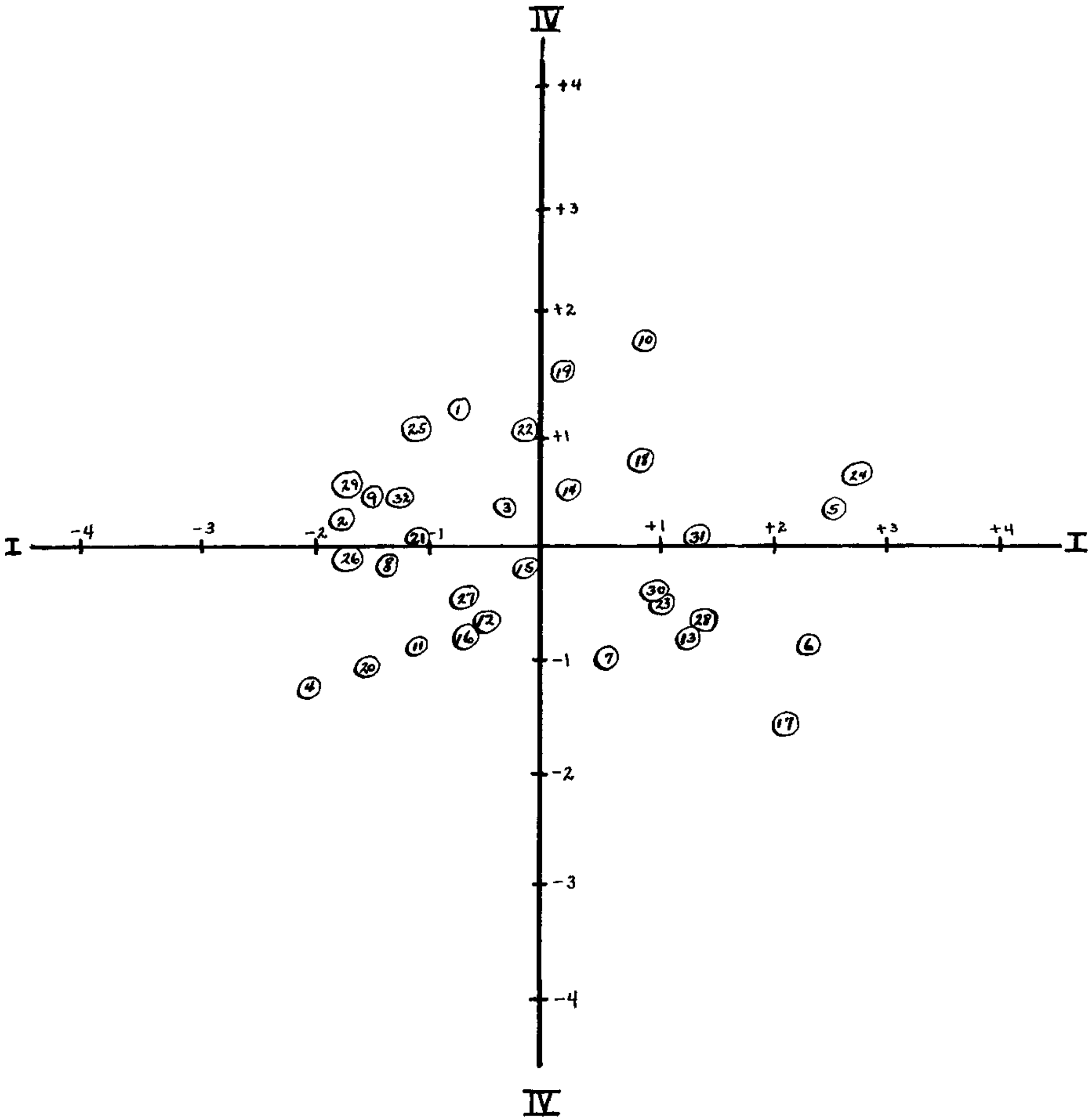

Figure 5. Factor I versus Factor IV for 32 Complex Sounds

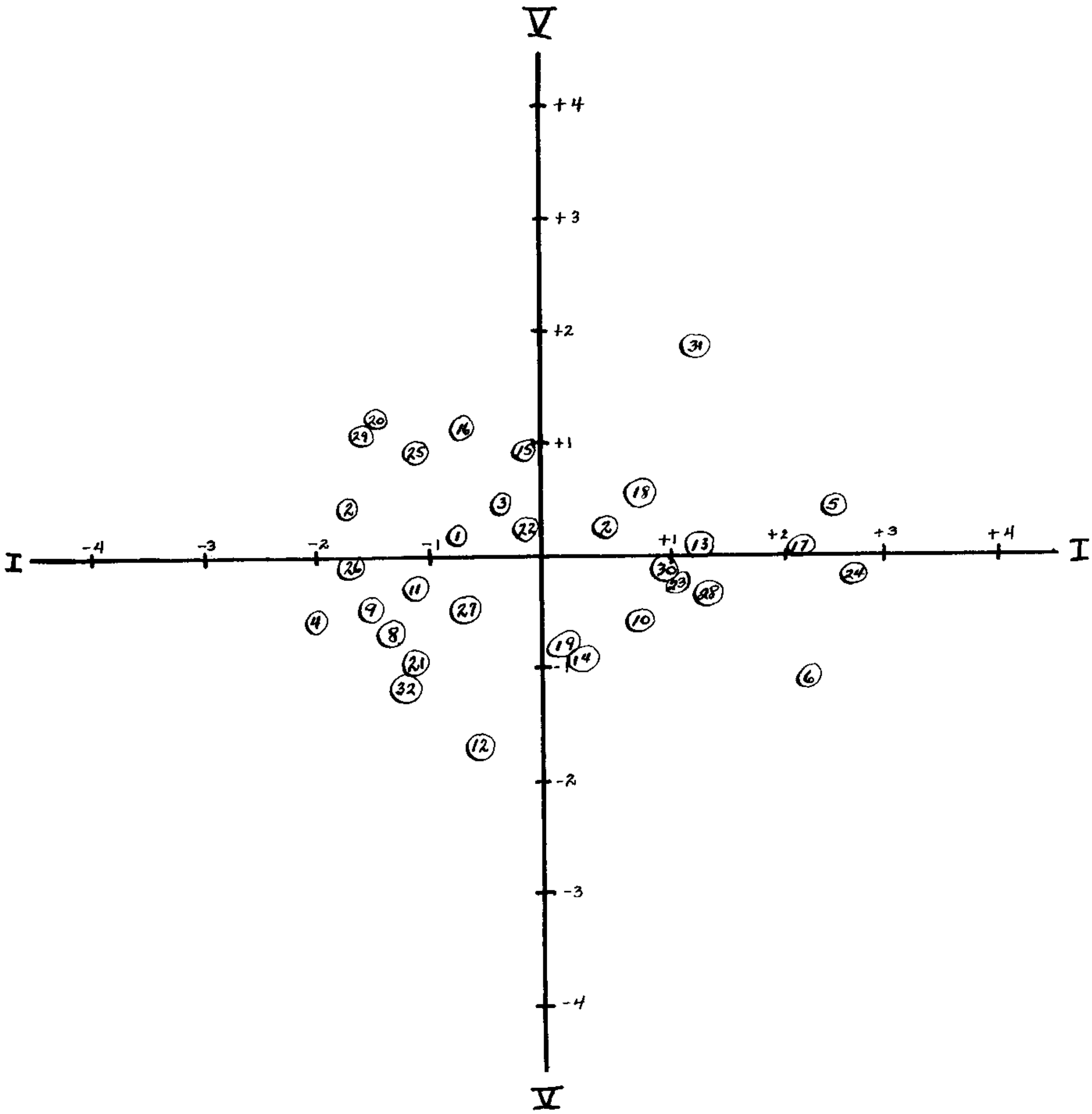

Figure 6. Factor I versus Factor V for 32 Complex Sounds

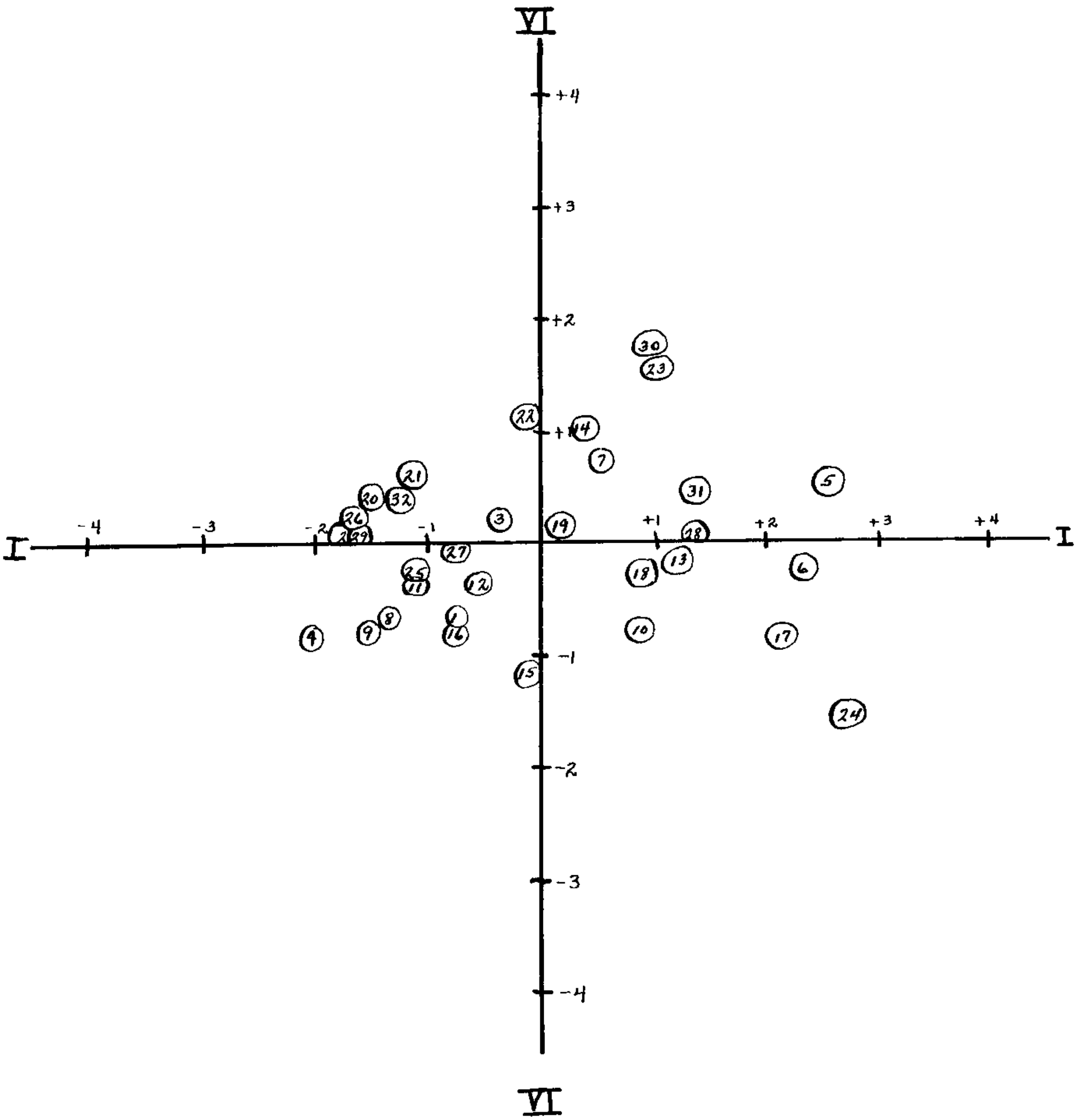

Figure 7. Factor I versus Factor VI for 32 Complex Sounds

The second method used in the attempt to identify and name the multidimensional scaling dimensions was that of presenting a series of sounds to observers and asking the observers to describe the dimension which appeared to be represented by the series of sounds. The sounds of each of the series represented successive points on each of the first six multidimensional scaling factors. The observers were in good agreement that the sounds for factor I described a highness to lowness continuum. This is in accord with relating this factor to unidimensional attributes where it appeared to be a density, brightness, and pitch dimension. Factor II was termed by some observers as a ringing to muffled dimension. Factor III was termed a clarity scale and also involved ringing to muffled aspects. Factor IV appeared to have an aspect of loudness according to the observers and was also described as being on a tight to loose continuum. Factor V sounds were noted as describing a progression from a poorly defined tone to a well defined tone, or, in the language of the observers, as having a forced or an unforced quality. The variability of a sound, smoothness, and dullness were responses given by observers to sounds which represented a continuum for factor VI. It was not felt that this attempt to obtain names for dimensions was successful because for the most part, the observers, with exception of factor I, did not agree in their descriptions of a dimension. The observers' task of naming the dimensions was probably confounded by the language which they have available for describing sounds. People tend to generalize when they label sounds as soft-loud, high-low, and bright-dull, and, since each of the sounds varied in several dimensions, the task of isolating a single dimension and naming it precisely was difficult. The results of this study would seem to indicate that for this set of complex sounds between three and six psychological dimensions were present. In dealing with complex sounds it is difficult to know whether or not all possible dimensions are present with adequate representation of sounds on each dimension. More conclusive results than those presented here could only result from a series of multidimensional studies in which various groups of complex sounds were scaled. The likelihood of success of this approach would depend upon the adequacy of the sounds used to represent all possible dimensions.

Multidimensional Scaling of Speech Sounds

The multidimensional scaling method, as outlined in the two preceding studies, was applied to another set of complex sounds. The stimuli were 32 voice produced sounds of the English language. The sounds were produced by an individual trained in phonetics who practiced each sound until the sound could be continued for at least five seconds with level and quality of each sound held relatively constant. The sounds used were:

a æ ɔ o ʊ u ʌ r l w i j
ɪ e ɛ s f θ z ʒ v ɣ m n
ŋ ʃ t d k g p b

The 32 speech sounds were programmed in all possible pair combinations, 496 pairs. The duration of each sound was 1.5 seconds with a rise-fall time of 100 milliseconds. An interval of 0.5 seconds separated the two members of each pair and each pair of sounds was preceded by an identifying carrier number. The interval between the onset of each carrier number was approximately five seconds. The sounds were played in reverse direction from the master tape which contained the speaker's original production to the tape on which the stimulus pairs were programmed which meant that the sounds heard by the observers were backward presentations of the sounds. The order of presentation of the stimulus pairs was randomly determined.

The observers judged each pair of sounds for similarity-dissimilarity on a nine-point scale. Practice stimuli preceded the test sounds to acquaint the subjects with the task and the sounds which were to be judged for similarity. One hundred subjects served as observers and heard the sounds through earphones.

Analysis of the data resulting from the observers' judgments of similarity of stimulus pairs was completed on the IBM 1103A computer. The analysis method was the same as described in the two preceding multidimensional studies. The input to the computer was the 496 x 9 matrix in which the cell values indicated the frequency with which each of the scale categories, representing nine degrees of similarity, were selected by observers for each of the 496 stimulus pairs. Comparative distances between stimuli were obtained using the categorical judgment model of successive intervals. The B* matrix was determined using an estimated additive constant, c = 6.3700641, which was the largest element in the comparative distance between stimuli matrix. The B* matrix was solved for eigen values and eigen vectors. The eigen values were: 339.95, 110.93, 70.46, 67.27, 65.29, 62.93, 62.49, 58.71, 54.23, 52.41, 50.99, 49.83, 46.86, 45.54, 44.55, 43.36, 41.96, 40.94, 40.25, 38.71, 38.11, 35.30, 33.84, 33.62, 30.86, 29.76, 26.05, 22.24, 20.26, 19.02, 14.70, and .00009.91. Thirty eigen values and their corresponding eigen vectors were retained and the analysis was continued through four successive B* matrix determinations and corresponding approximations to the additive constant. The final approximation to the constant was, c = 3.1209517. The final set of eigen values were: 223.10, 63.31, 36.39, 34.13, 33.49, 31.46, 31.34, 28.88, 26.10, 24.70, 23.68, 22.76, 20.70, 19.97, 19.22, 18.45, 17.69, 16.76, 16.33, 15.40, 14.98, 12.85, 11.93, 11.54, 9.95, 9.49, 5.99, 4.16, 2.17, and 1.81. The final B* matrix was factored to obtained the final matrix of projections. The adequacy of the factorization was evaluated by producing an additional B* matrix from the factor matrix, B* = FF', and by comparing the sum of the squares of the B* matrix elements to the sum of the squares of the factor matrix. The factorization was exceptionally good as may be noted by the two values, 66116.104 and 66116.107.

The matrix of projections, for 16 factors, for the 32 speech sounds is given in Table XIV. In Figures 8-12 the projections for the first six factors are shown graphically. Inspection of the eigen values where a considerable decrease in values may be noted, 223.10, 63.31, 36.39, and 34.13, and of the factor loadings for factors I and II as compared to the loadings for successive factors, would seem to indicate that a good part of the variance for the B* matrix could be accounted for by the first two factors. Factor I, see Table XIV and Figure 8, is defined by three groupings of sounds, vowels, voiced consonants, and unvoiced consonants. The loadings for factor II indicates primarily a consonant dimension with the voiced consonants differentiated from the vowels and the unvoiced consonants. Factor III is a vowel dimension and in terms of sound production appears to represent a continuum from least to greatest opening of the oral cavity. Although factor loadings continue to be relatively high for some sounds through factor VI and beyond, as may be noted in Table XIV, it is questionable whether or not those beyond factor III should be considered as relevant.

Table XIV.

Matrix of Projections for 32 Speech Sounds

Sounds	Factors I	II	III	IV	V	VI	VII	VIII
a	-1.84998	-0.26764	2.14862	-0.37746	0.28796	-0.18143	-1.51631	1.08647
æ	-1.67710	-0.37799	1.57955	0.03809	0.26003	-0.59066	1.31622	0.76112
ɔ	-1.97834	-0.77495	1.56690	-1.15295	-0.58887	1.40505	-1.54274	0.41487
o	-1.90476	-0.56088	0.79416	-1.22956	-0.91513	0.03300	-0.87931	-0.94794
ʊ	-1.75452	-0.45293	0.40114	-0.11856	-0.82372	-0.21402	-1.79638	-1.46699
u	-1.75825	-0.69293	-1.98508	-0.01343	0.81473	-0.40946	-0.17815	-0.86551
ʌ	-1.70790	-1.43638	-0.01386	0.54114	-0.73907	-0.27596	-0.29193	2.84017
r	-1.99258	-0.82567	-0.63480	0.07407	0.10598	-2.11043	1.17645	0.10081
l	-1.89557	-0.86962	-0.66985	-0.29794	0.09128	-1.34524	-1.72274	-0.26095
w	-2.01519	-0.50203	-1.41287	-0.87092	0.12109	-0.74194	0.93040	-0.32862
i	-1.75054	-0.67943	-0.86378	1.26071	1.58898	1.95946	-0.04865	0.90561
j	-1.90410	-0.45961	-1.12162	0.27151	1.07528	2.84884	0.79881	1.17496
ɪ	-2.00976	-0.38893	1.54188	1.96834	0.47755	0.17977	0.25360	-1.33608
e	-1.87696	-0.74769	1.35228	1.43959	-0.17849	-0.38171	1.19447	-0.96203
ɛ	-1.65914	-1.09544	1.27045	1.49307	-0.94171	-0.15211	2.06640	-0.25635
s	5.19551	0.36682	-0.38802	1.61774	0.46256	-0.58960	0.27670	0.10032
f	4.98315	0.49192	0.62868	0.98489	-1.71817	1.30707	-0.69624	-0.31330
θ	5.06143	0.23182	-0.88197	1.57450	-1.14830	-0.65574	-0.60425	-0.31963
z	1.45748	2.69571	-0.48985	0.40620	1.70836	-0.01571	-0.80939	-0.14746
ʒ	0.21253	2.28618	1.04086	0.25206	1.80837	-1.41012	-1.00933	0.11321
v	0.42750	3.42575	-0.47191	0.19403	-0.73697	-0.15739	0.25459	1.24189
ɣ	-0.61402	3.24632	0.23529	0.05884	-1.59245	-0.21285	0.29196	0.79118
m	-1.75672	-0.34698	-1.26508	-0.80334	-0.40160	0.11348	-0.33681	-0.77372
n	-1.73728	-0.37625	-1.31937	0.17888	-1.14218	0.51891	-0.57792	-0.16118
ŋ	-1.86064	-0.53899	-1.87634	1.12233	0.33747	-0.37256	-0.82957	-0.26892
ʃ	3.92187	-1.58496	0.83693	-1.26195	2.19564	-0.78207	0.22312	0.59253
t	4.45980	-1.10676	0.60660	-0.89177	1.54213	0.31176	0.43735	-0.82697
d	-0.34824	2.42646	0.37487	-1.32279	0.79058	0.88650	0.12525	-0.32465
k	4.54685	-1.94302	-0.58792	-1.52721	-1.21249	-1.01068	0.22166	1.70435
g	-0.91348	0.89695	-0.16620	-1.28367	-0.47412	2.82291	1.38933	0.06458
p	4.98796	-1.63182	-0.01522	-0.43583	-0.52330	1.48394	0.13664	-1.03135
b	-0.28920	1.59486	-0.21515	-1.88906	-0.53236	0.28005	1.74724	-1.30034

Table XIV (cont.)

Matrix of Projections for 32 Speech Sounds

Sounds	Factors IX	X	XI	XII	XIII	XIV	XV	XVI
ɑ	0.23469	-0.26672	1.68759	-0.22926	0.32673	-1.39042	-0.60086	1.67955
æ	-0.83327	1.96213	1.04871	1.90786	0.13440	1.42235	-0.62076	-1.31632
ɔ	-0.50787	0.22231	0.74183	0.12616	-1.46495	-0.83940	0.16882	-0.75957
o	-0.73208	-0.97504	-0.99852	0.29954	-0.36293	1.97648	-0.61085	0.16294
ʊ	-1.14654	0.83890	-1.16843	0.47742	0.54598	-0.02209	0.62182	1.79236
u	0.36076	0.44443	0.86387	-0.74663	0.41463	-1.20557	0.29324	-0.66778
ʌ	-0.96571	-1.46155	-0.38497	-0.61687	1.26221	-0.51984	0.69692	-1.04536
r	-0.79333	-0.47858	-1.13902	-0.43143	-1.92833	-0.58657	-1.68900	0.54762
l	0.38146	1.74607	-0.36887	-1.59938	0.68154	0.80590	0.42898	-1.02772
w	-0.09550	-0.69545	0.97830	0.08641	-1.17900	-0.14810	1.77048	-0.00971
i	-0.31902	0.79106	-0.90410	-1.10777	-0.24988	1.16752	0.32816	0.76690
j	0.01327	0.09465	-0.56610	0.46442	-0.00772	0.12746	-0.36653	0.69663
ɪ	1.45144	-0.88883	-0.11029	-0.93598	0.08114	0.64299	-1.11540	-0.69098
e	0.41224	0.38512	-1.18021	0.23817	0.55038	-1.65216	-0.37613	-0.24045
ɛ	-0.17891	0.57586	0.90359	-0.57961	0.12839	0.03814	1.67295	0.62347
s	-1.61984	-0.75207	0.09464	0.92704	0.14707	0.18852	-0.20877	0.11639
f	1.66184	-0.23328	0.08171	-0.03650	-0.27842	0.09631	0.28342	-0.34404
θ	-0.72142	-0.26404	0.15073	0.35678	0.62899	-0.04792	0.07368	0.14547
z	0.08228	2.01818	-0.02994	0.64046	0.15049	-0.98943	-0.52575	-0.36429
ʒ	0.39002	-0.78656	-0.71034	0.29958	-0.18322	0.71916	1.68017	0.15154
v	0.35677	0.34991	0.44544	-1.05496	-0.67042	0.13151	-0.68230	1.00828
ð	-0.00671	-0.07629	-0.04834	-1.23440	-0.73642	0.21887	-0.05936	-0.52754
m	0.03111	-0.77334	0.76074	-0.65818	1.53767	-0.02917	-1.44590	-0.13442
n	2.12985	-0.28774	-0.26607	1.77866	-1.07503	-0.03067	0.19623	-0.64368
ŋ	-0.15856	-0.80266	1.32411	1.47346	-0.33690	0.38856	-0.16982	0.24794
ʃ	1.44553	-0.58082	-1.26966	0.11644	-0.44370	-0.50477	0.24100	-0.04585
t	0.18600	-0.60814	2.04240	-1.07708	0.08231	0.99023	-0.32266	0.35105
d	-1.40552	-1.21319	-0.44772	0.22570	0.61550	-0.43411	-0.03143	-1.39852
k	0.79953	1.17645	-0.32989	0.10047	0.29706	0.42922	-0.45841	0.33118
g	1.15175	-0.43951	-0.72947	0.96506	1.98082	0.07531	0.18872	0.67427
p	-1.46430	0.51784	-0.62062	-0.50307	-0.70624	-0.85721	0.17211	-0.50837
b	-0.13971	0.46210	0.14900	-0.13105	0.05875	-0.16097	0.46838	0.42740

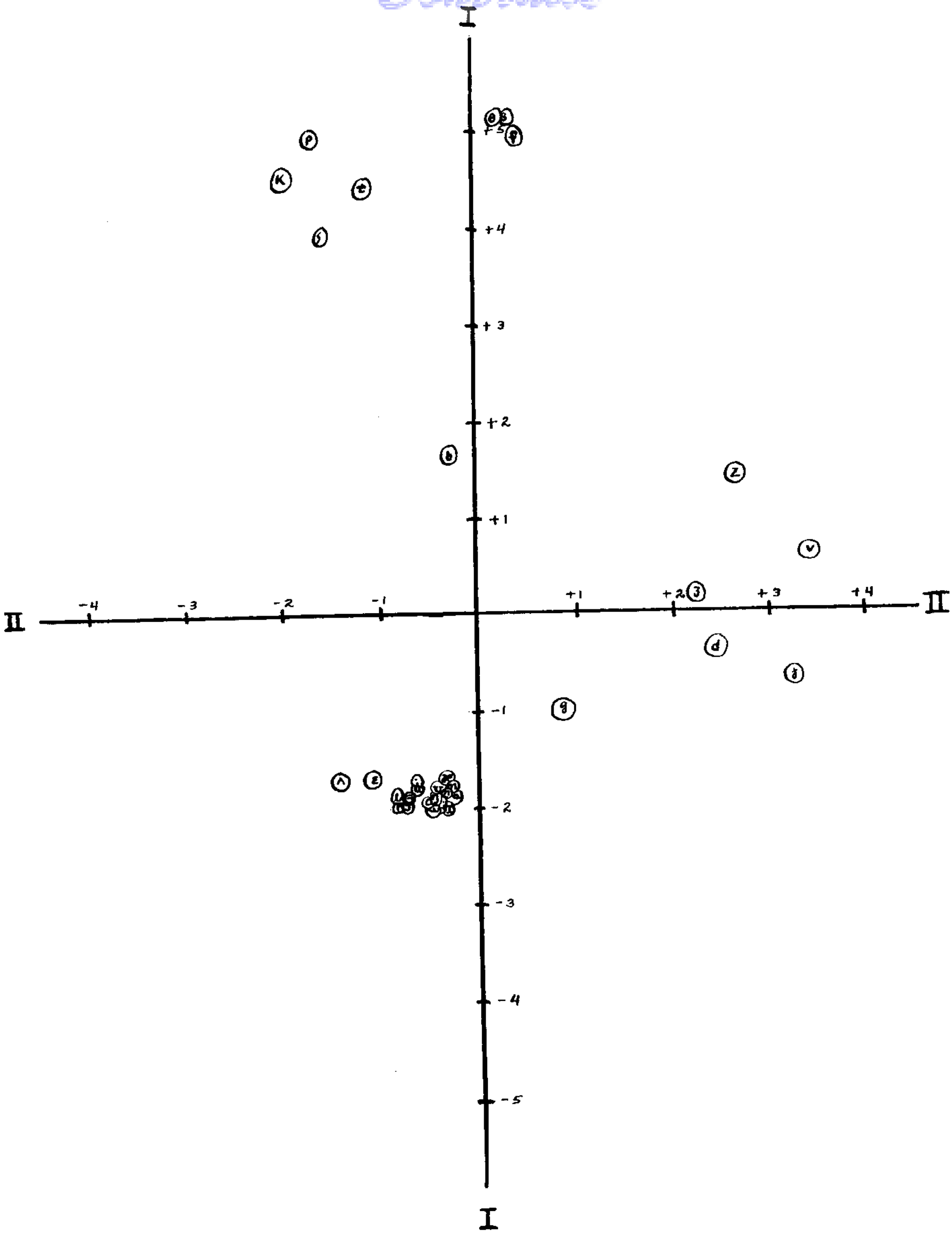

Figure 8. Factor I versus Factor II for 32 Speech Sounds

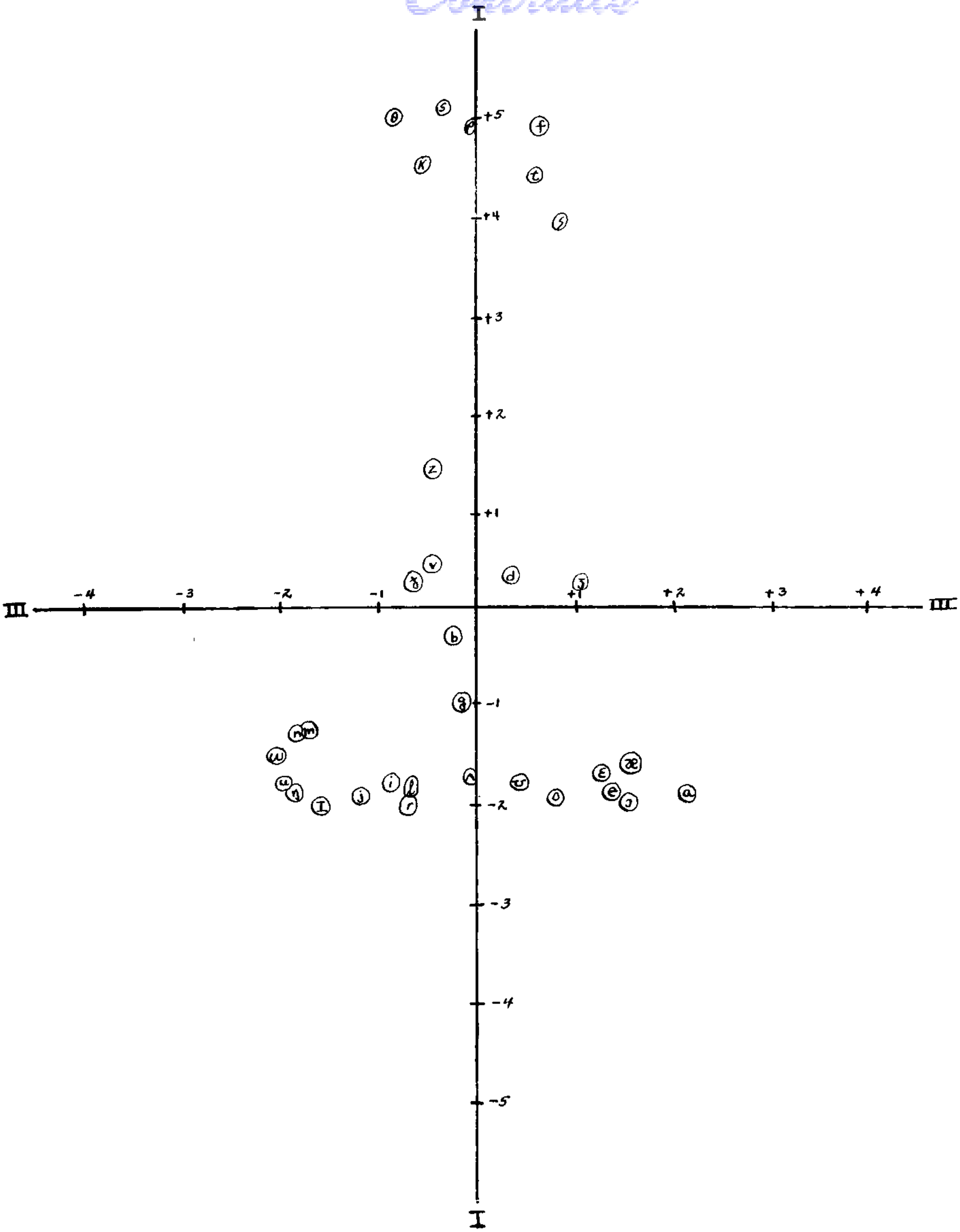

Figure 9. Factor I versus Factor III for 32 Speech Sounds

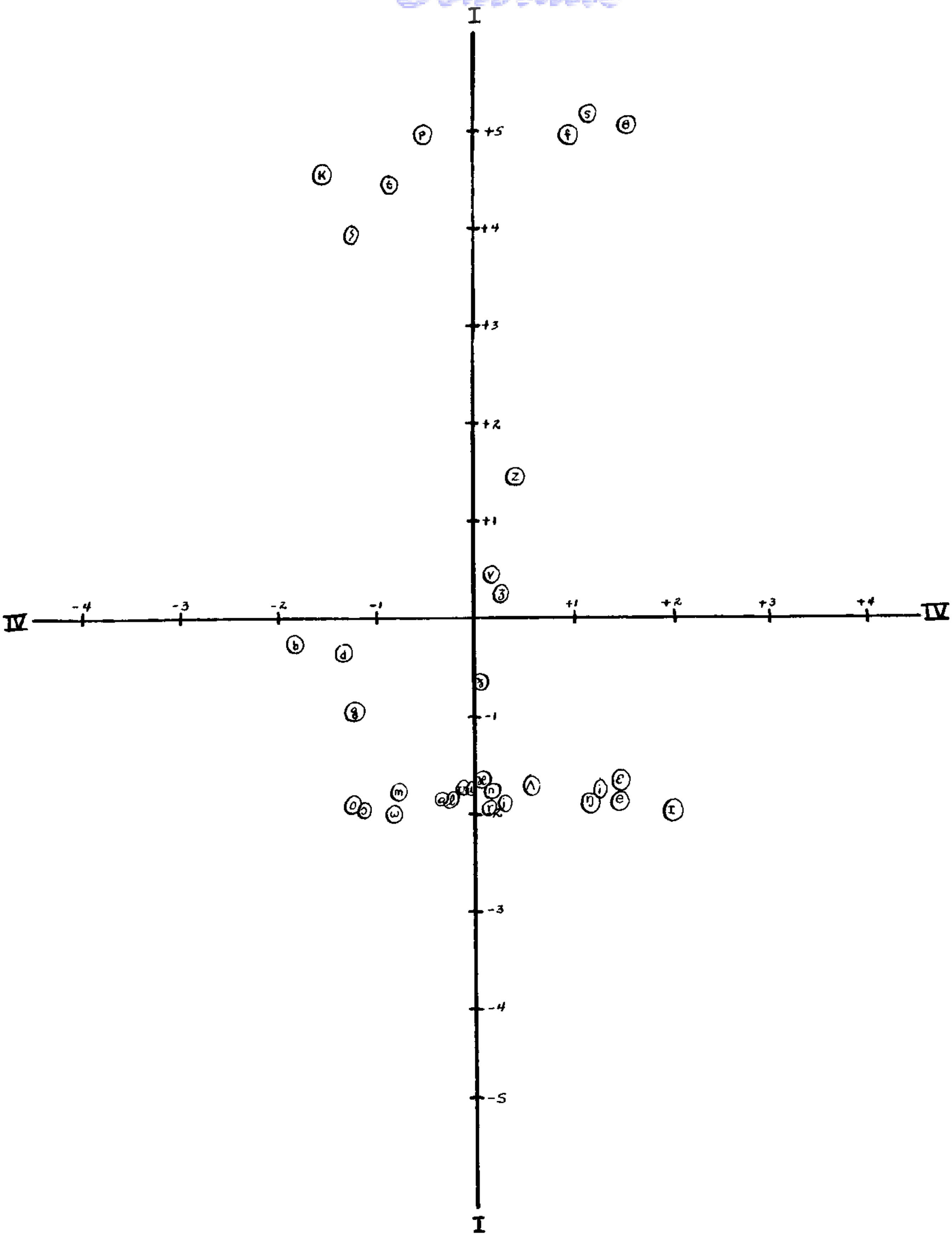

Figure 10. Factor I versus Factor IV for 32 Speech Sounds

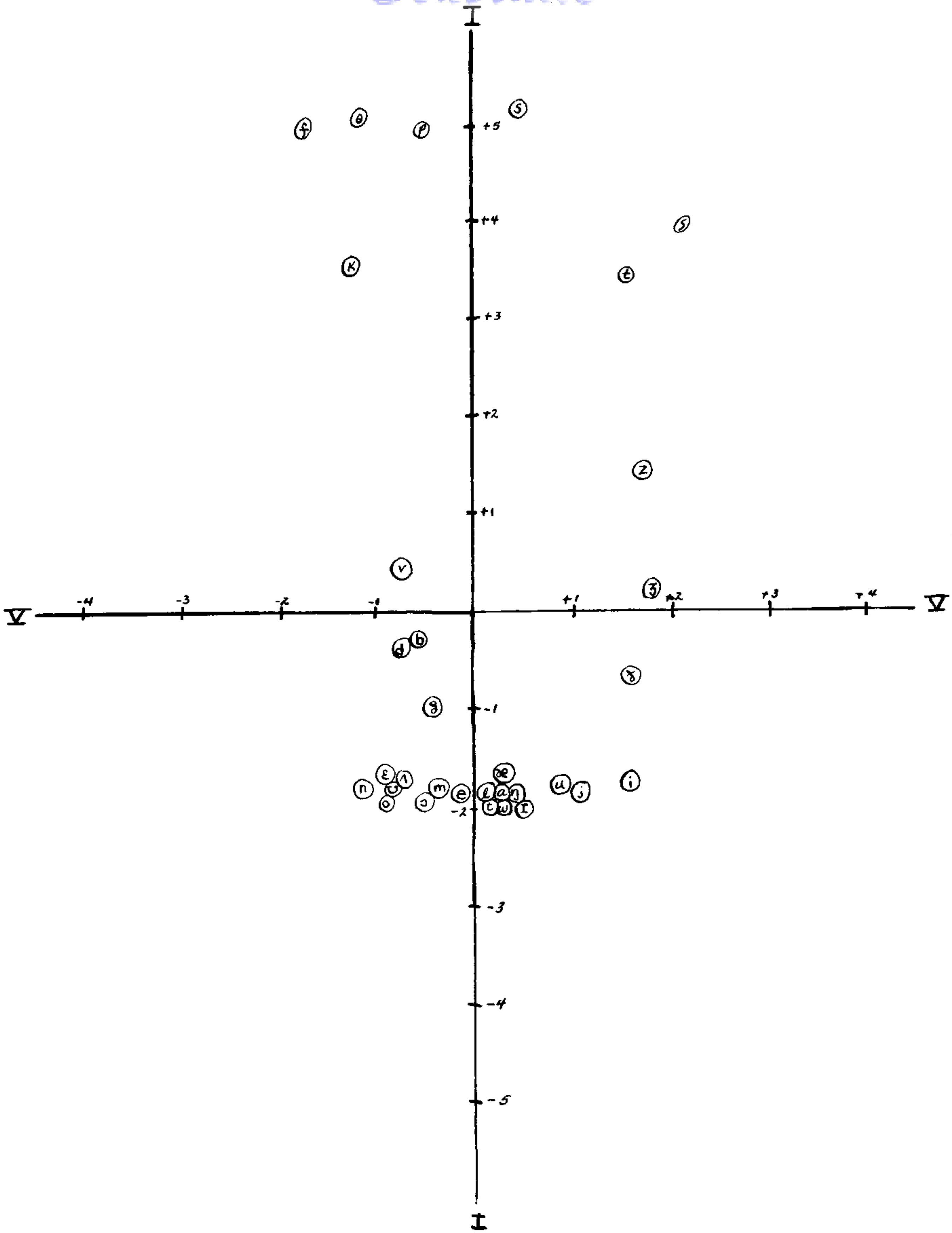

Figure 11. Factor I versus Factor V for 32 Speech Sounds

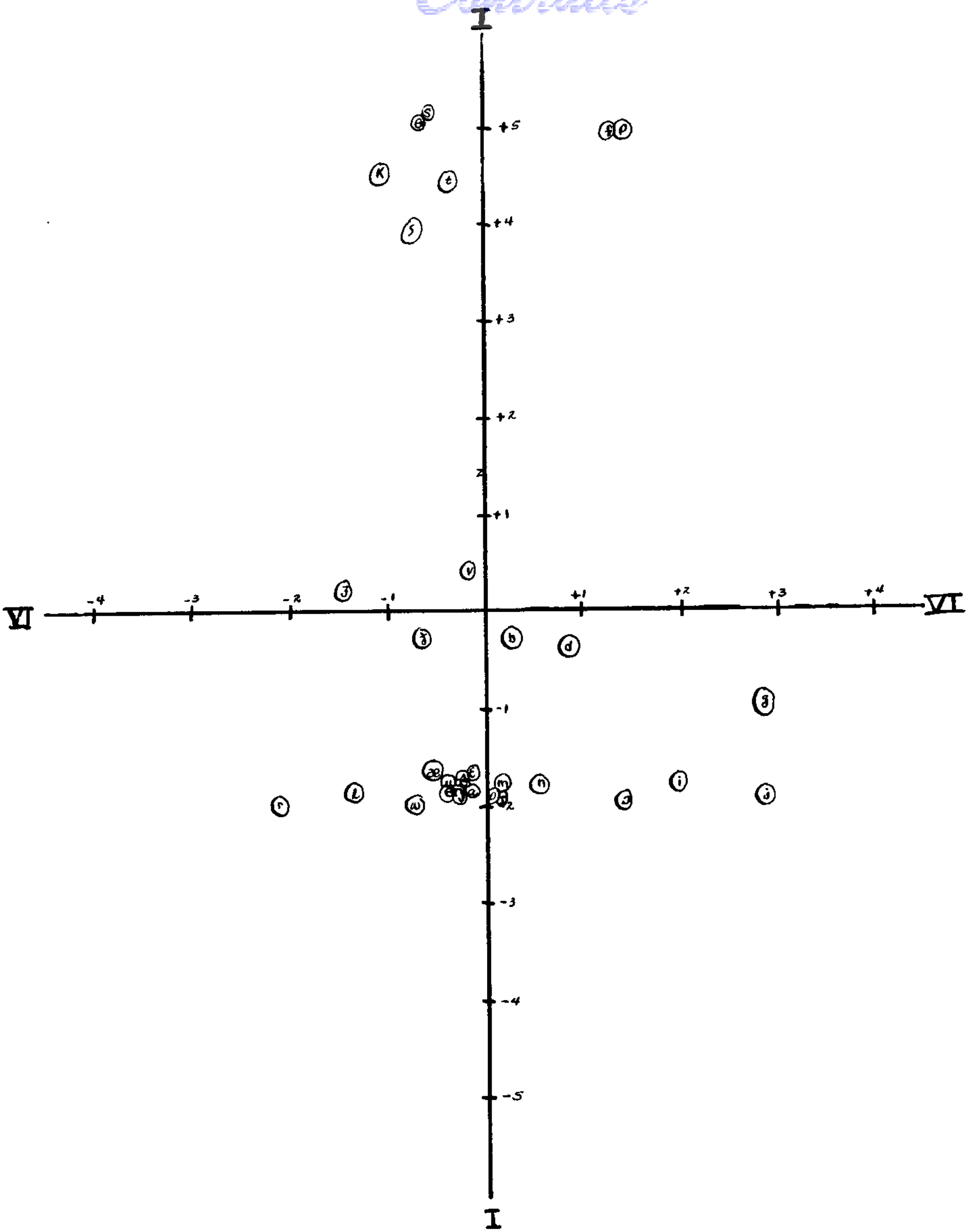

Figure 12. Factor I versus Factor VI for 32 Speech Sounds

Multidimensional Scaling of Visual and Auditory Stimuli

The purpose of this experiment was to compare the results of multidimensional scaling for the same stimuli which were separately seen and separately heard by the observers. The stimuli were heard by the observers through earphones and, for the visual portion of the experiment, viewed on an oscillograph. The stimuli consisted of nine multitones. The stimuli were programmed in all pair combinations which resulted in 36 pairs. The duration of each stimulus was two seconds with a brief interval separating the members of each pair. Each pair was preceded by an identifying carrier number. There was an interval of 10 seconds between the onset of each carrier number. The pairs of stimuli were presented in random order.

Twenty individuals judged the stimulus pairs for similarity-dissimilarity on a nine-point scale. Each observer made three judgments of similarity for the stimulus pairs for both the visual and auditory presentations. In an attempt to minimize order effect, presentation of the two conditions was systematically varied. The first judgment by each observer of each presentation, visual and auditory, was practiced to acquaint the subject with the stimuli and the experimental task.

Comparative distances between stimuli were determined for both sets of data, visual and auditory, using the multidimensional scaling procedure of successive intervals. B* matrices were determined and were factored by the centroid method (74). Four factors were extracted for each B* matrix. The factor matrices for the auditory and visual presentations are shown in Table XV. The projections of the stimuli for the four factors are shown in Figures 13-15. The circled numbers represent the auditory stimuli and the boxed numbers, the visual stimuli. Although the visual and the auditory structure in the space defined by the four dimensions for the nine stimuli was similar, there was not complete accord between the two stimuli sets. The agreement between the two sets of stimuli was most pronounced for factor I and in least agreement for factor II. The observers were questioned after each experimental session as to their bases for judging the visual stimuli for similarity-dissimilarity. Their responses included the relative height of the stimuli, spacing between successive oscillograph tracings, regularity or irregularity of the tracings, and rate of change in the pattern of the tracings. If the structure for the two sets of stimuli had been in close accord, indicating that much of the auditory information was also present in the oscillograph time-amplitude presentation of the stimuli, it would have been possible to make inferences relating physical parameters to psychological parameters. The correspondence between the two sets of stimuli appeared to relate loudness and pitch to the oscillograph display of amplitude and frequency. The interest in this study was on those psychological dimensions of complex sounds other than loudness and pitch and in this sense was not successful in relating physical to psychological dimensions.

CHANGES IN AUDITORY PERCEPTION UNDER CONSTANT STIMULATION

There is a considerable amount of evidence which indicates that sounds which are heard by an observer for a period of time changes in loudness and in pitch. It is conceivable that sounds undergo other perceptual changes and that these changes may relate to the dimensions of auditory experience. In consideration of this possibility a series of studies were conducted which involved observers listening to a constant sound for a period of time, either 15 or 30 minutes, and reporting on

Table XV.

Matrix of Projections for Nine Complex Sounds

Auditory Presentation

Item	I	Factors II	III	IV
1	-1.8787	2.8160	-0.0933	-0.1574
2	2.6055	-0.3937	0.3518	0.2468
3	-0.6038	0.3992	3.2713	-1.2227
4	-0.0061	-1.5527	0.6373	2.2611
5	-2.4086	-0.8590	-0.0325	0.8290
6	2.7662	1.6417	-1.0000	0.6336
7	4.2227	-0.6691	-1.2338	-1.8616
8	-2.2085	-0.4284	-1.3833	-0.6375
9	-2.5329	-0.8088	-0.3133	-1.1719

Visual Presentation

Item	I	Factors II	III	IV
1	-2.7415	-1.9451	-1.9021	-0.1569
2	0.1010	2.1926	0.2599	2.6961
3	-0.2192	2.5869	-0.1229	2.5743
4	0.4673	4.2565	0.8796	1.1481
5	-0.6030	-0.8278	-1.9003	-0.4094
6	4.5820	-0.3283	1.3942	-0.9092
7	4.1410	-1.6806	2.1153	-0.4348
8	-0.5381	-0.9632	-1.8489	-0.3256
9	-2.3446	-0.3510	-1.6972	-0.4263

the changes which occurred with respect to the sound. The sounds were pure tones, bands of white noise, and multitones. The sounds were presented at various sound pressure levels. When questioned after the period of listening about the changes which had occurred in a sound, the observers consistently reported changes in loudness, in pitch, in the number of components present, in the relationship of the fragmenting components, and in the smoothness or roughness of the sound. Attempts were made to formulate a predictive time model which related to these changes. These attempts were not successful because, when the subjects were required to write, or report verbally while listening, the perceptive experience was altered. Responses under these conditions were inconsistent between subjects and for the same subject on successive listening trials.

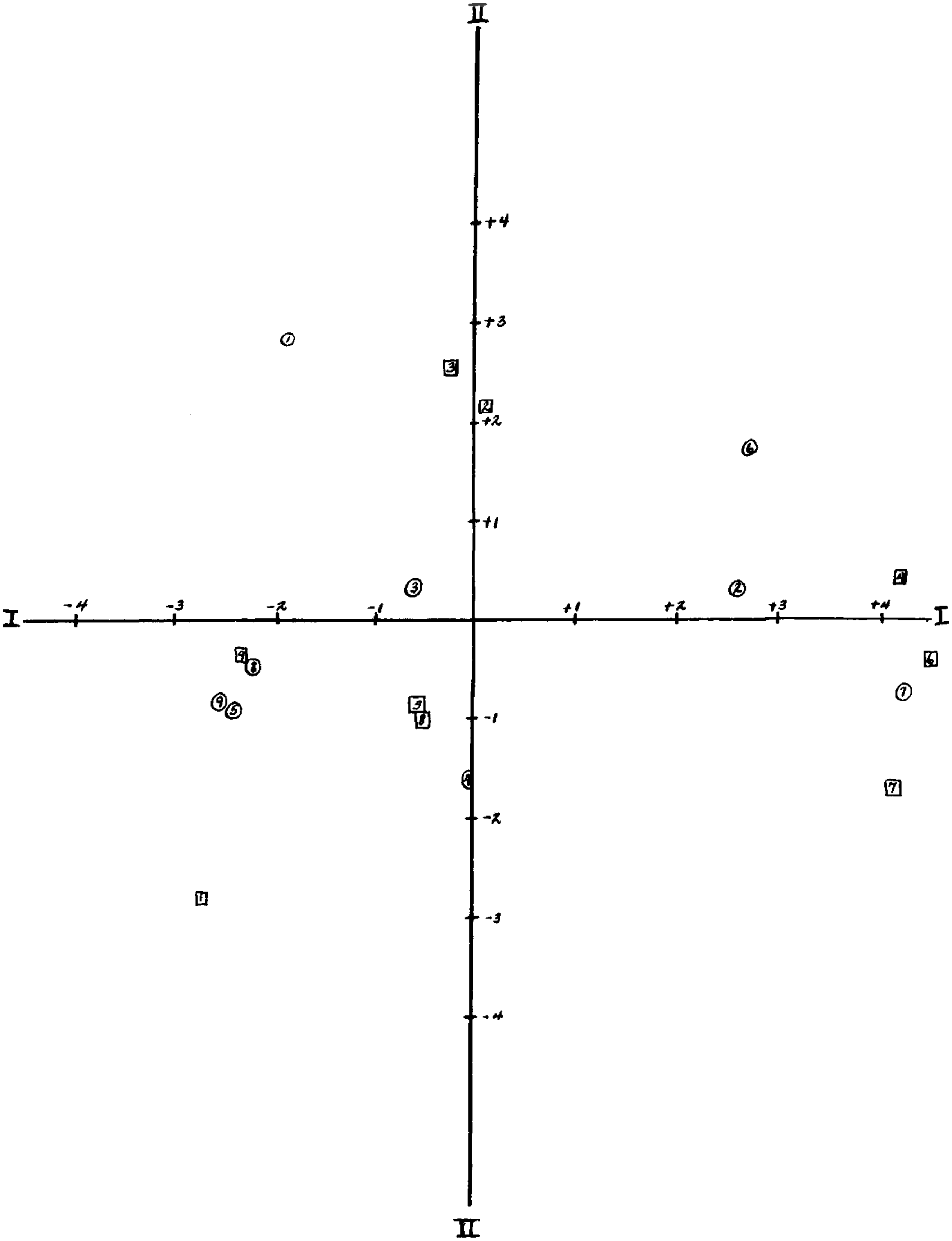

Figure 13. Factor I versus Factor II for 9 Complex Sounds

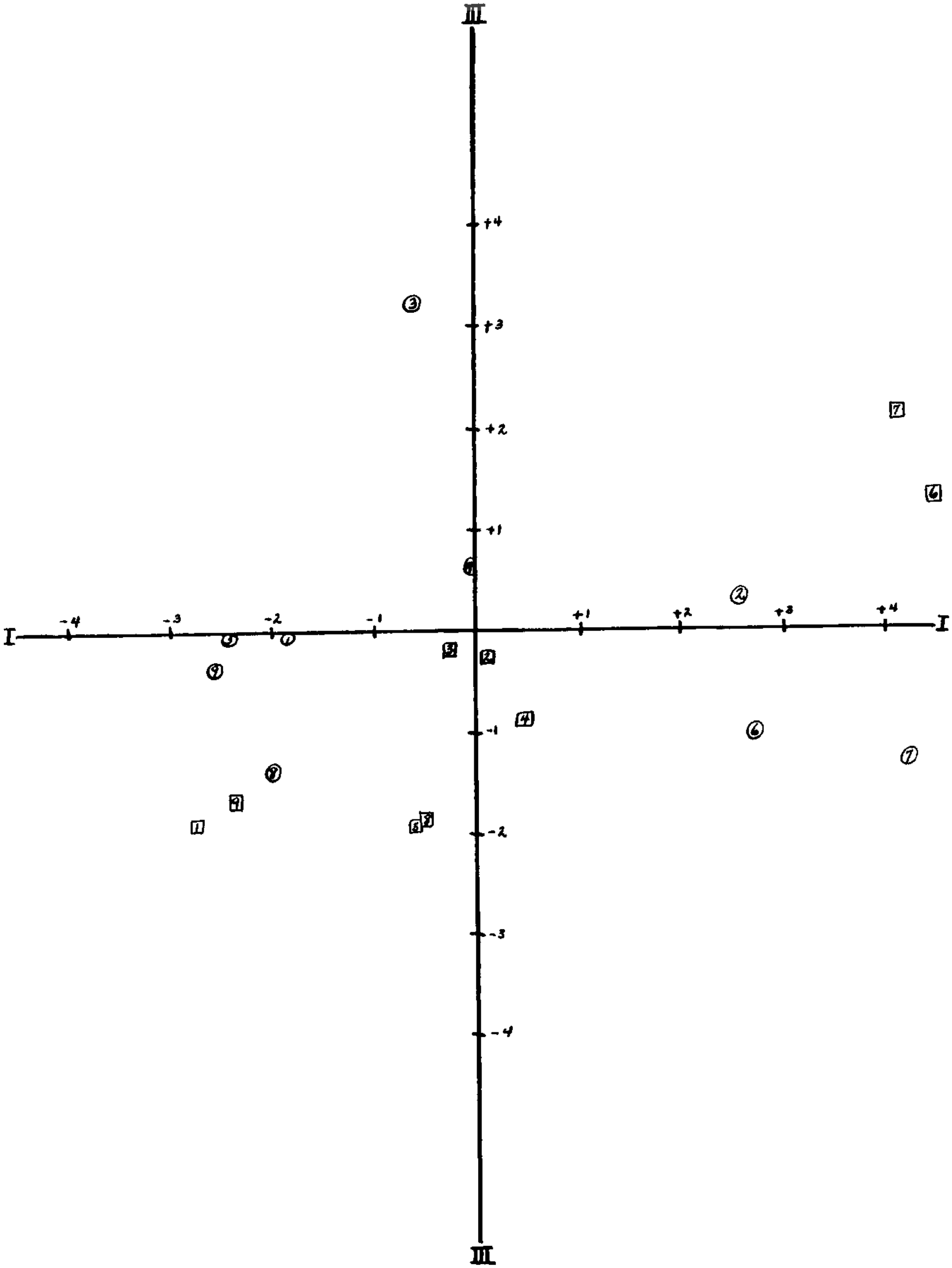

Figure 14. Factor I versus Factor III for 9 Complex Sounds

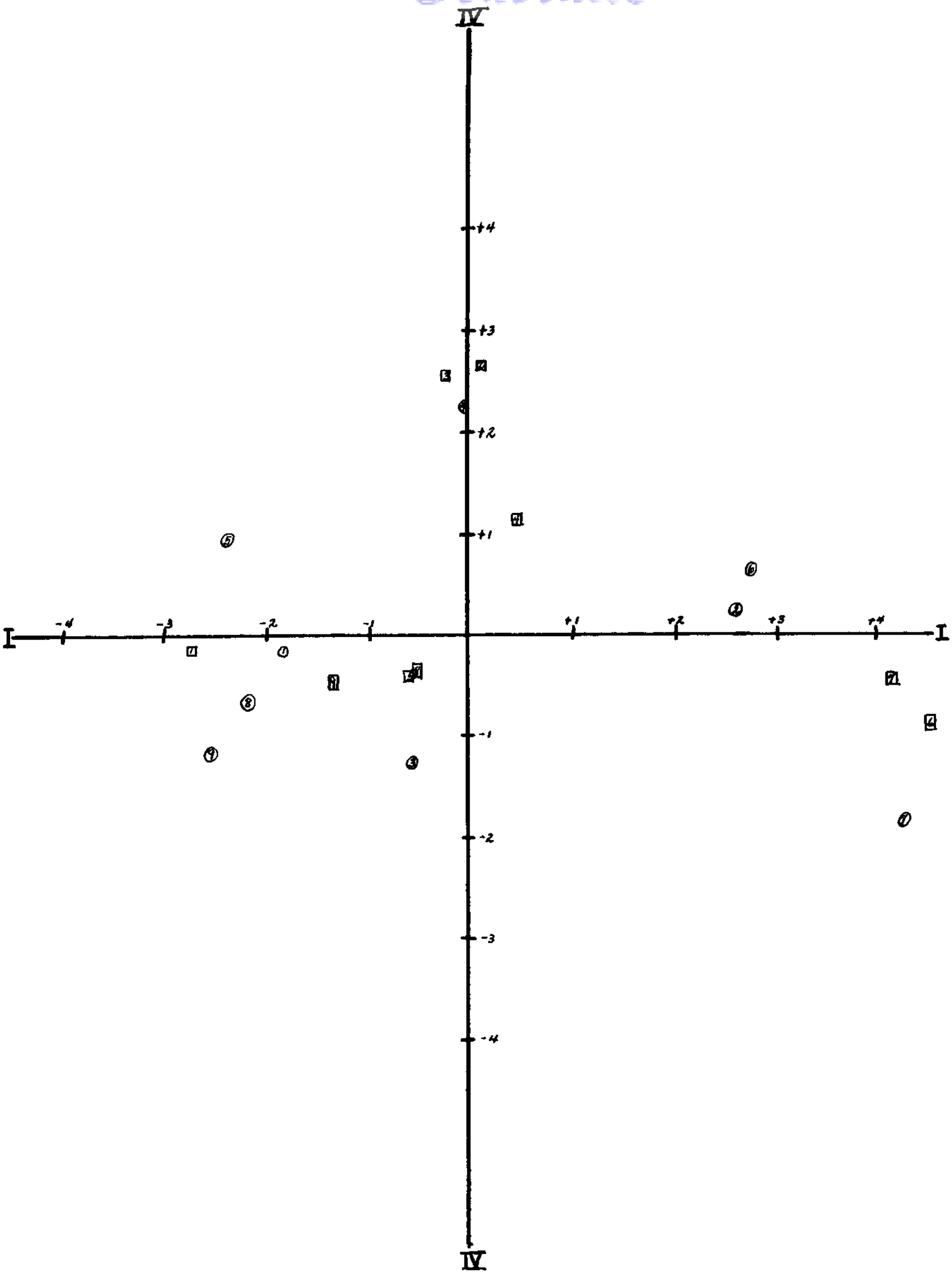

Figure 15. Factor I versus Factor IV for 9 Complex Sounds

SUMMARY AND CONCLUSIONS

The purpose of the program of research summarized in this report was to investigate and evaluate experimental approaches to the determination of psychological parameters for complex sounds. Approaches to the problem were discussed with reference to measurement model, observables in human behavior, and auditory stimulus. Measurement models discussed included: unidimensional scaling with reference to scaling of loudness, pitch, volume, brightness, and density; multidimensional scaling; information theory; and information processing. With respect to observables in human behavior, studies concerned with auditory abilities and loudness and pitch shifts in time were summarized. Speech and music synthesis, and the transformation of auditory stimuli into other sensory modes were discussed in relation to auditory stimulus.

In the attempt to isolate and define psychological parameters of complex sounds, a series of experiments were conducted which involved language of auditory experience, scaling of stimulus sets on equal-interval scales, paired-comparisons scaling, scaling by direct magnitude estimations, observer generated scales, differentiation of auditory patterns, inference of dimensions from measures of channel capacity of observers, evaluating stimulus sets of sounds within a multidimensional scaling model, and changes in auditory perception under constant stimulation.

The most promising approach for the isolation and definition of psychological parameters of complex sounds was the multidimensional scaling model. In a test of the adequacy of the method, pitch and loudness, the expected dimensions, were isolated for pure tones. In the scaling of complex sounds, for the several stimulus sets evaluated in this program of research, between three and six dimensions appeared to emerge. A major problem in evaluating the results of multidimensional scaling of sounds is that of insuring that the set of sounds scaled represent differing values on all possible dimensions. Since this information is not necessarily reflected in the scaling results, the alternate approach is to scale a number of stimuli sets and compare the dimensions which have been obtained in the several evaluations.

REFERENCES

1. Asher, J. E., Hanley, T. D., and Steer, M. D. A factor analysis of twelve physical measures of voice. Purdue U., Lafayette, Indiana. Technical Report Navtradevcen 104-2-48.

2. Attneave, F. Dimensions of similarity. Amer. J. Psychol., 1950, 63, 516-556.

3. Banister, H. Auditory theory: a criticism of Professor Boring's hypothesis. Amer. J. Psychol., 1927, 38, 436-440.

4. Boring, E. G., and Stevens, S. S. The nature of tonal brightness. Procedures National Academy of Science, Washington, 1936, 22, 514-521.

5. Broadbent, D. E. A mechanical model for human attention and immediate memory. Psychol. Rev., 1957, 64, 205-215.

6. Broadbent, D. E. Perception and Communication. New York: Pergamon Press, 1958.

7. Chang, S. H. Speech analysis. Northeastern U., Boston, Massachusetts. Report No. AFCRC TR-58-107. February, 1958.

8. Christman, R. J. Shifts in pitch as a function of prolonged stimulation with pure tones. Amer. J. Psychol., 1954, 67, 484-491.

9. Clark, M., Jr. Proposed keyboard musical instrument. J. acoust. Soc. Amer., 1959, 31, 403-419.

10. Comerci, F. A. Perceptibility of flutter in speech and music. J. Soc. Mot. Pict. Engrs., 1955, 64, 117-122.

11. Davis, H., Silverman, S. R., and McAuliffe, D. R. Some observations on pitch and frequency. J. acoust. Soc. Amer., 1951, 23, 40-41.

12. Delattre, P., Cooper, F. S., Liberman, A. M., and Gerstman, L. J. Speech synthesis as a research technique. Precedings of the VIIth International Congress of Linguists (1952), London, 1956, 545-561.

13. Dix, M. R., Hallpike, C. S., and Hood, J. D. Auditory adaptation in the human subject. Nature, London, 1949, 164, 59-60.

14. Dudley, H. Phonetic pattern recognition vocoder for narrow-band speech transmission. J. acoust. Soc. Amer., 1958, 30, 733-739.

15. Egan, J. P. Perstimulatory fatigue as measured by heterophonic loudness balances. J. acoust. Soc. Amer., 1955, 27, 111-120.

16. Egan, J. P. Message repetition, operating characteristics, and confusion matrices in speech communication. Indiana U., Bloomington, Indiana. Report No. AFCRC-TR-57-50. June, 1957.

17. Fletcher, H. Loudness, pitch and the timbre of musical tones and their relation to the intensity, the frequency and the overtone structure. J. acoust. Soc. Amer., 1934, 6, 59-69.

18. Fletcher, H., and Munson, W. A. Loudness, its definition, measurement and calculation. J. acoust. Soc. Amer., 1933, 5, 82-108.

19. Garner, W. R. Some statistical aspects of half-loudness judgments. J. acoust. Soc. Amer., 1952, 24, 153-157.

20. Garner, W. R. An equal discriminability scale for loudness judgments. J. exp. Psychol., 1952, 43, 232-238.

21. Garner, W. R. An informational analysis of absolute judgments of loudness. J. exp. Psychol., 1953, 46, 373-380.

22. Garner, W. R. A technique and a scale for loudness measurement. J. acoust. Soc. Amer., 1954, 26, 73-88.

23. Garner, W. R. Advantages of the discriminability criterion for a loudness scale. J. acoust. Soc. Amer., 1958, 30, 1005-1012.

24. Garner, W. R. On the lambda loudness function, masking and the loudness of multicomponent tones. J. acoust. Soc. Amer., 1959, 31, 602-607.

25. Guelke, R. W., and Huyssen, R. M. J. Development of apparatus for the analysis of sound by the sense of touch. J. acoust. Soc. Amer., 1959, 31, 799-809.

26. Gundlach, R. Tonal attributes and frequency theories of hearing. J. exp. Psychol., 1929, 12, 187-196.

27. Gundlach, R., and Bentley, M. The dependence of tonal attributes upon phase. Amer. J. Psychol., 1930, 42, 519-543.

28. Hake, H. W., and Garner, W. R. The effect of presenting various numbers of discrete steps on scale reading accuracy. J. exp. Psychol., 1951, 42, 358-366.

29. Halverson, H. M. Tonal volume as a function of intensity. Amer. J. Psychol., 1924, 35, 360-367.

30. Hanley, C. N. Factorial analysis of speech perception. J. speech and hearing Dis., 1956, 21, 76-87.

31. Harris, J. D. The decline of pitch discrimination with time. J. exp. Psychol., 1952, 43, 96-99.

32. Harris, J. D. A search toward the primary auditory abilities. Med. Res. Lab., U. S. Naval Submarine Base, New London, Connecticut. Memorandum Report No. 57-4, BuMed Project NM 22 01 20.2.1. April, 1957.

33. Jeffress, L. A. The pitch of complex tones. Amer. J. Psychol., 1940, 53, 240-250.

34. Jeffress, L. A. Variations in pitch. Amer. J. Psychol., 1944, 57, 63-76.

35. Jerger, J. F. Auditory adaptation. J. acoust. Soc. Amer., 1957, 29, 357-363.

36. Karlin, J. E. A factorial study of auditory function. Psychometrika, 1942, 7, 251-279.

37. Klemmer, E. T., and Frick, F. C. Assimilation of information from dot and matrix patterns. J. exp. Psychol., 1953, 45, 15-19.

38. Klingberg, F. L. Studies in measurement of the relations among sovereign states. Psychometrika, 1941, 6, 335-352.

39. Kohler, I. Auditory guidance studies. Institut fur experimentelle Psychologie der Universitat Innsbruck, Innsbruck, Austria. Report No. AFOSR TR 56-43. July, 1956.

40. Lichte, W. H. Attributes of complex tones. J. exp. Psychol., 1941, 28, 455-480.

41. Lichte, W. H., and Gray, R. F. The influence of overtone structure on the pitch of complex tones. J. exp. Psychol., 1955, 49, 431-436.

42. Licklider, J. C. R. A duplex theory of pitch perception. Experimentia, 1951, 7, 128-134.

43. Messick, S. J. The perception of attitude relationships: a multidimensional scaling approach to the structuring of social attitudes. Educational Testing Service Research Bulletin No. RB-54-27. Princeton, New Jersey, September, 1954.

44. Metfessel, M. Sonance as a form of tonal fusion. Psychol. Rev., 1926, 33, 459-466.

45. Miller, G. A., and Heise, G. A. The trill threshold. J. acoust. Soc. Amer., 1950, 22, 637-638.

46. Miller, G. A., and Nicely, P. E. An analysis of perceptual confusions among some English consonants. J. acoust. Soc. Amer., 1955, 27, 338-352.

47. Morgan, C. T., Garner, W. R., and Galambos, R. Pitch and intensity. J. acoust. Soc. Amer., 1951, 23, 658-663.

48. Munsell Book of Color, Abridged Edition. Baltimore: Munsell Color Co., Inc., 1945.

49. Olson, H. F., and Belar, H. Electronic music synthesizer. J. acoust. Soc. Amer., 1955, 27, 595-612.

50. Pollack, I. Loudness as a discriminable aspect of noise. Amer. J. Psychol., 1949, 62, 285-289.

51. Pollack, I. Auditory flutter. Amer. J. Psychol., 1952, 65, 544-554.

52. Pollack, I. The information of elementary auditory displays. J. acoust. Soc. Amer., 1952, 24, 745-749.

53. Pollack, I. The information of elementary auditory displays. II. J. acoust. Soc. Amer., 1953, 25, 765-769.

54. Potter, R. K., Kopp, G. A., and Green, H. C. Visible Speech. New York: Van Nostrand Co., 1947.

55. Rich, G. J. A study of tonal attributes. Amer. J. Psychol., 1919, 30, 121-164.

56. Richardson, M. W. Multidimensional psychophysics. Psychol. Bull., 1938, 35, 659-660, Abstract.

57. Seashore, C. E. The psychology of musical talent. Boston: Silver, Burdett and Co., 1919.

58. Shepard, R. N. Stimulus and response generalization: tests of a model relating generalization to distance in psychological space. J. exp. Psychol., 1958, 55, 509-523.

59. Sherman, D., and Goodwin, F. Pitch level and nasality. J. speech and hearing Dis., 1954, 19, 423-428.

60. Sherman, D. and Linke, E. The influence of certain vowel types on degree of harsh voice quality. J. speech and hearing Dis., 1952, 17, 401-408.

61. Solomon, L. N. Semantic approach to the perception of complex sounds. J. acoust. Soc. Amer., 1958, 30, 421-425.

62. Solomon, L. N. Search for physical correlates to psychological dimensions of sounds. J. acoust. Soc. Amer., 1959, 31, 492-497.

63. Stevens, S. S. Tonal density. J. exp. Psychol., 1934, 17, 585-592.

64. Stevens, S. S. The volume and intensity of tones. Amer. J. Psychol., 1934, 46, 397-408.

65. Stevens, S. S. The attributes of tones. Procedures National Academy of Science, Washington, 1934, 20, 457-459.

66. Stevens, S. S. A scale for the measurement of a psychological magnitude: loudness. Psychol. Rev., 1936, 43, 405-416.

67. Stevens, S. S. The measurement of loudness. J. acoust. Soc. Amer., 1955, 27, 815-829.

68. Stevens, S. S. Calculation of the loudness of complex noise. J. acoust. Soc. Amer., 1956, 28, 807-832.

69. Stevens, S. S., and Volkmann, J. The relation of pitch to frequency: a revised scale. Amer. J. Psychol., 1940, 53, 329-353.

70. Stevens, S. S., Volkmann, J., and Newman, E. B. A scale for the measurement of the psychological magnitude pitch. J. acoust. Soc. Amer., 1937, 8, 185-190.

71. Tanner, W. P. Theory of recognition. J. acoust. Soc. Amer., 1956, 28, 882-888.

72. Thomas, G. J. Equal-volume judgments of tones. Amer. J. Psychol., 1949, 62, 182-201.

73. Thomas, G. J. Volume and loudness of noise. Amer. J. Psychol., 1952, 65, 588-593.

74. Thurstone, L. L. Multiple-factor analysis. Chicago: The University of Chicago Press, 1947.

75. Thurstone, L. L., and Chave, E. J. The measurement of attitude. Chicago: The University of Chicago Press, 1929.

76. Turnbull, W. W. Pitch discrimination as a function of tonal duration. J. exp. Psychol., 1944, 34, 302-316.

77. Veniar, F. A. Signal detection as a function of frequency ensemble. I. J. acoust. Soc. Amer., 1958, 30, 1020-1024.

78. Veniar, F. A. Signal detection as a function of frequency ensemble. II. J. acoust. Soc. Amer., 1958, 30, 1075-1078.

79. Veniar, F. A. Effect of auditory cue on discrimination of auditory stimuli. J. acoust. Soc. Amer., 1958, 30, 1079-1081.

80. Ward, W. D., Glorig, A., and Sklar, D. L. Dependence of temporary threshold shift at 4 kc on intensity and time. J. acoust Soc. Amer., 1958, 30, 944-954.

81. Webster, J. C. Pitch shifts accompanying certain auditory threshold shifts. J. acoust. Soc. Amer., 1954, 26, 754-758.

82. Wever, E. G., and Truman, S. R. The course of the auditory threshold in the presence of a tonal background. J. exp. Psychol., 1928, 11, 98-112.

83. Zoll, P. M. The relation of tonal volume, intensity, and pitch. Amer. J. Psychol., 1934, 46, 99-106.