A PARALLELOGRAM MODEL OF TIMBRE ANALOGIES

Thesis for the Degree of M. A.

MICHIGAN STATE UNIVERSITY
DAVID EDWARD EHRESMAN

1977

ABSTRACT

A PARALLELOGRAM MODEL OF TIMBRE ANALOGIES

By

David Edward Ehresman

Recent multidimensional scaling studies have found that of the acoustical properties on which tones produced by musical instruments differ, two or three are important in the perception of timbre. These findings were replicated and the dimensions were used as the basis for a parallelogram model of timbre analogies.

Fifteen naturalistic tones were synthesized by using an analysis based additive synthesis technique. The complex time varying amplitude and frequency/phase functions obtained during the analysis step were simplified by replacing them with straight line segment approximations during the synthesis step. Five musically sophisticated and five musically untrained subjects rated the dissimilarity of all possible pairs of the 15 tones. The multidimensional scaling of this data was interpreted in two dimensions. The most potent dimension was interpreted in terms of the spectral energy distribution of the tones. The second dimension was related to the attack portion of the tones, that is,

whether the onset of the higher harmonics was synchronous or asynchronous. These interpretations agree with previous research.

The parallelogram model of timbre analogies is based on a mapping of the stimulus tones onto a multidimensional space. The model assumes that for an analogy of the form A:B::C:? there is some ideal analogy point, I, that will complete a parallelogram. The prediction of the model is that in solving timbre analogies a subject will choose the alternative, D, which is closest to I in the multidimensional space. Three alternatives to the parallelogram model were tested. One possibility is that the subjects are unable to use the directional information that is the basis of the parallelogram model. In this situation, a subject might choose as the best solution to the analogy the alternative D which is most similar to B, the terminal tone of the first half of the analogy. A second possibility is that subjects only use the most salient dimension in solving the analogies by projecting the parallelogram onto that axis. This model predicts that subjects will choose as the best solution the alternative D which is closest to the ideal analogy point in that <u>one</u> dimension. The third alternative model is a combination of the first two. It predicts that the similarity of the terminal tones along the one dimension is the basis for choosing the best solution.

David Edward Ehresman

Using the 15 tones from the first experiment, forty timbre analogies were formed; each analogy had four alternative solutions. Nine subjects from the scaling phase of the study rank ordered the four alternatives as to which best completed the timbre analogy. The parallelogram model best predicted subjects' solutions to the timbre analogies. The effects of musical training were not reflected in either the scaling solution or performance on the analogy task.

A PARALLELOGRAM MODEL OF TIMBRE ANALOGIES

By

David Edward Ehresman

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1977

ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. James Zacks, Dr. Gordon Wood, and Dr. William Hartmann for their help in serving as members of my thesis committee. I am also very grateful to Dr. David Wessel, my committee chairman, whose guidance, assistance, and inspiration made this work possible. My special thanks to to Dr. Judy Frankmann and Joey Mazzella for their editorial assistance and to my wife, Mary Anne, for her understanding support. Finally, I would like to thank Toni Tryon for her willing help in preparing this manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE		PAGE

INTRODUCTION

The spectral energy distribution of a complex tone and the temporal properties of its attack provide major cues in the perception of its timbre. Recent multidimensional scaling studies report timbre spaces with these properties as salient dimensions. A parallelogram model of timbre analogies is based on such a multidimensional timbre space. The assumption is made that for an analogy of the form A:B::C:?, where A, B, and C are points in the timbre space, there is some ideal analogy point, I, that completes a parallelogram. This model predicts that the alternative which is closest to this ideal analogy point will be chosen as the best solution to the analogy.

Timbre

Timbre is not a clearly defined term; it is usually described negatively in terms of the qualities that are left after loudness and pitch have been determined. The definition given by the American Standards Association (1960) is such a definition by exclusion: "Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar." Similarly presented presumably implies that the stimuli are of equal duration and have the same spatial location.

The classical theory of timbre perception originated with Helmholtz. His argument, based on Ohm's Acoustical Law (Ohm, 1843), stated that differences in timbre are a result of the presence and the strength of harmonics in the tone and that the phase relationships among the harmonics make little difference (Helmholtz, 1877/1954). It is important to note that Helmholtz studied only the steady-state portion of complex tones, choosing to ignore the dynamically changing phenomena which are characteristic of naturalistic auditory tones.

Helmholtz was able to show that the steady-state portion of musical and vocal sounds is composed of sets of harmonics and that the ear can distinguish a number of these harmonics individually. After Ohm, he reasoned that the ear performs a Fourier analysis on a tone and thereby identifies the amplitude pattern of the resulting series of harmonically related sinusoids which forms the basis of timbre judgment.

A modification of Helmholtz's classical theory makes use of the notion of formant regions. A formant is a frequency range in which the amplitudes of the harmonics are considerably higher than the amplitudes of the harmonics in the neighboring regions; a formant shows up as a peak in the spectral envelope. This model, which originated in speech perception research, contends that it is the formants, not the harmonics, that provide the major cues in timbre judgments (Fletcher, 1934; Bartholomew, 1945; Slawson, 1968).

The classical theories of timbre dealt only with steady-state tones. Musical tones, however, are usually considered to consist of three segments: (1) the attack, the portion of the tone in which it builds in amplitude, (2) the steady-state, the portion which is reached at the end of the attack, in which the tone remains stable, and (3) the decay, the portion of the tone in which its amplitude decreases until it has finished sounding. An important question is to what extent the perception of timbre is determined by the attack and decay portions of a tone where many dynamic changes occur in the spectral distribution.

Several studies indicate that the acoustical details in the attack transient are very important for the identification of the instrument which produced a tone. In one of these studies (Saldanha and Corso, 1964), trained musicians identified the instrument which produced a tonal stimulus when the tone consisted of (1) the initial transients and a short steady-state (1/3 sec.), (2) the initial transients, a short steady-state, and the decay transients, (3) the initial transients, a long steady-state (9 sec.), and the decay transients, (4) a short steady-state without the characteristic attack or decay, or (5) a short steady-state and the decay transients. A tone for each condition was obtained from eight wind instruments and two string instruments. The absence of the decay transients (Group 1 vs. Groups 2 and 3; Group 4 vs. Group 5) appeared to have minimal effect on the ability to recognize the instrument (i.e., label the timbre). However, the absence of the characteristic attack transient (Groups 1, 2, and 3 vs. Groups 4 and 5) was detrimental to performance.

Berger (1964) found similar results for 10 wind instruments. He asked musicians to identify the instrument which produced a tone when the stimulus consisted of (1) the unaltered tone, (2) the steady-state portion of the tone, or (3) the fundamental component of the tone with the harmonics filtered out. As expected, the unaltered tones were easier to identify than those without the transients, which in turn were easier to identify than those with the harmonics filtered out. Wedin and Goude (1972) also noted that attack transients were important in the identification of the timbre of a tone. In addition, Saldanha and Corso (1964) found that the timbre of tones played with vibrato, a temporally dynamic phenomenon, was easier to identify than the timbre of tones played without vibrato.

This work in timbre perception was valuable to investigators in the related field of music synthesis. With a growing awareness that temporal events play an important role in timbre perception, these researchers sought a way to synthesize naturalistic tones with dynamic attack transients. Tones were analyzed to determine the temporal envelope (time vs. amplitude) for several or all of the harmonics in the tone. Part or all of this information was then used to synthesize tones.

Strong and Clark (1967a, 1967b) determined the steady-state spectral envelope (frequency vs. amplitude) and several temporal envelopes for three subsets of harmonics associated with each of nine wind instruments. Using this information, they synthesized tones with the aid of a digital computer. Music students were able to identify the synthesized tones with 66% accuracy as compared to 85% accuracy for the natural tones. After observing the effects of systematically exchanging envelopes among instruments, Strong and Clark concluded that tones from some instruments had unique spectral envelopes and that for these tones the spectral envelope was more important for correct identification than was the temporal envelope. Tones from other instruments did not have unique spectral envelopes, and for these tones the temporal envelope was as important or more important than the spectral envelope in the identification process.

Grey (1975) approached the synthesis problem somewhat differently. First, he determined the time varying amplitude and frequency functions for each of the harmonics present. Then he synthesized tones using all of the information to test the validity of the analysis (full data tones). Next he systematically simplified the information used in the synthesis to determine the effect of particular types of information. The simplifications he used were: (1) representing the complex time variant amplitude and frequency functions with small numbers of straight line segments (line segment approximation tones), (2) excluding any clearly delineated initial attack segments which contained low-amplitude inharmonicities (cut attack tones), and (3) substituting constant frequencies for the time variant frequency functions while retaining the time variant line segment approximations for the amplitude functions (constant frequency tones). Grey (1975) tested for discriminability between the five types of tones, i.e., the original tones in digitized form, the full data tones, and the three types of simplified tones; he also asked his musically trained listeners to rate the subjective distances between tones. The results of both of these measures indicated that the full data tones were an adequate representation of the original tone. Verbal reports of the listeners indicated that differences between the original and the full data tones were difficult to detect, and when detected, the stimuli were described as tones from the same instrument played with a different articulation or style of playing. The line segment approximation tones were very similar to the full data tones; however, the constant frequency tones and the cut attack tones were too discriminable to be of general use. Thus, it appears that perceptually convincing renditions of naturalistic tones can be obtained when the details of amplitude and frequency variation in each harmonic are approximated with relatively few linear segments.

Several studies have used multidimensional scaling (MDS) techniques (Shepard, 1962a, 1962b; Kruskal, 1964a, 1964b;

Ca

ti

si

to

in

Eu

tw

hy

ti

a r

onl

ie

rat

ins

ded

The

tio

ene

pre

othe

The

sing

tend

stri

Prop

segm

Carroll and Chang, 1970) to gain a better understanding of timbre perception. Using matrices of dissimilarities (or similarities) between objects, these MDS techniques attempt to represent the dissimilarities as distances between points in an n-dimensional space (usually a two or three dimensional Euclidean space). The resulting "picture" can be viewed in two ways: (1) as a very useful data reduction suggesting hypotheses for new lines of research, or (2) more speculatively, as a model of the way the objects are perceived.

Plomp and Steeneken (1969) were the first to use MDS as a model of timbre perception. However, their study used only steady-state tones. Tones which included the transients were used by Wessel (1973), who had music students rate the dissimilarity of tones played on nine orchestral instruments. Using MDS as a data reduction tool, he embedded the instruments in a two dimensional Euclidean space. One dimension differentiated among timbres by the distribution of energy in the steady-state region of the tones. The energy of the tones on one end of this dimension was located predominately in the lower harmonics, while tones on the other end had more energy located in the higher harmonics. The second dimension was more difficult to interpret using a single physical characteristic of the tones. The tones tended to be grouped by family (i.e., brass, woodwinds, and strings). This dimension appeared to be related to temporal properties of the tone, i.e., differences in the attack segment. Wessel and Grey (in press) also scaled the similarity judgments of nine instrument tones reported by Wedin and Goude (1972). The results were similar to those described above and strongly supported the important role of attack transients in providing perceptual distinctions among instruments.

Grey (1975) used MDS techniques to analyze the similarity of the 16 "line segment approximation" tones described earlier. Grey noted three ways in which these tones differed from the tones used by Wedin and Goude (1972) and Wessel (1973). They were (1) shorter in duration, (2) synthesized naturalistic tones rather than actual natural tones, and (3) experimentally equalized for pitch, loudness, and perceived duration. His data reduction yielded not two, but three dimensions.

One dimension, the spectral energy distribution, paralleled the first dimension found in previous studies (Wessel, 1973; Wedin and Goude, 1972 as scaled by Wessel and Grey, in press). A second dimension reflected patterns in the onset-offset portion of the tones. At one extreme of this dimension all of the upper harmonics entered and exited at approximately the same time or with synchrony; at the other extreme the upper harmonics entered and exited gradually or asynchronously. This dimension also related fairly well to musical instrument families. The third dimension also focused on the attack segment of the tones. Tones were differentiated by the presence of high frequency, low amplitude, usually inharmonic energy, during the attack segment as opposed to low frequency inharmonic energy, or at least the absence of high frequency inharmoniousness in the attack. Since two of Grey's dimensions were interpreted in terms of the attack transients, it appears that they are encompassed by Wessel's second dimension. The consistency of findings and the complementary nature of the results in these studies clearly pin-point important attributes of tones which need to be taken into consideration in future studies of timbre.

Analogical Reasoning

Rumelhart and Abrahamson (1973) have presented an intuitively appealing theoretical model of analogical reasoning based on MDS techniques. They assume that the elements to be used in forming analogies have been embedded in a multidimensional space. Their model states that for an analogy of the form A:B::C:?, there is some "ideal analogy point" (Rumelhart and Abrahamson, 1973, p. 4), "I", that completes the analogy such that line segments connecting the four points in the multidimensional space will form the sides of a parallelogram. In other words, there is some vector, CI, which is parallel to and equal in length to the AB vector. The coordinates of this ideal analogy point, I, can be computed from the following formula:

$$I(j) = C(j) + B(j) - A(j), \quad j = 1,n \qquad (1)$$

where I(j), C(j), B(j), and A(j) refer to the coordinate on the jth dimension of points I, C, B, and A respectively, and n is the dimension of the multidimensional space.

This model, which will be referred to as the parallelogram model of analogical reasoning, predicts that for an analogy of the form A:B::C:D1, D2, D3, or D4, the probability that a particular alternative will be chosen as the best solution to the analogy is a monotonic decreasing function of the distance between that point and the point "I" (the ID distance) in the multidimensional space.

Rumelhart and Abrahamson found support for this model by using a three dimensional space of animal names obtained from a scaling study by Henley (1969). A second experiment reported in the same paper found further support for the model. Forming analogies from the same set of animal names they (1) replicated the first experiment, (2) showed that the probability of choosing a particular alternative did not depend on the particular analogy problem, provided that the distances between the ideal solution point and the alternative were approximately equal, (3) found support for the idea that the monotonic decreasing function which relates the probability that an alternative will be chosen as the best solution to the ID distance is an exponential one, and (4) found that the 2nd, 3rd, and 4th best solutions were also predicted by the distance between the alternative and the ideal analogy point.

An implicit assumption of the parallelogram model is that subjects are able to judge the similarity of the vectors involved in an analogy. If subjects are unable to appreciate the directional information implied by the concept of vectors, they may solve the analogies by choosing the alternative D (the endpoint of the CD vector) which is most similar to B (the endpoint of the AB vector). This alternative hypothesis will be referred to as the similarity of terminal tones model.

A second assumption is that subjects are able to use multidimensional information in solving an analogy. However, due to the complexity of the timbre analogy task, subjects may resort to using only the most potent dimension in solving an analogy. This possibility gives rise to two more alternative hypotheses.

The first of these potent dimension hypotheses is based on the parallelogram model. It states that subjects project the parallelogram onto the most potent dimension and proceed as in the parallelogram model. This hypothesis predicts the choice of the alternative D which is closest to the ideal analogy point along the one dimension as the best solution.

The second potent dimension hypothesis is based on the similarity of terminal tones hypothesis. The prediction of this model is that the alternative D which is closest to tone B along the potent dimension will be selected as the best solution to an analogy.

The purpose of the following two experiments is to test whether Rumelhart and Abrahamson's parallelogram model will predict subjects' choices of the best solutions to timbre analogies more accurately than the three alternative hypotheses. As discussed previously, recent MDS solutions of tones of different timbre (Grey, 1975; Wedin and Goude, 1972 as scaled by Wessel and Grey, in press; Wessel, 1973) have resulted in two or three interpretable axes. In Experiment 1, a scaling of 15 tones of different timbre gave rise to a two dimensional timbre space that is comparable to those found by other researchers. Experiment 2 used this scaling solution to test Rumelhart and Abrahamson's (1973) parallelogram model of analogical reasoning in the timbre domain.

SYNTHESIS OF STIMULI

In the following experiments, subjects were asked to make judgments about 15 tones with different timbres. The specific tones that were used were 15 of Grey's (1975) line segment approximation tones, as discussed in the Introduction. These tones were originally played on the following musical instruments: oboe (2 different instruments and players), English horn, bassoon, $E^{\flat}$ clarinet, bass clarinet, flute, alto saxophone (2 tones from one instrument, played at p and mf) soprano saxophone, trumpet, French horn, and cello (3 tones from one instrument, played normally, muted sul tasto, and sul ponticello). All 15 tones were played near the pitch of $E^{\flat}$ above middle C (approximately 311 Hz), at approximately the same loudness level, and with durations between 280 and 400 milliseconds. These tones were then analyzed by Grey using the heterodyne filter method (Moorer, 1973), which produces time variant amplitude and frequency/ phase functions for each harmonic. The tones were then resynthesized by summing a set of harmonic sinusoids that were controlled in time by the amplitude and frequency functions obtained in the analysis stage. This process is called analysis based additive synthesis. The line segment approximation tones were synthexized by replacing the extremely complex time variant amplitude and frequency/phase functions with a small number of straight line segments and using these to control the harmonic sinusoids.

The tones used in this study were synthesized at Michigan State University using a specialized additive synthesis program implemented on a PDP-11/40 digital computer (See Appendix). The data were supplied by John Grey in the form of break-point tables (i.e., a list of the endpoints of each line segment) of amplitude and frequency/phase functions for each harmonic of each of the tones. The values of the amplitude and frequency/phase functions (as determined from the break-point table) for each time period were used to control the sampling from a sinusoid for each harmonic and these samples were summed to yield the waveform value for that time period. This process was repeated until the entire tone was determined and stored on a magnetic disk in digital form. Figures 1-15 display the waveforms of these tones. Time is on the X axis, amplitude is on the Y axis, and frequency (with the fundamental in the background) is on the Z axis. A sampling rate of 25,000 samples per second was used. During the experiments the tones were played via a 16 bit digital to audio converter (DAC), constructed by Three Rivers Computer Corporation (Kriz, 1975).

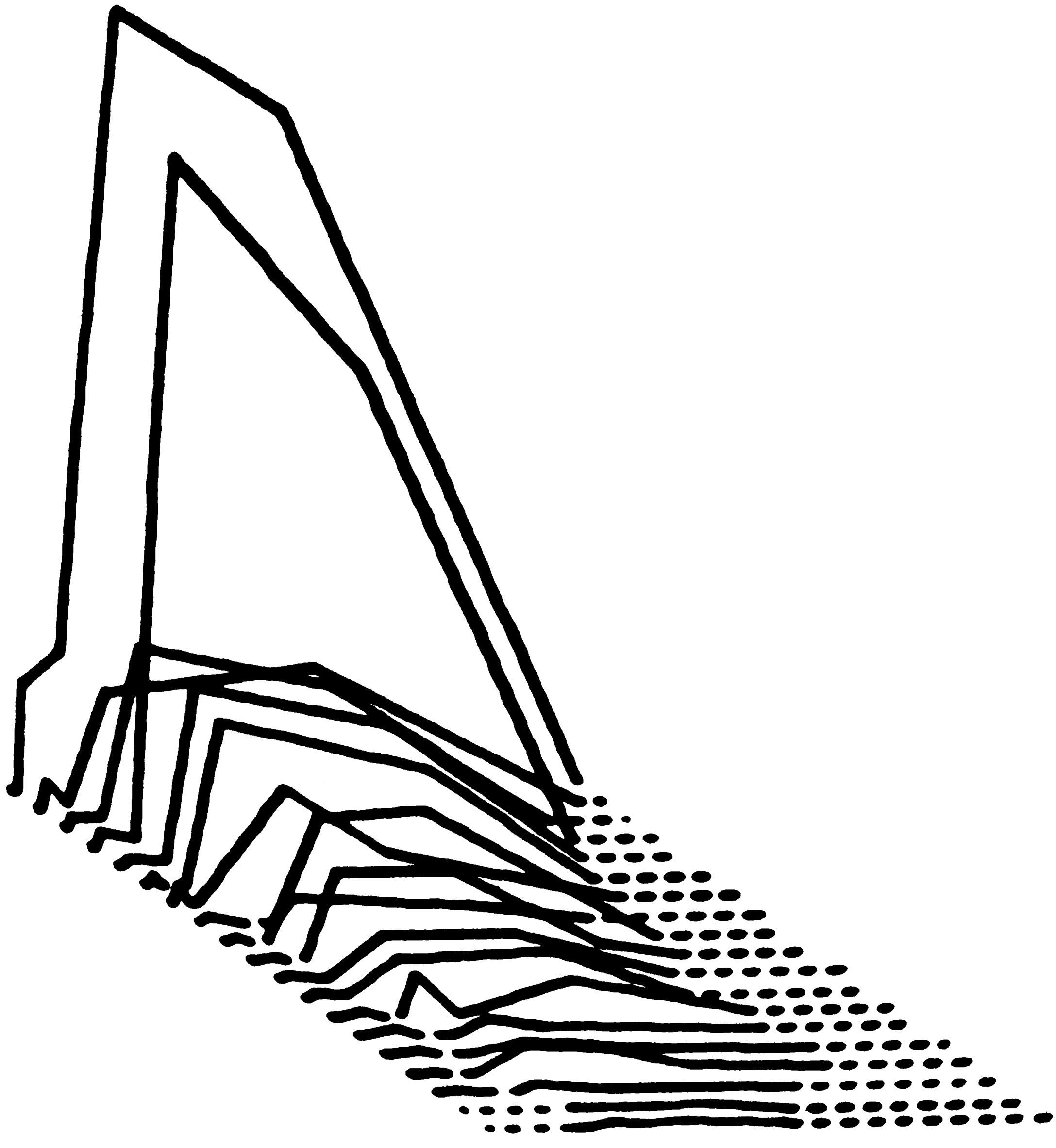

Figure 1. Oboe 1 (01) amplitude envelope.

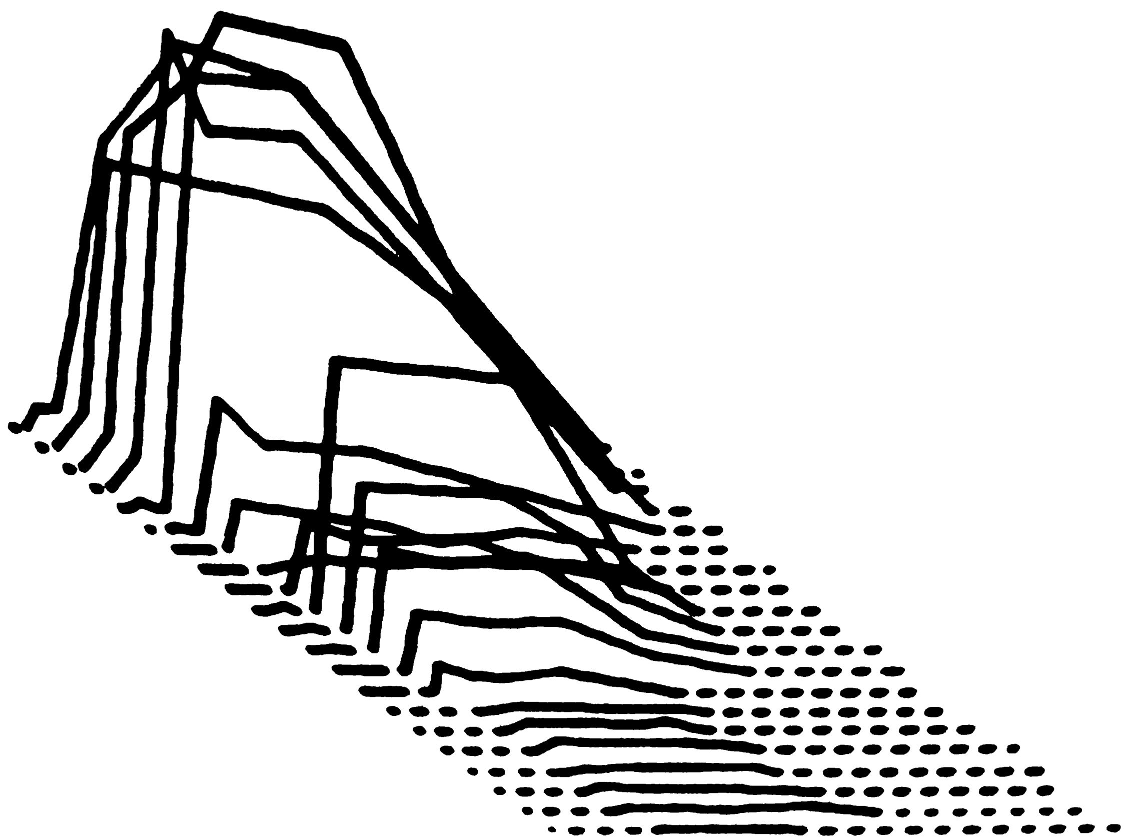

Figure 2. Oboe 2 (02) amplitude envelope.

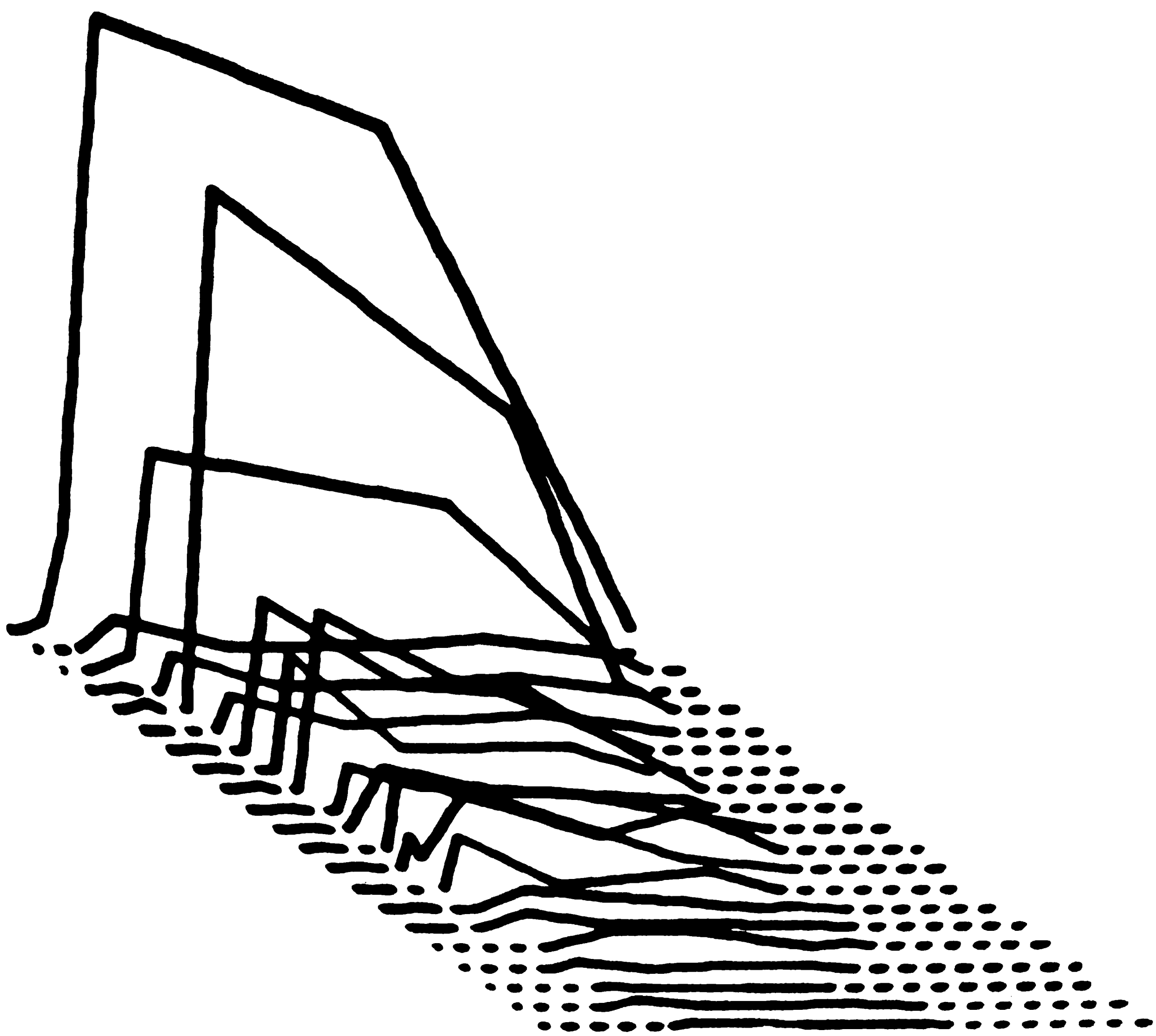

Figure 3. Clarinet 1 (C1) amplitude envelope.

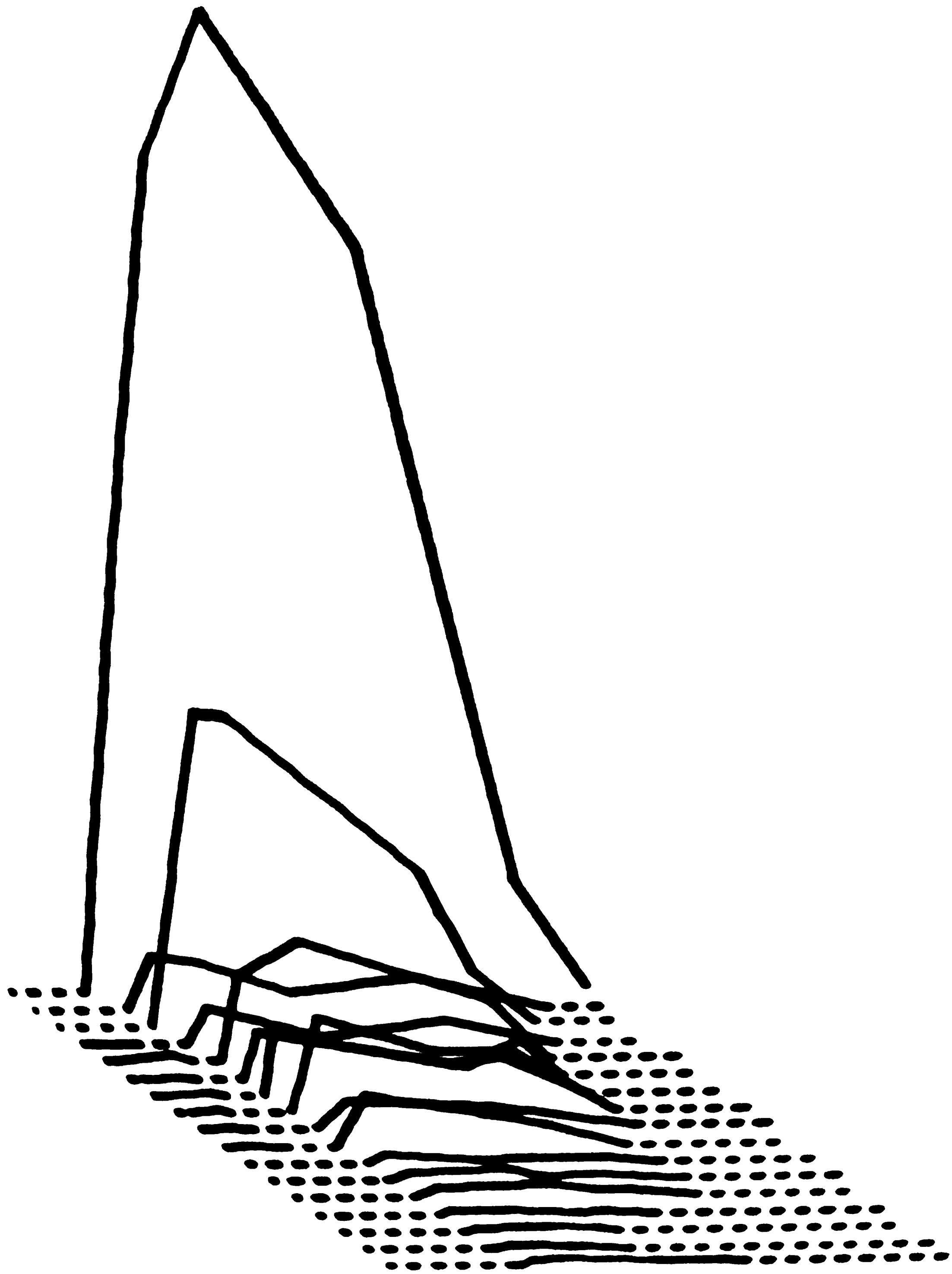

Figure 4. Clarinet 2 (C2) amplitude envelope.

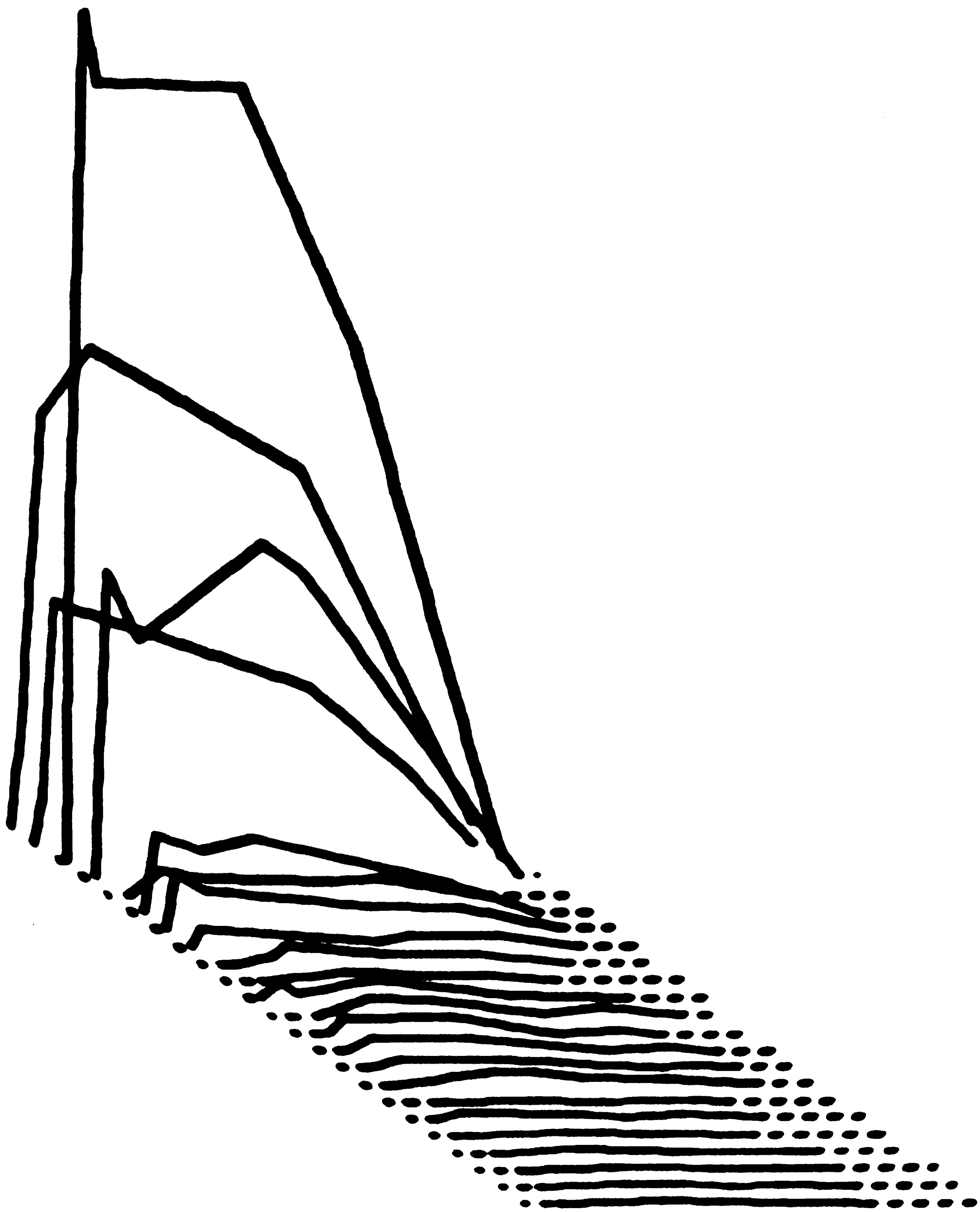

Figure 5. English horn (EH) amplitude envelope.

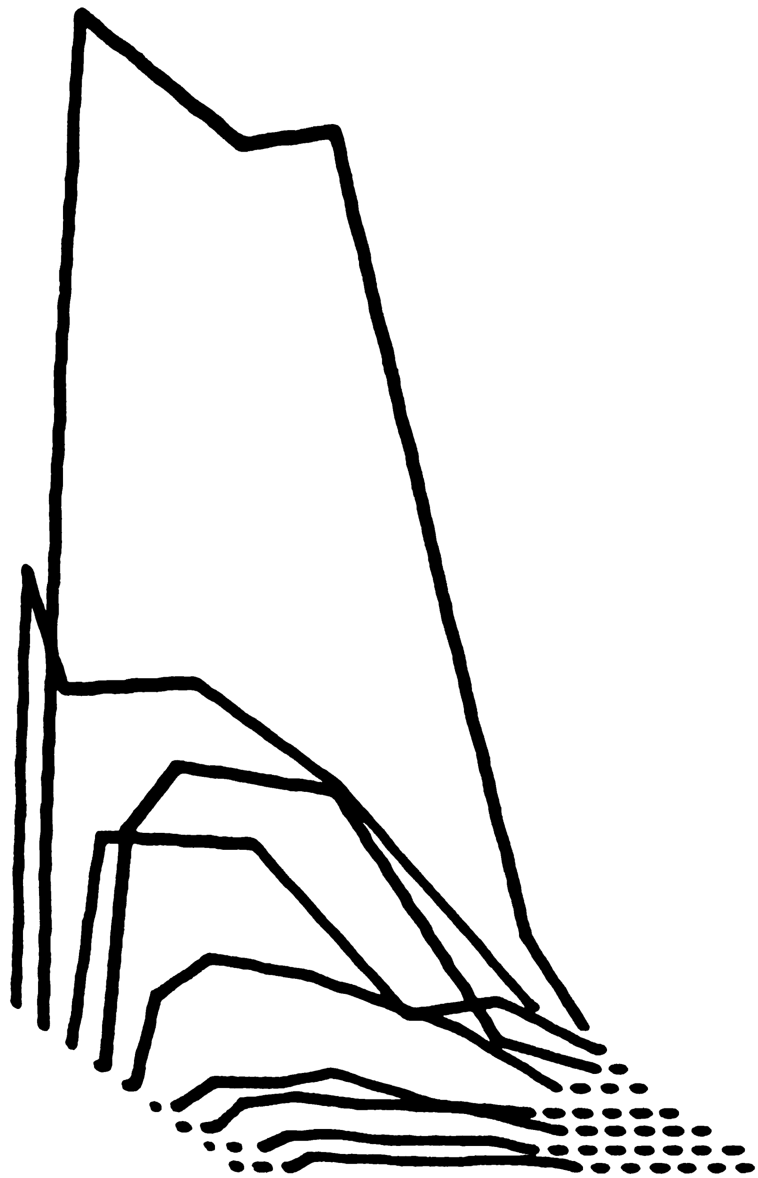

Figure 6. Bassoon (BN) amplitude envelope.

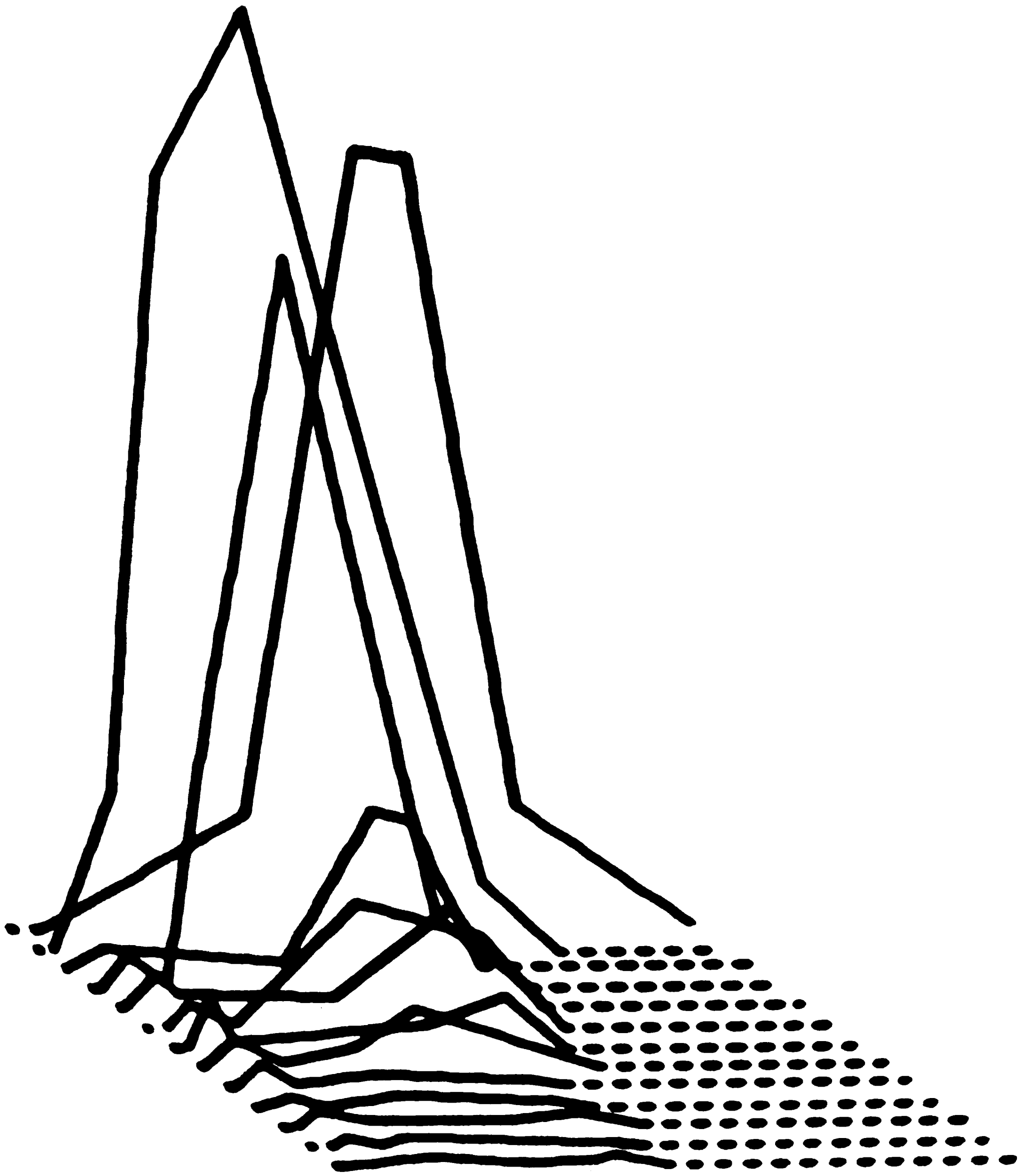

Figure 7. Flute (FL) amplitude envelope.

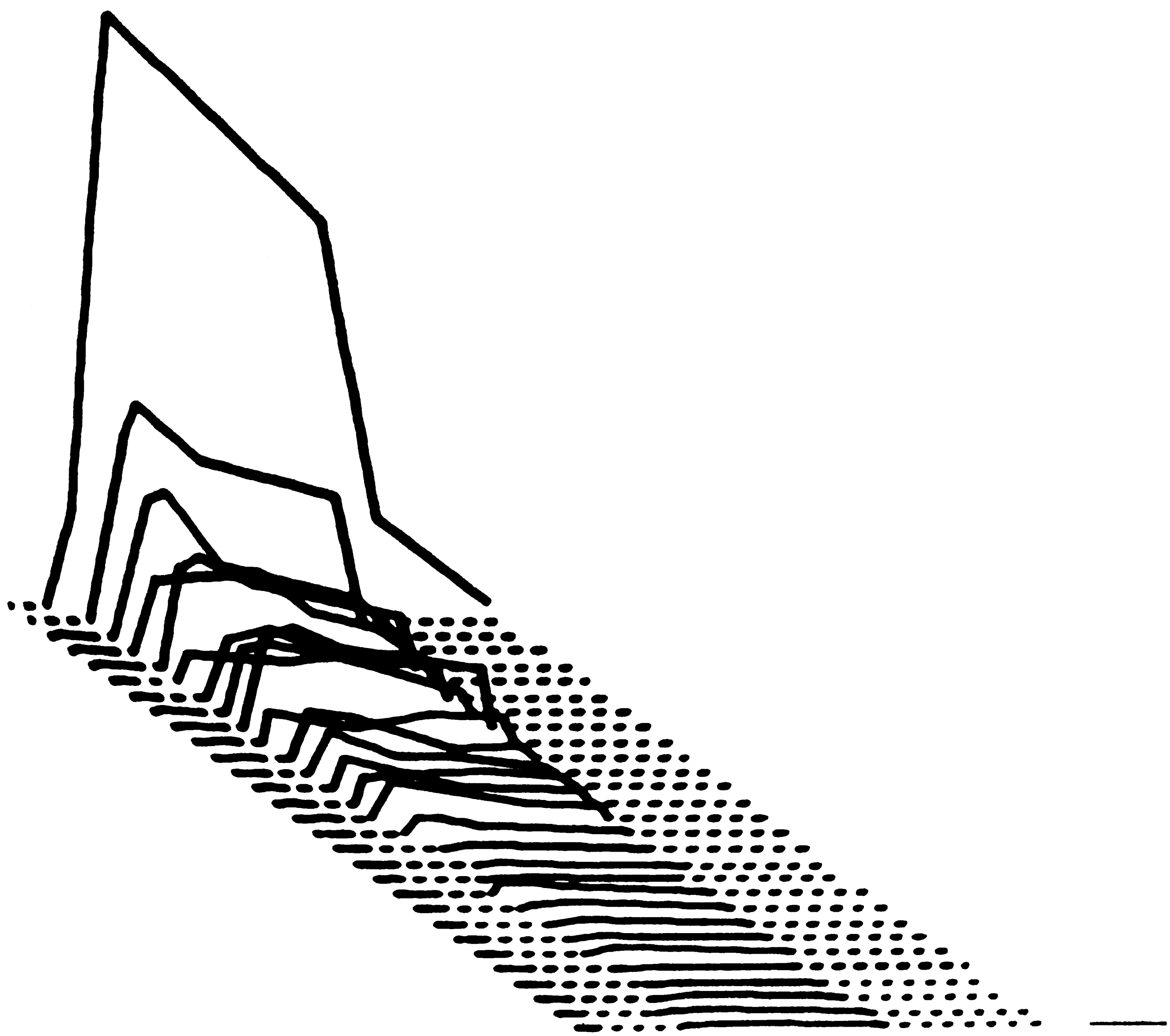

Figure 8. Saxophone 1 (X1) amplitude envelope.

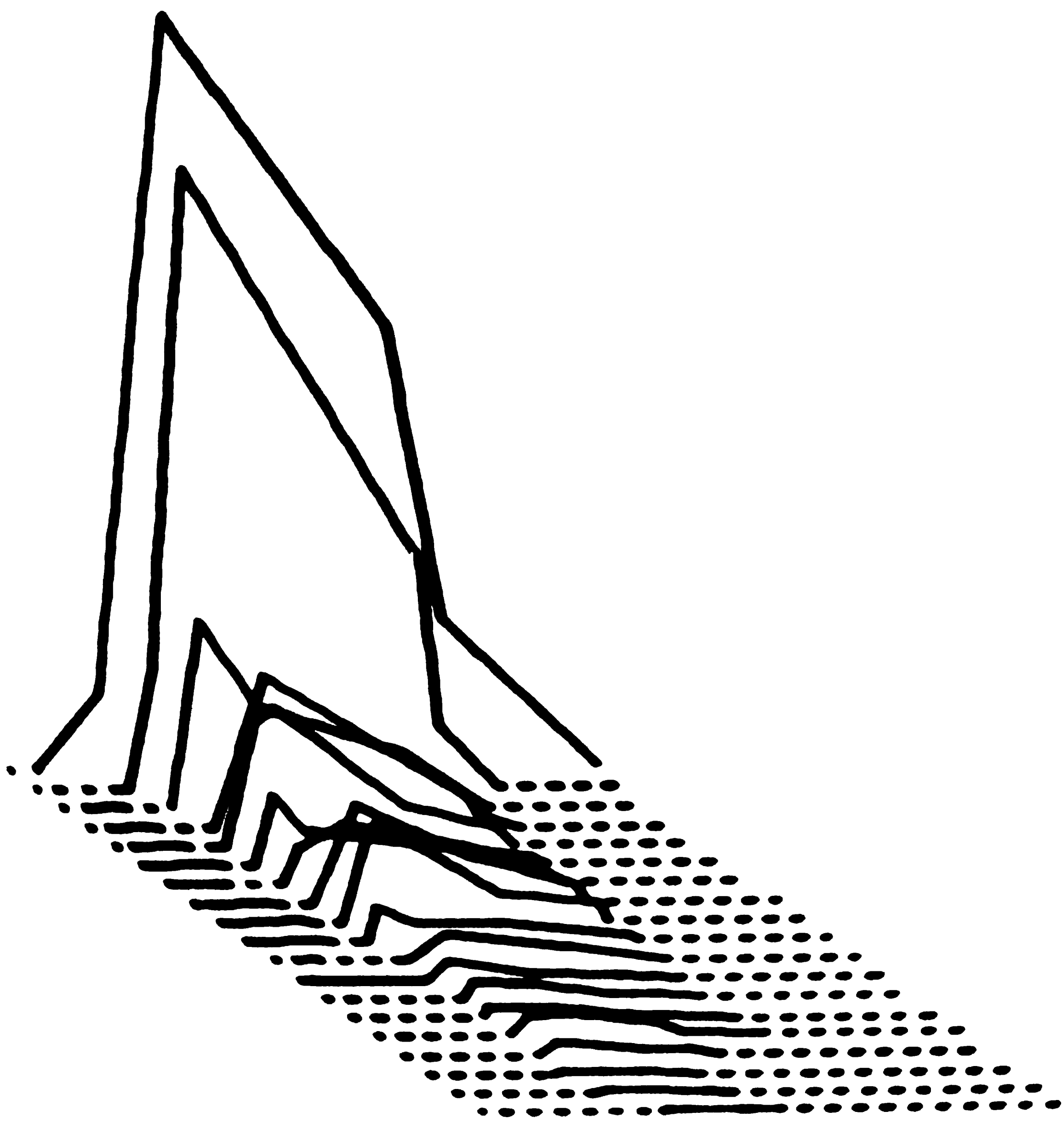

Figure 9. Saxophone 2 (X2) amplitude envelope.

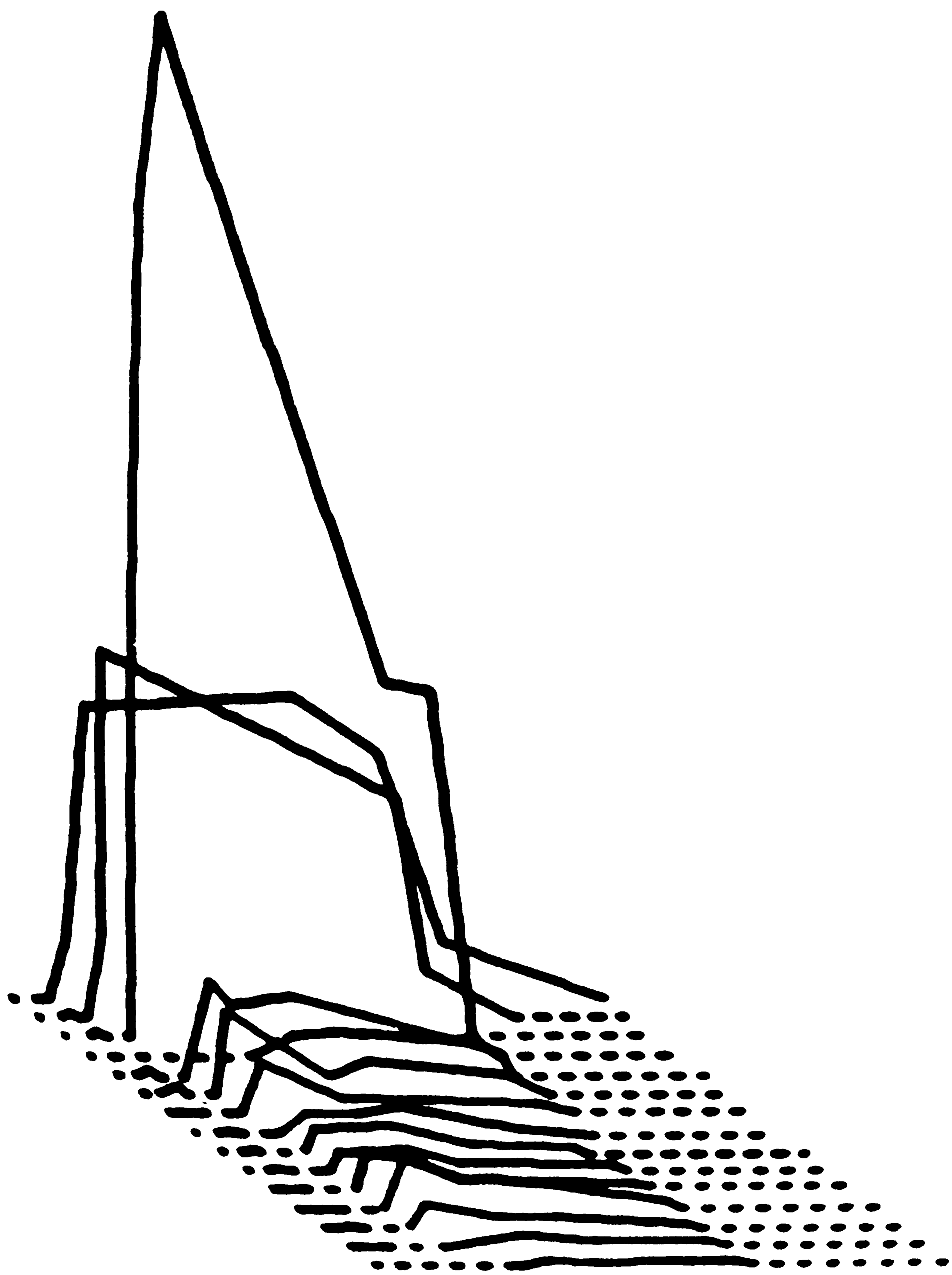

Figure 10. Saxophone 3 (X3) amplitude envelope.

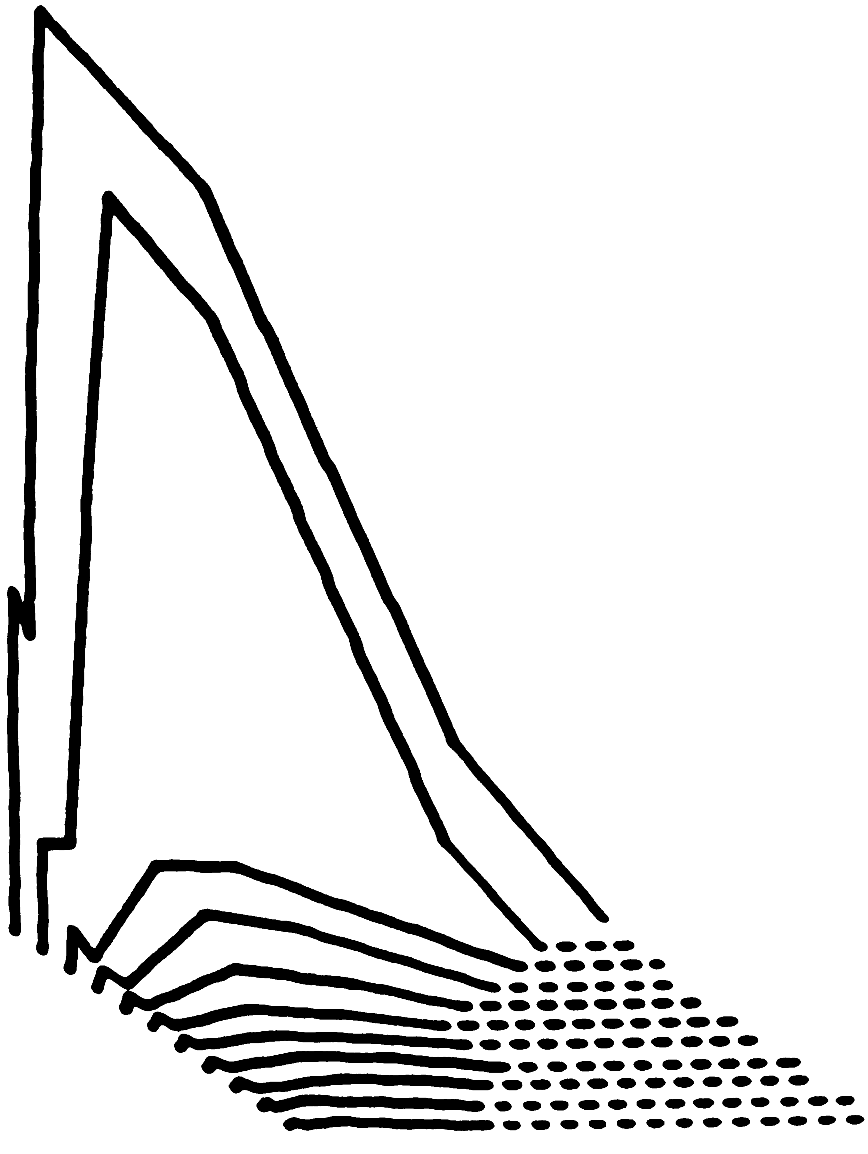

Figure 11. French horn (FH) amplitude envelope.

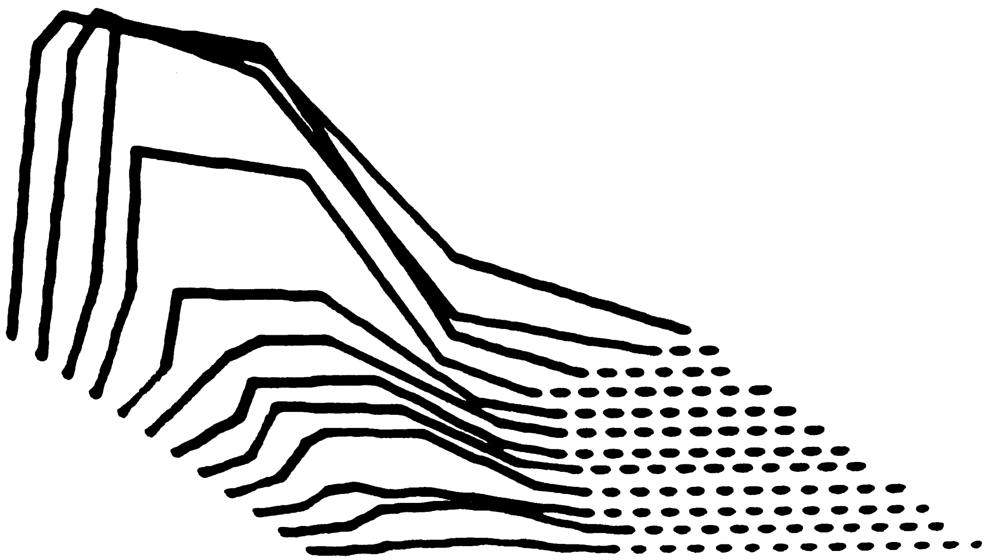

Figure 12. Trumpet (TP) amplitude envelope.

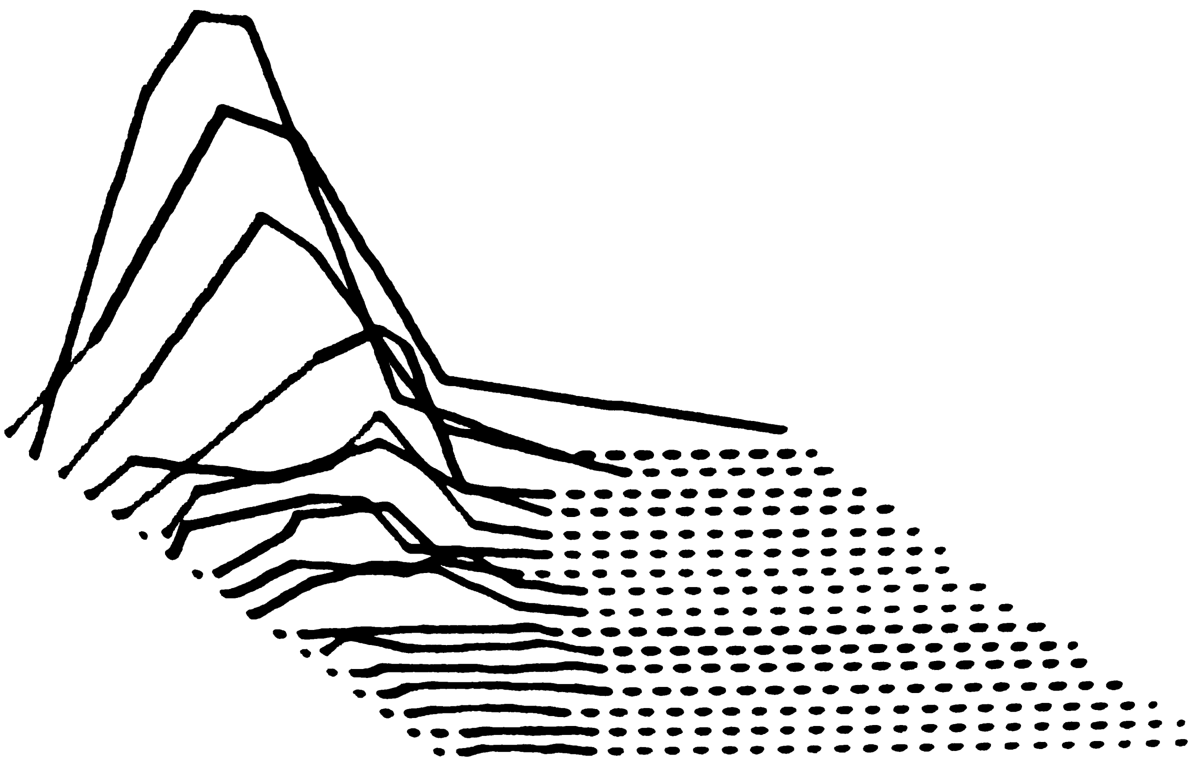

Figure 13. Cello 1 (S1) amplitude envelope.

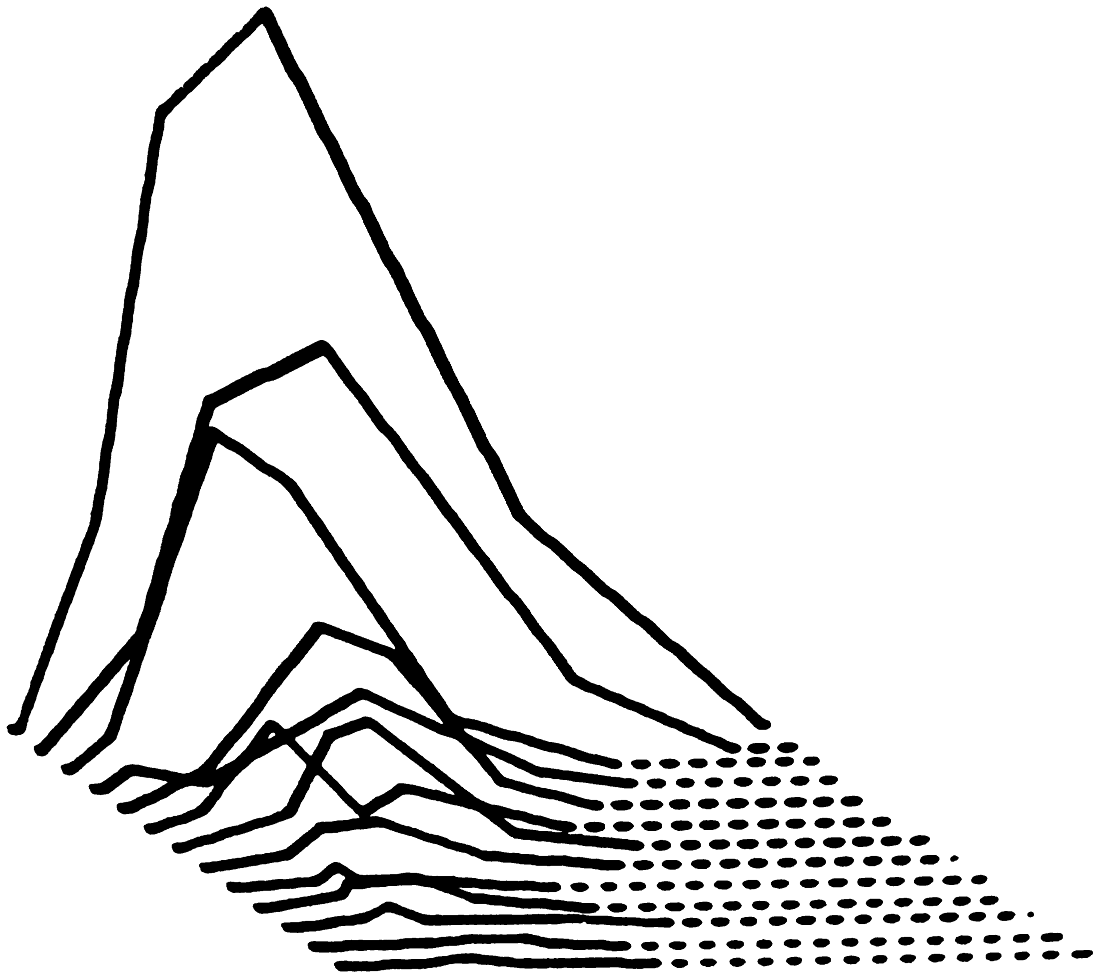

Figure 14. Cello 2 (S2) amplitude envelope.

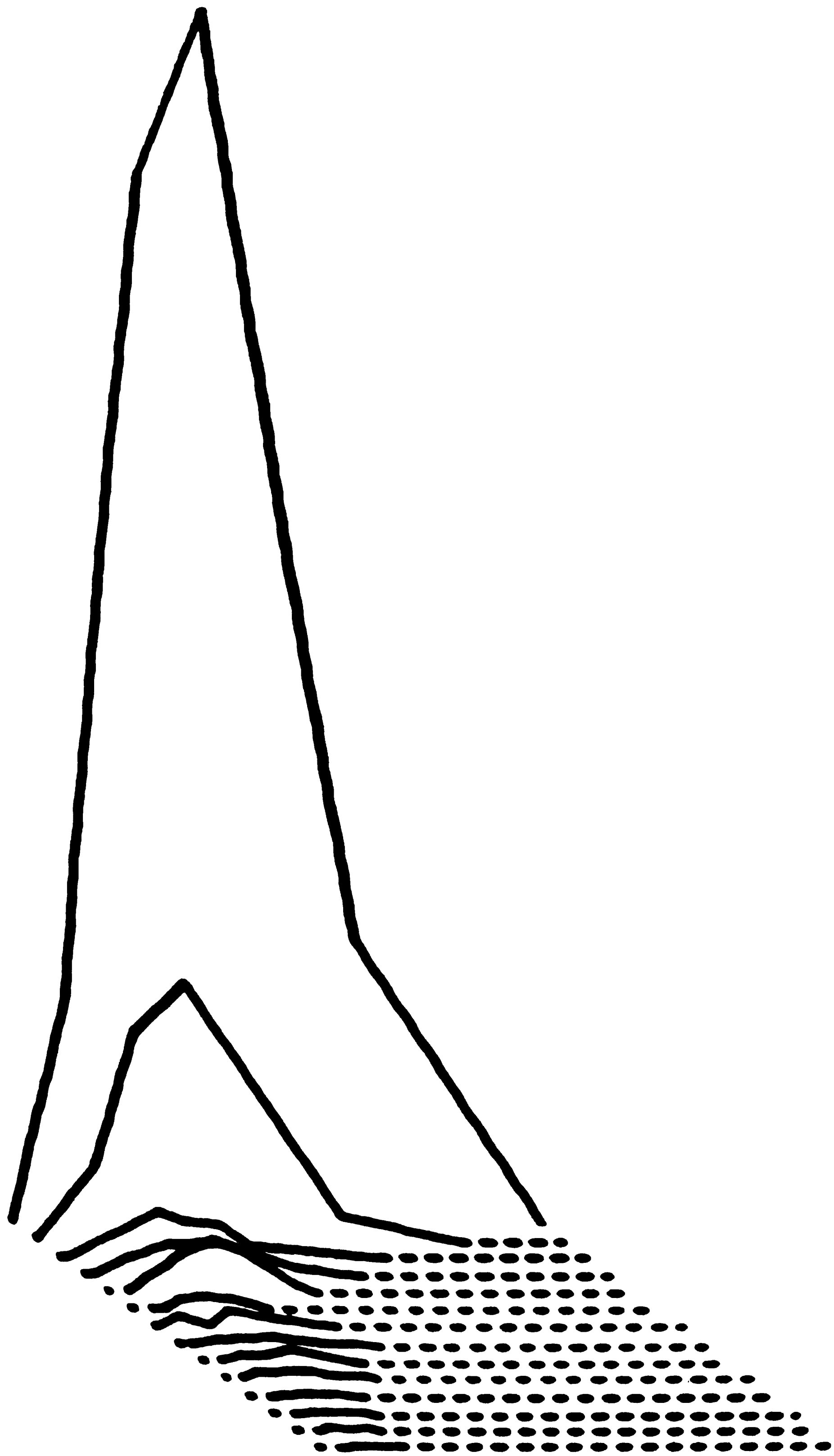

Figure 15. Cello 3 (S3) amplitude envelope.

EXPERIMENT 1: DISSIMILARITY JUDGMENTS OF NATURALISTIC TONES

Stimuli

The 210 possible pairs of the 15 tones were formed and randomized for each subject.

Procedure

Subjects were asked to judge how dissimilar each pair of tones was. The tones were presented to subjects in a sound chamber via the DAC over a Philips 532 Motional Feed-back loudspeaker. Subjects sat approximately two and one half feet from the speaker. The tones were presented in pairs; in order to hear a pair, the subject pressed a button switch which was interfaced with the PDP-11/40. The subject was allowed to listen to each pair as many times as he or she desired and then registered a dissimilarity judgment by using a linear potentiometer. Each judgment (the position of the potentiometer) was read by the PDP-11/40 via a DR11C interface when a second button switch was depressed, and was stored on a magnetic disk for later analysis. Subjects were given 20 practice trials to become familar with the procedure.

Subjects

Five musically sophisticated and five musically untrained subjects were recruited from students and faculty at Michigan State University.

Results and discussion

The dissimilarity judgments were analyzed using two MDS programs, KYST (Kruskal, 1964a, 1964b; Young and Torgerson, 1967; Shepard, 1962a, 1962b; Torgerson, 1958) and INDSCAL (Carroll and Chang, 1970). Data from the musically sophisticated subjects and from the musically untrained subjects were analyzed separately using the KYST program. Three, two and one dimensional solutions were obtained for both sets of data. The goodness-of-fit of these solutions is assessed by stress measures which are shown in Table 1. Large values of

Table 1. KYST stress values.

Number of Dimensions	Stress	
	Musically Sophisticated	Musically Untrained
Three	0.21	0.22
Two	0.28	0.30
One	0.43	0.43

stress indicate a poor solution. For both groups, the goodness of fit for the three dimensional solution was not substantially better than that for the two dimensional solution. Stress for the one dimensional solutions was markedly higher than that for the two dimensional solutions. Therefore, the two dimensional solution seemed most appropriate. The Shepard diagrams (solution distances versus the data fitted with the monotone regression line) for both two dimensional solutions are shown in Figure 16. Although the stress was within acceptable limits, there was still considerable scatter.

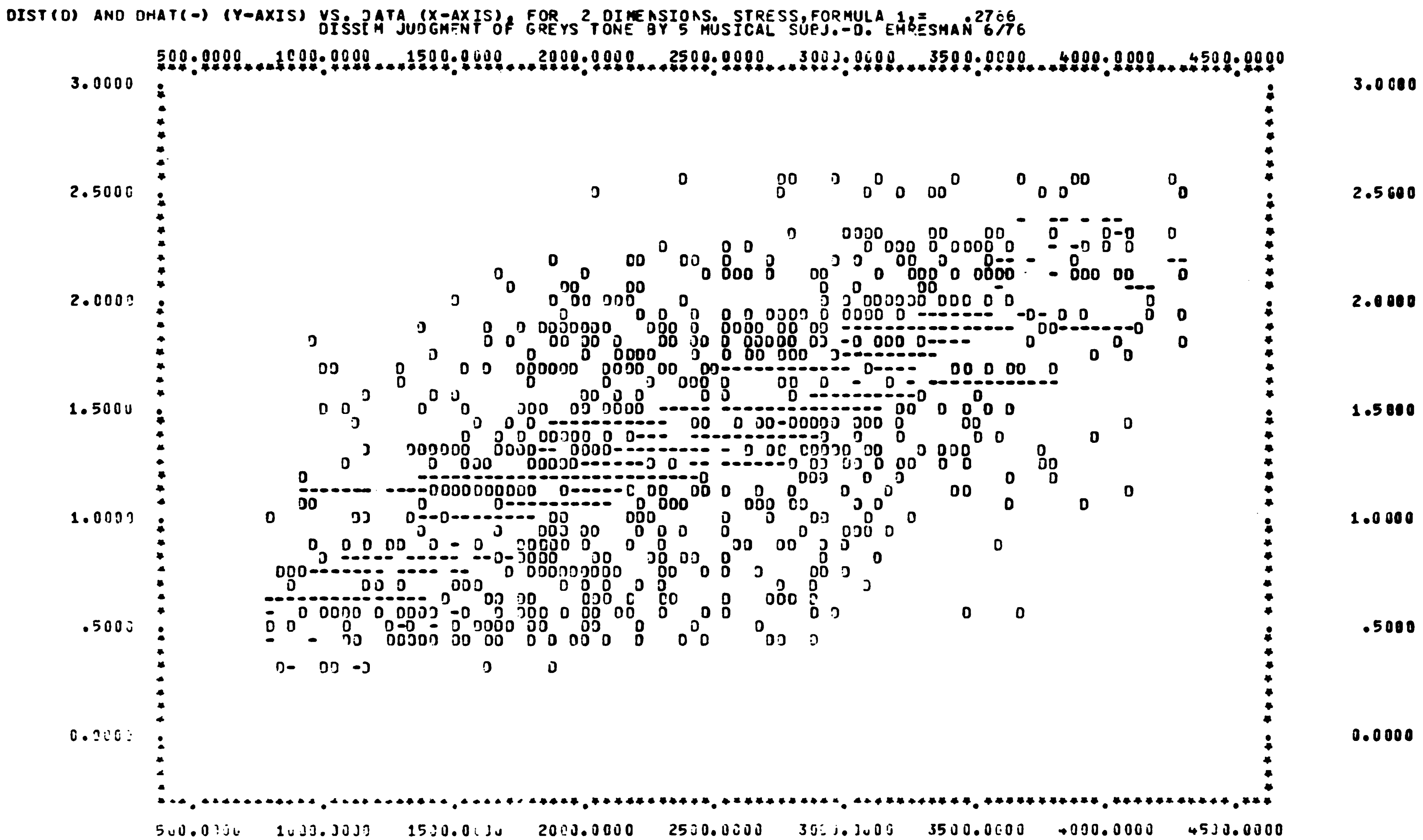

Figure 16a. Composite Shepard diagram for five musically sophisticated subjects.

Figure 16b. Composite Shepard diagram for five musically untrained subjects.

The goodness-of-fit measure for INDSCAL is the correlation between the data and the distances in the solution. Again the three dimensional solution was little better than the two dimensional solution (Table 2); however the correlation did decrease when a one dimensional solution was obtained.

Table 2. INDSCAL goodness-of-fit correlations.

Number of Dimensions	Correlation
Three	0.65
Two	0.62
One	0.54

An examination of the subjects' weight space for the three dimensional INDSCAL solution (Figure 17), yields yet another reason for choosing the two dimensional solution: Subject U4 was the only one to place any weight on the third dimension. Since there was no substantial improvement in goodness-of-fit for the third dimension and since only one subject used the third INDSCAL dimension, the two dimensional solution seemed to be a more appropriate representation of the data.

Figures 18 and 19 show the INDSCAL two dimensional subjects' weight space and group timbre space. The horizontal dimension of the timbre space closely corresponds to the first dimension found by Wessel and to Grey's Y dimension. At one extreme are the tones, from instruments such as the French horn and the cellos, which have most of their energy located in the lower harmonics. At the other extreme are the tones, such as those produced by the saxophones and oboes,

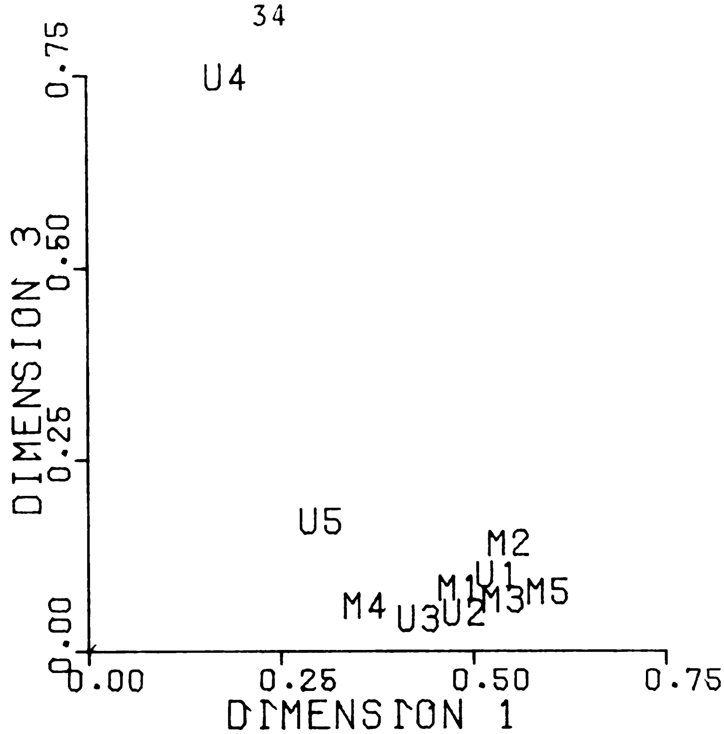

Figure 17. Two dimensional subspace of the INDSCAL three dimensional weight space.

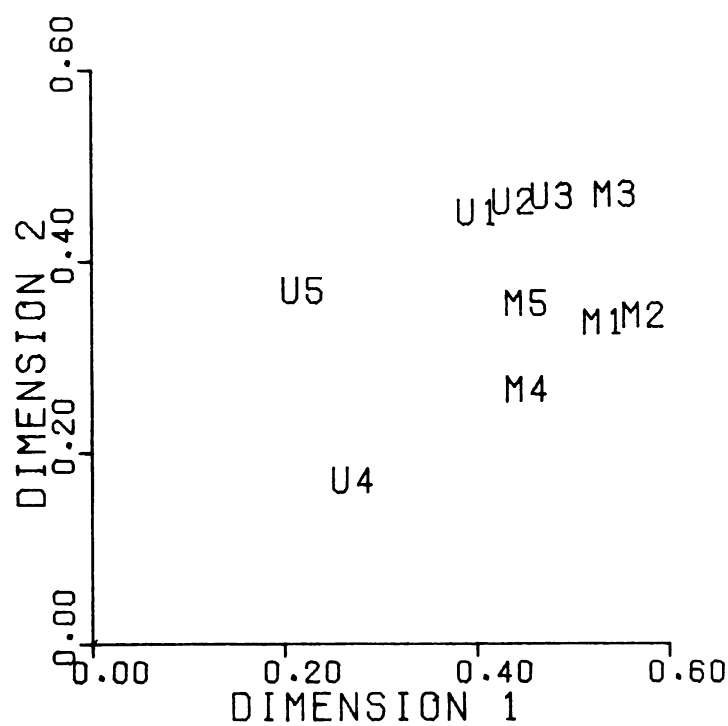

Figure 18. INDSCAL two dimensional weight space.

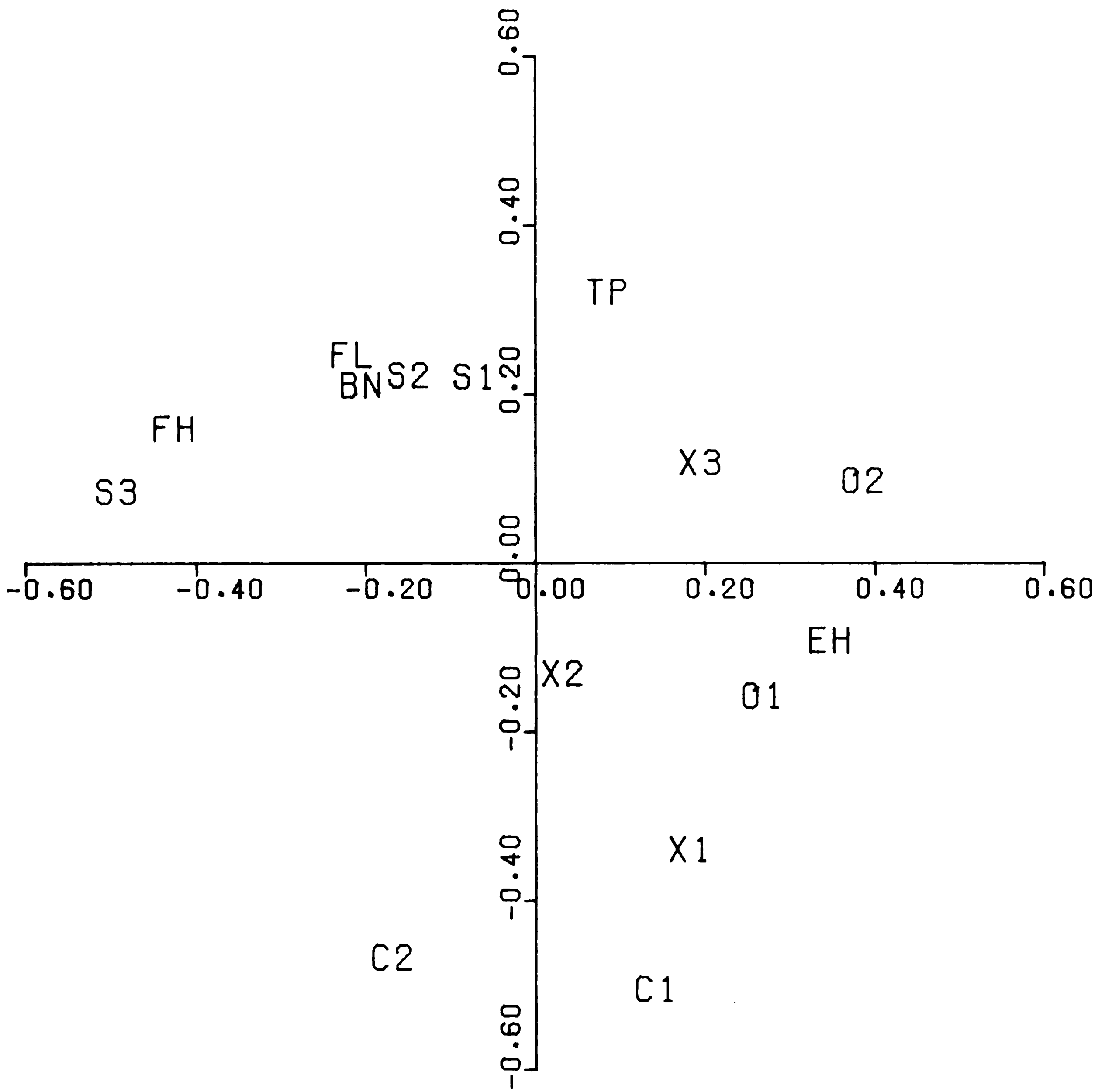

Figure 19. INDSCAL two dimensional timbre space.

which have more of their energy located in the higher harmonics. One can think of this as a mellow to bright continuum.

The first moment of the average amplitude of the harmonics can be used as a quantitative measure of this energy distribution dimension. The average amplitude of the kth harmonic, AA(k), was computed as:

$$AA(k) = \frac{\int f(t)\,dt}{b - a}, \tag{2}$$

where f(t) = the amplitude of the kth harmonic at time t, a = the start time of the tone, and b = the stop time of the tone. The first moment of the average amplitudes for a tone was calculated as:

$$M = \frac{\sum_{k=1}^{n} k\,AA(k)}{\sum_{k=1}^{n} AA(k)}, \tag{3}$$

where M = the first moment, k = the harmonic number, and n = the number of harmonics in the tone. This measure of energy distribution is highly correlated ($r = 0.85$) with the X coordinates of the INDSCAL solution (Figure 20).

As in previous work, the second dimension appears to be determined by the attack portion of the tones. The onsets of the upper harmonics in the tones at one extreme of this dimension tend to be asynchronous; at the other extreme the onsets look much more synchronous. The tones at the asynchronous extreme, from instruments such as the clarinets and the saxophones, have a roughness or raggedness in the attack;

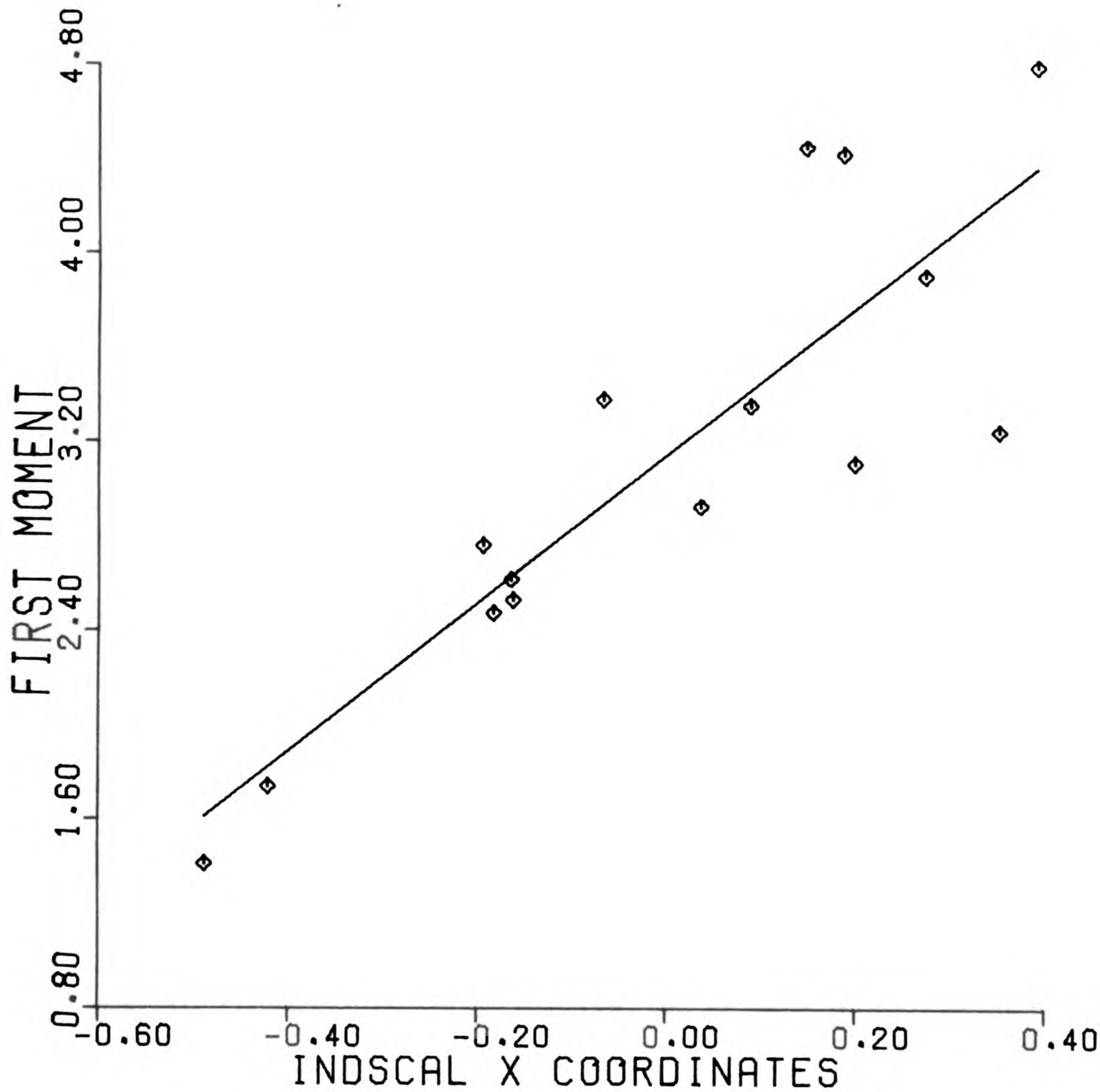

Figure 20. INDSCAL X coordinates versus the first moment of the average amplitude of the harmonics.

the tones which are more synchronous in the onsets of the upper harmonics, created by instruments such as the trumpet and the bassoon, have a much sharper or clearer attack.

A weighted standard deviation of the start times of the upper harmonics, beginning with harmonic six, was chosen as a quantitative measure of this dimension. The standard deviation was calculated as:

$$\mathrm{WSD} = \frac{\Sigma\ (t(k) - \bar{t})^2 * \mathrm{AMP}(k) \ / \ \mathrm{TOTAMP}}{n - 5}, \tag{4}$$

where WSD = the weighted standard deviation, $t(k)$ = the onset time of the kth harmonic, $\bar{t}$ = the mean start time of the upper harmonics, n = the number of harmonics in the tone, AMP(k) = the energy in harmonic k, and TOTAMP = the total amount of energy in the tone. The start of a harmonic was defined to be the time at which its amplitude was 1.5 (about 20 dB). These weighted standard deviations correlate highly ($r = -0.78$) with the Y coordinates of the INDSCAL solution (Figure 21).

A comparison of the two dimensional INDSCAL solution with a two dimensional projection of Grey's Y and X dimensions (Figure 22) reveals that these two scaling solutions are very similar. This supports the position that the scaling solution is a reasonable one. However, there is a puzzling discrepancy between Grey's interpretation of the synchronous-asynchronous dimension and the one just offered. While both studies obtained this dimension, Grey claimed that the woodwinds tend to be sychronous while the strings and brasses tend to be asynchronous. This study

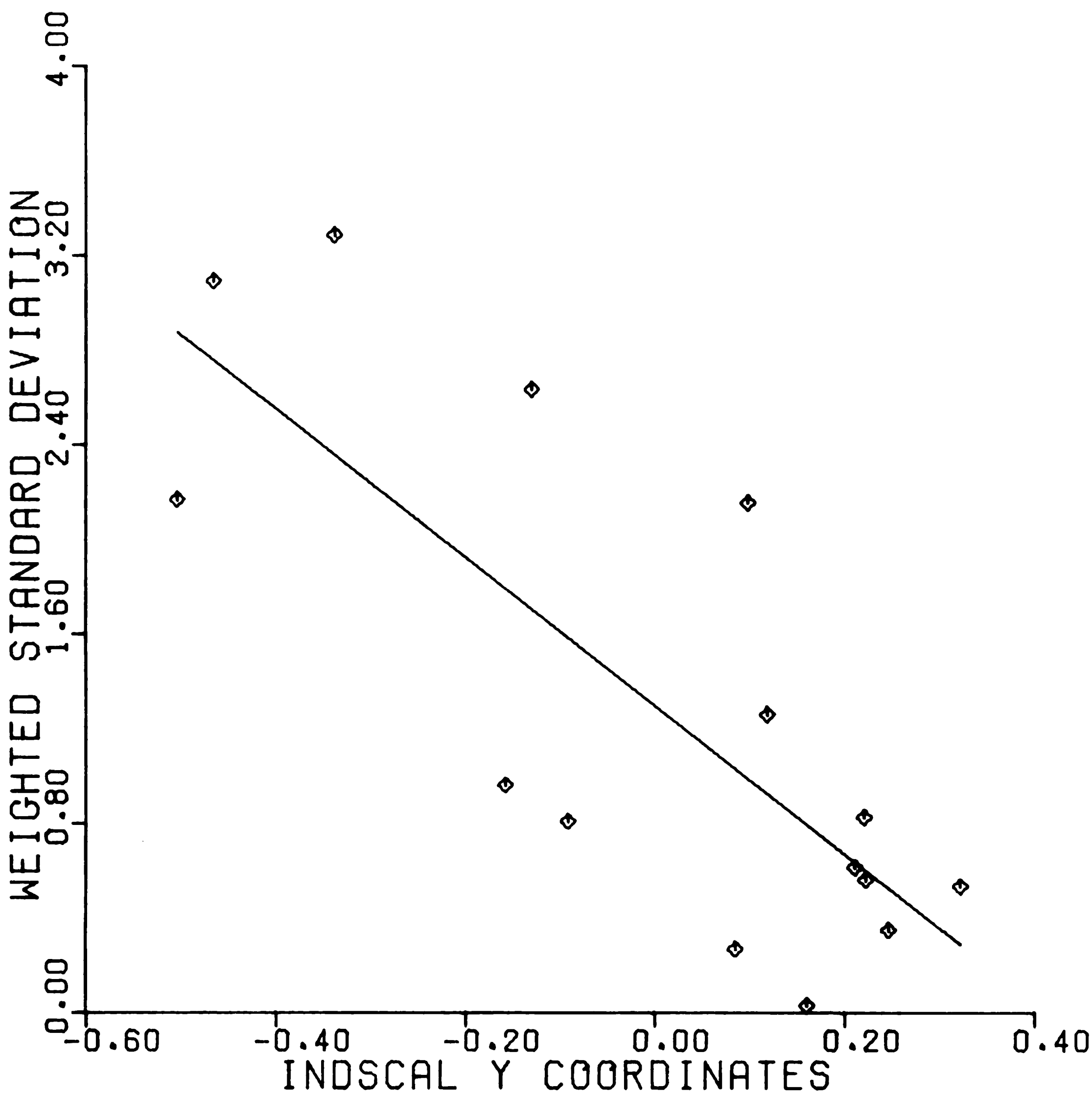

Figure 21. INDSCAL Y coordinates versus a weighted standard deviation of the onset times of the upper harmonics.

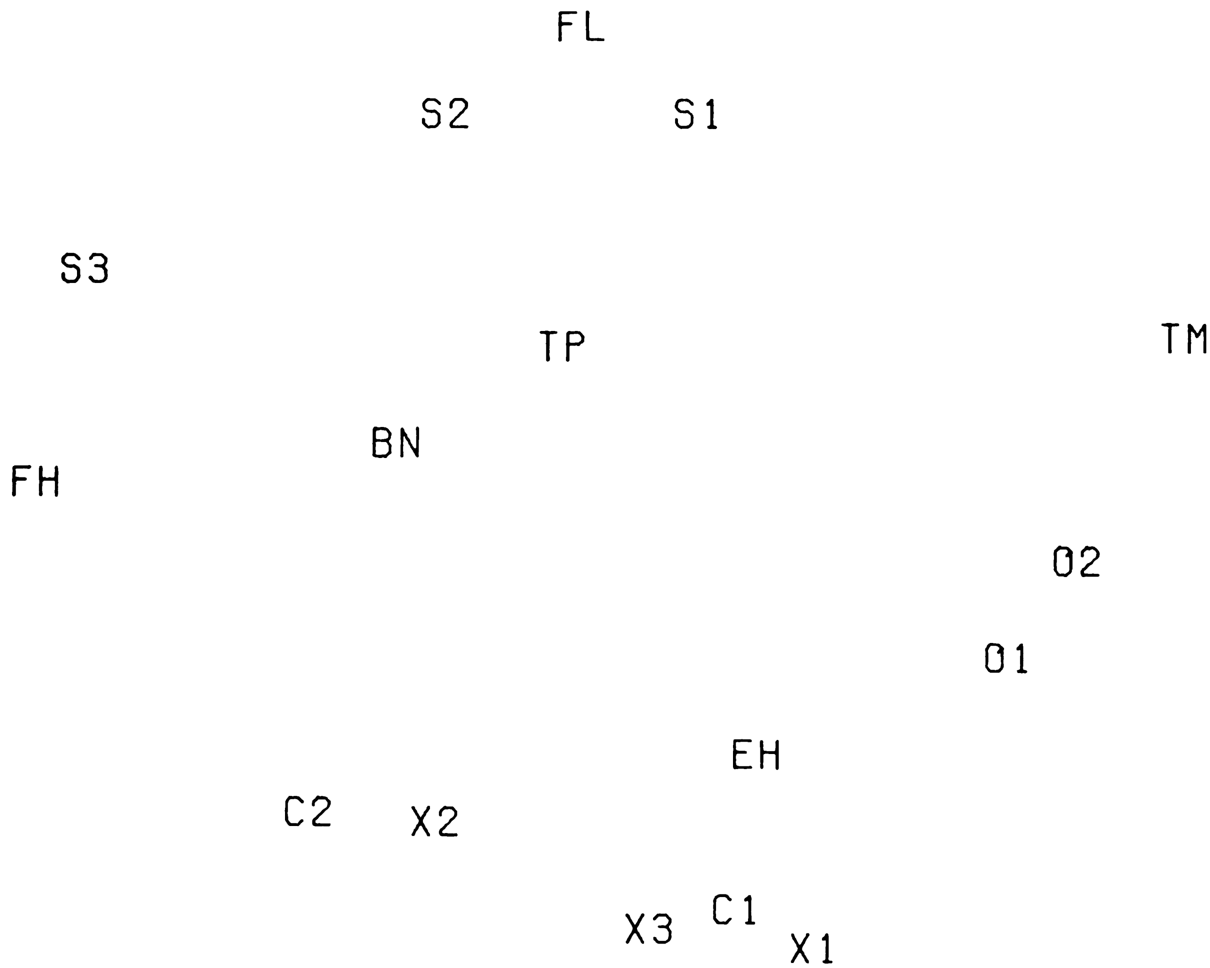

Figure 22. A two dimensional subspace of Grey's (1975) three dimensional INDSCAL timbre space.

claims exactly the opposite, that the woodwinds tend to be asynchronous while the strings and brasses tend to be synchronous. It is important to note that Grey arrived at his interpretation by looking at spectrograms of his tones but failed to give any quantitative measure to support his interpretation.

One possible problem with the quantitative interpretation provided here is that some of the harmonics have frequency glides. That is, at the start of the tones some frequencies are not at their proper harmonic values but glide to that value as the tone progresses. Thus a second way of defining the start of a harmonic is that time at which the frequencies reach some proportion of their proper harmonic value. Exploration with this definition of start time failed to provide as good a quantitative interpretation as did the amplitude definition of start time. The modified measure also failed to resolve the discrepancy since the strings and brasses still tended to be synchronous while the woodwinds tended to be asynchronous.

The start of an harmonic could also be defined to be that time at which its frequency reaches some proportion of its proper value and its amplitude exceeds some threshold. Exploration with this measure also failed to equal or improve upon the original measure, and again the discrepancy remained unresolved.

As can be seen from the subjects' weight space shown in Figure 18, there were no systematic differences between the musically sophisticated subjects and those with no musical training. Three of the untrained subjects and the five musical subjects clustered together at a point which indicates that they were giving an approximately equal weight to both dimensions. Subject U5 gave less weight to the first dimension that did the others and subject U4 gave little weight to either dimension. This conclusion is further substantiated by the two KYST solutions (Figure 23). Not only are the stress values very nearly the same, but the solutions themselves are similar to each other. Note also that the two dimensional KYST solutions are roughly comparable to the INDSCAL solution. These convergent solutions serve to increase our confidence in the interpretations that have been provided.

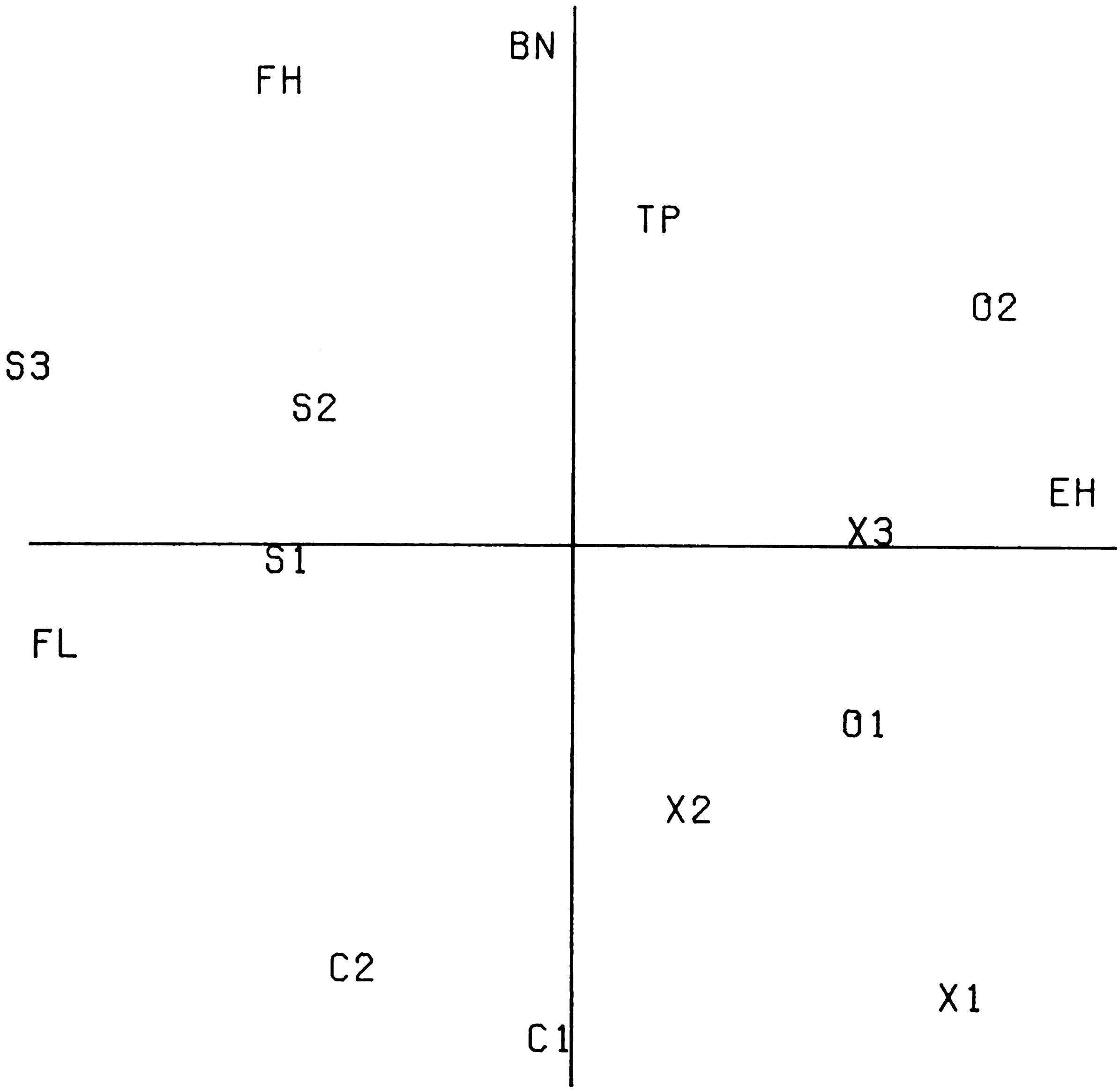

Figure 23a. Two dimensional KYST timbre space for musically sophisticated subjects.

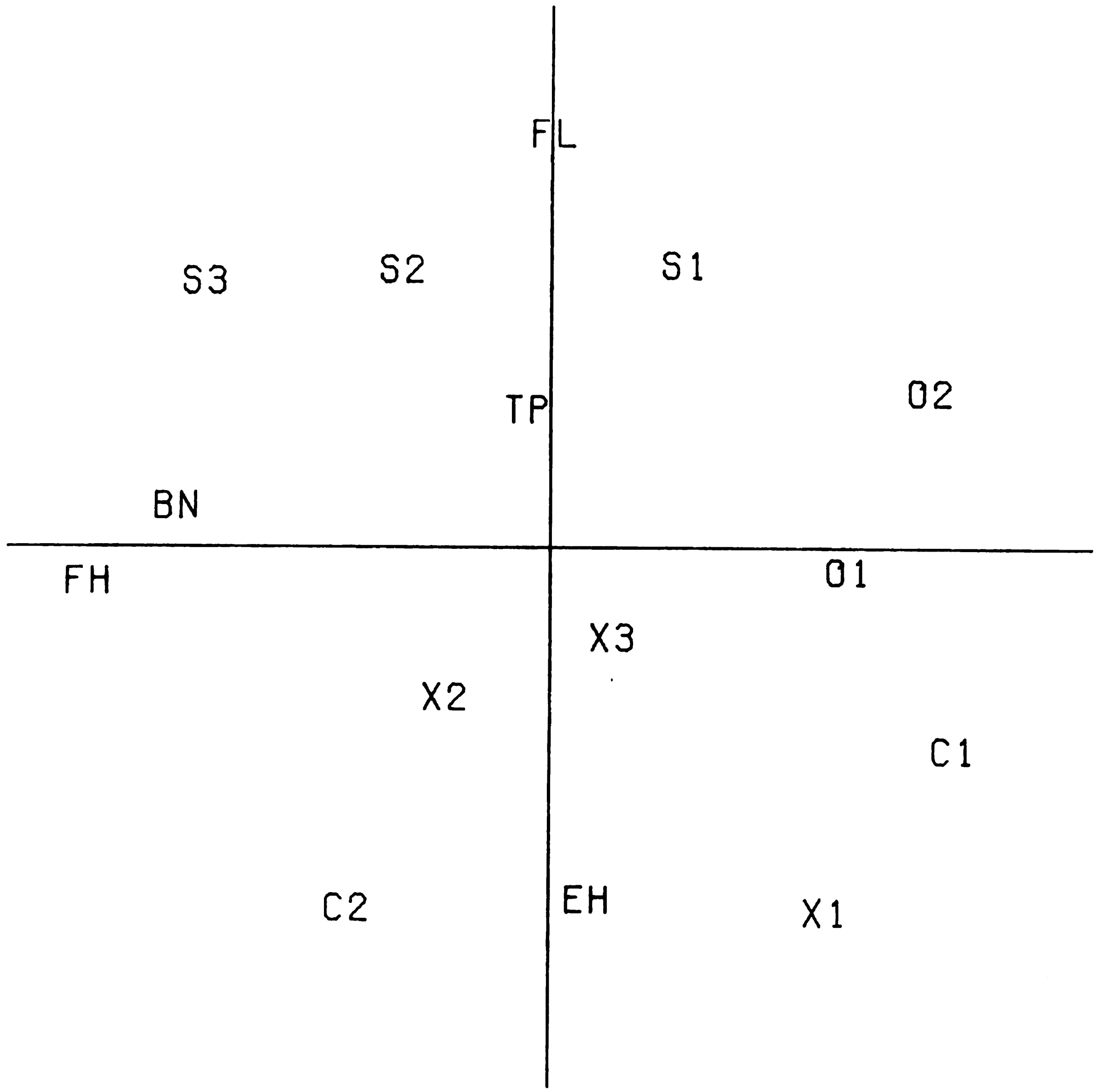

Figure 23b. Two dimensional KYST timbre space for musically untrained subjects.

EXPERIMENT 2: A TEST OF THE PARALLELOGRAM MODEL

Stimuli

In this experiment, Rumelhart and Abrahamson's (1973) parallelogram model of analogical reasoning was tested using the timbre space derived in Experiment 1. Twenty timbre analogies of the form A:B::C:D1, D2, D3, or D4 were formed as follows: The 15 tones were arranged in random order; the first three in this order were chosen as A, B, and C of the first analogy, the second three were used to form the second analogy, and so on. When the list was exhausted the tones were rerandomized and the procedure repeated until 20 analogies had been formed. For each of the analogies thus formed, the coordinates of the ideal analogy point, "I", were calculated and the distances between each of the remaining 12 tones and "I" were computed. Four alternative solutions to the analogy were chosen such that each analogy had an alternative in each of the following ranges: 0.00-0.25 units from "I", 0.25-0.50 units from "I", 0.50-0.75 units from "I", and 0.75-1.00 units from "I". The units which were used are the ones produced by the INDSCAL program as shown in Figure 19. If it was not possible to choose alternatives to meet these conditions, that analogy was discarded and another one formed as above. If more than one tone fell within a given range, the one closest to the lower boundary was chosen. The four chosen alternatives were also ordered randomly for each subject.

If A is to B as C is to D, then it ought to be the case that A is to C as B is to D. In terms of the parallelogram model, the above two analogies are exactly the same parallelogram. Therefore, for each of the 20 analogies formed above, another analogy was formed which had the same components but had the second and third elements reversed. In other words, the analogies were of the form A:C::B:D1, D2, D3, or D4.

Procedure

The analogies were presented to subjects using the audio equipment described in Experiment 1. A trial consisted of the four alternative forms of an analogy (A:B:C:Dn, where Dn is one of the four alternative solutions). Subjects selected one of the alternative forms by depressing one of four button switches; each alternative form was randomly associated with one switch. Subjects listened to each alternative as many times as he or she wished. Subjects were asked to rank order each of the four alternatives as to how well it completed the analogy. The rank order was indicated by rearranging the order of the four button switches until they were in the appropriate order. Subjects then pressed a fifth switch which signaled to the computer that the rank order was ready to be entered; subjects then pressed each of the four switches in the appropriate order and this was read by the PDP-11/40 via the DR11C interface and was stored on a magnetic disk for later analysis.

Subjects

Nine of the ten people who served as subjects in Experiment 1 served as subjects in this experiment. Subject M3 was unable to participate.

Results and discussion

The parallelogram model predicted subjects' responses on the timbre analogy task better than any of the alternative hypotheses. Of the alternative hypotheses, the potent dimension similarity of terminal tones model was the least satisfactory. This model predicts that the alternative D closest to the tone B along the potent dimension will be chosen as the best solution. However, the correlation between the one dimensional BD distance and the proportion of subjects who ranked that particular alternative as the best solution was quite low ($r = -0.31$).

The other two alternative hypotheses were slightly better. The potent dimension parallelogram model predicts that the alternative D which is chosen as the best solution to an analogy is that one with the shortest ID distance along the potent dimension. In this case, the one dimensional ID distances correlated poorly with the proportion of subjects who ranked that alternative D as the best solution ($r = -0.39$). The prediction of the similarity of terminal tones model is that the alternative D closest to tone B in the multidimensional space will be chosen as the best solution to an analogy. The correlation between the BD distances and the distances and the proportions of subjects chosing a particular D as the best solution was the same as the previous model (r = -0.39).

The parallelogram model predicts that the probability of chosing a given alternative as the best solution is inversely related to the ID distance. Table 3 lists the

Table 3. Rank order data averaged over all subjects and all analogies.

RANK DISTANCE OF THE ALTERNATIVE FROM I	SUBJECT-ASSIGNED RANK 1	2	3	4
1	0.422	0.303	0.156	0.119
2	0.322	0.283	0.217	0.178
3	0.169	0.267	0.358	0.206
4	0.086	0.147	0.269	0.497

proportion of responses, averaged over subjects and analogies, for which the Ith closest alternative to the ideal analogy point was ranked as the Jth best solution, where I is the row index and J is the column index. Column one of this table shows that the prediction was indeed fulfilled. In fact, the distance between an alternative and the ideal analogy point predicts not only the best solution, but the rank ordering of all four alternatives. Only one exception occurred (see column two). Figure 24 shows the scatter diagram of the ID distance versus the proportion of subjects who ranked that alternative D as the best solution. The product moment correlation coefficient is -0.52. This correlation is not too disappointing when one takes into consideration that the goodness of fit measure (r = 0.62) of timbre space places a kind of ceiling on subsequent

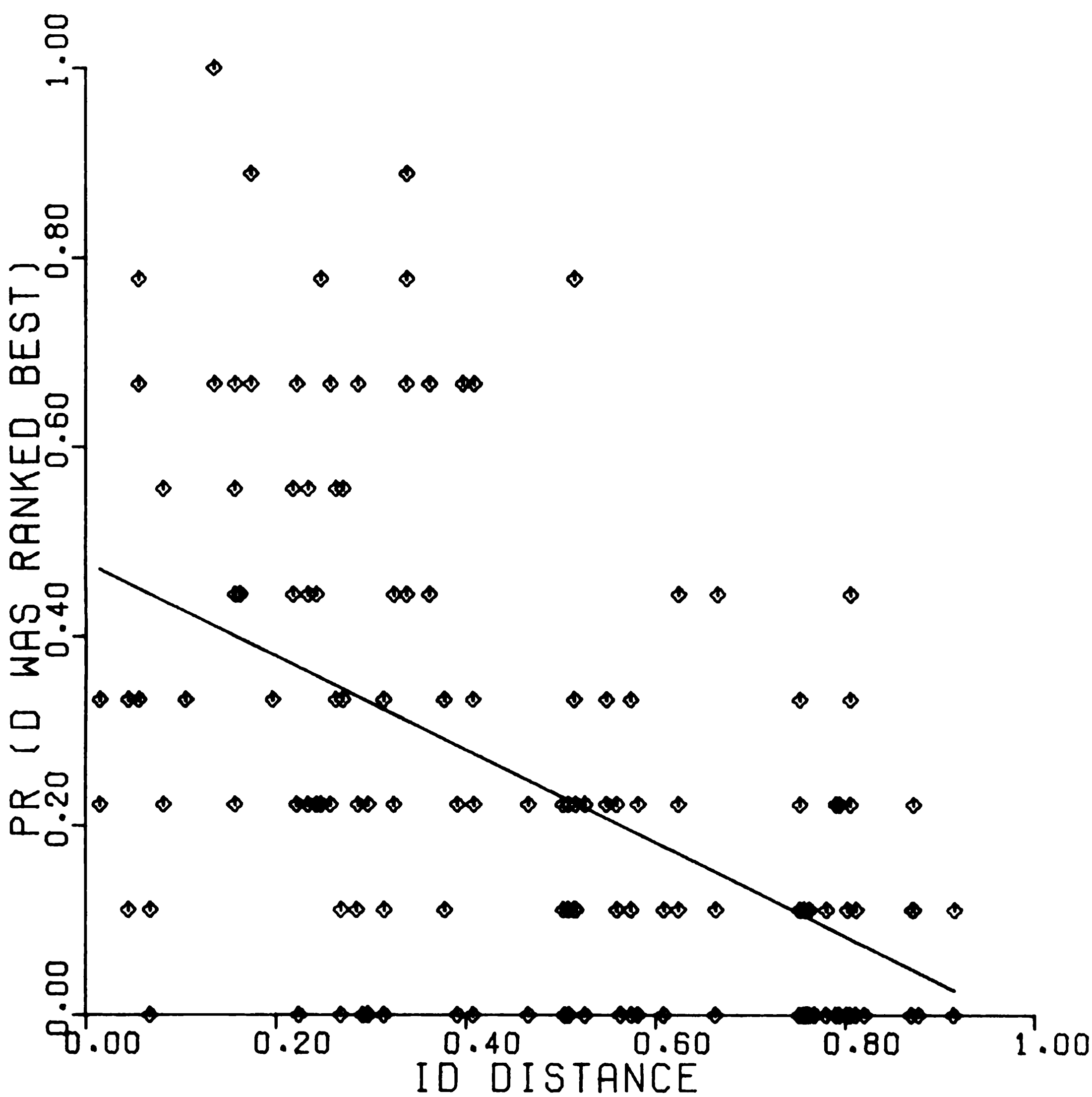

Figure 24. ID distance versus the proportion of subjects who ranked an alternative as the best solution.

correlations derived from the distances in a way analogous to the limit reliability places on validity in testing theory.

By assuming Luce's choice rule (Luce, 1959; Atkinson, Bower, and Crothers, 1965) the probability of choosing a given alternative as the best solution can be predicted. Luce's choice rule states that the probability of choosing any given alternative X_i from the set of alternatives X_1, . . . Xn is given by

$$Pr(X_i|X_1, \ldots, Xn) = p_i = v(d_i)/[\Sigma v(d_j)], \tag{5}$$

where d_i = distance between X_i and I, and v (d_i) is a monotonically decreasing function of its argument. The additional assumption was made that

$$v(x) = \exp(-\alpha x), \tag{6}$$

where x and α are positive numbers. The exponential function was chosen for two reasons: (1) Shepard (1957) found a good fit to an exponential generalization function over a similarly derived space and (2) Rumelhart and Abrahamson found a good fit to an exponential function in their work with the parallelogram model. Taking the natural logarithm of both sides of equation 5 yields

$$\ln(p_i) = \ln v(d_i) - \ln[\Sigma v(d_j)]. \tag{7}$$

Substituting exp $(-\alpha x)$ for v(x) (equation 6), we get

$$\ln p_i = -\alpha d_i - \ln[\Sigma \exp(-\alpha d_j)]. \tag{8}$$

This function states that the parameter, $-\alpha$, can be estimated by the slope of the regression line fit to the ID distance versus the natural logarithm of the proportion of subjects who ranked that alternative as the best solution (Figure 25). Alternatives which were never chosen as the best solution were eliminated while making this calculation. This yielded an estimate of $-\alpha = -1.33$. This parameter was then used to predict the proportion of subjects who ranked each alternative as the best solution. Figure 26 shows the observed versus the predicted proportion of subjects who ranked each alternative as the best solution. Again this correlation ($r = 0.52$) is quite acceptable when compared to the goodness of fit measure.

Given that there is no systematic clustering in the INDSCAL subjects' weight space, it is not surprising that there are no systematic differences in the analogy judgments attributable to musical training. Table 4 gives the proportion of responses, averaged over the musical subjects and the analogies, for which the Ith closest alternative to I was ranked as the Jth best solution. Table 5 gives these same results averaged over the musically untrained subjects. As one would expect, there were no systematic differences between these two tables and their entries were highly correlated ($r = 0.89$).

If the analogies work equally well along both dimensions then the analogies of the form A:C::B:D should work as well as those of the form A:B::C:D. Table 6 and Table 7 summarize the data for these two forms of the analogies averaged over all subjects. Since the subjects' weight space (Figure 18)

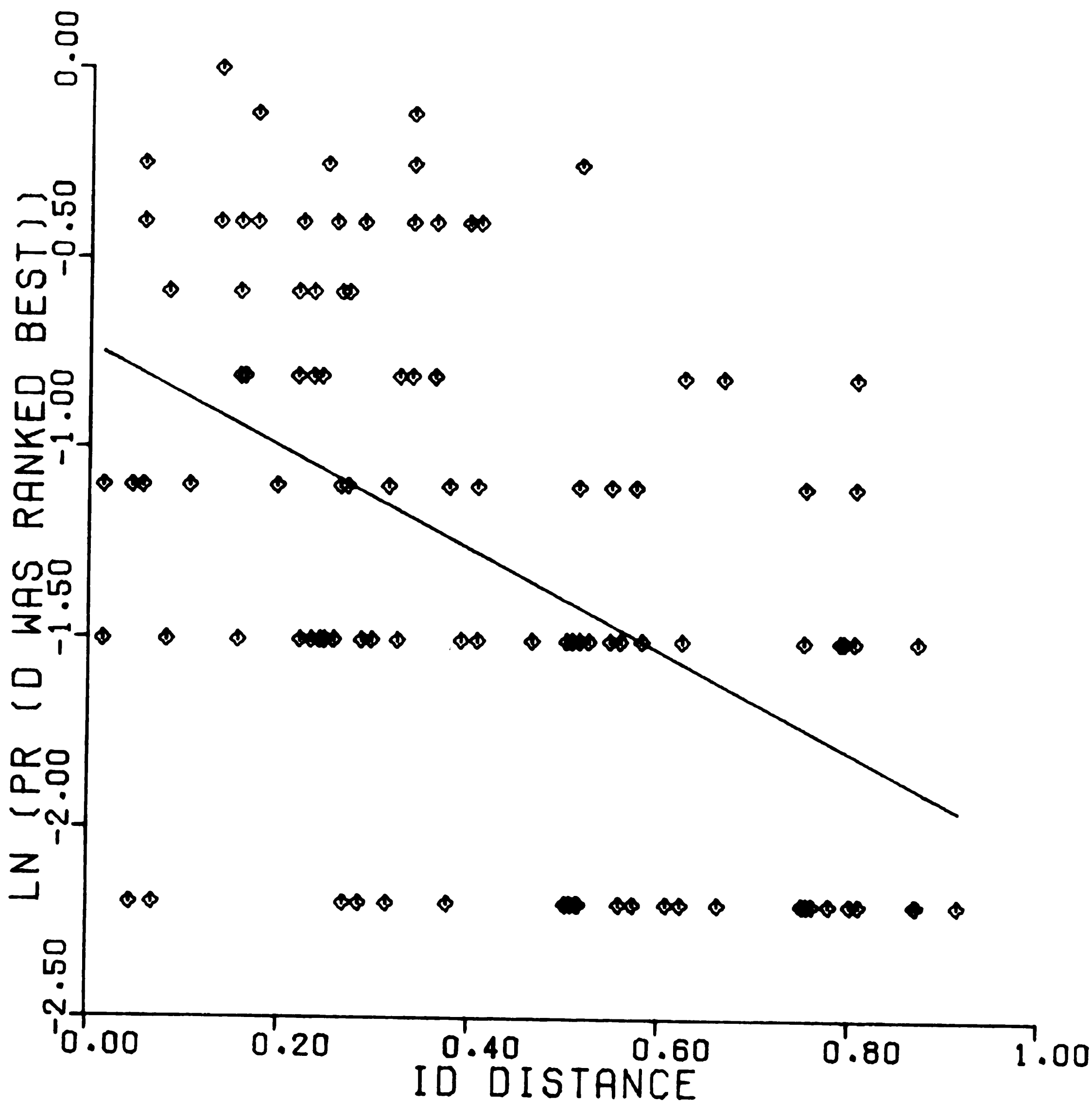

Figure 25. ID distance versus the logarithm of the proportion of subjects who ranked an alternative as the best solution.

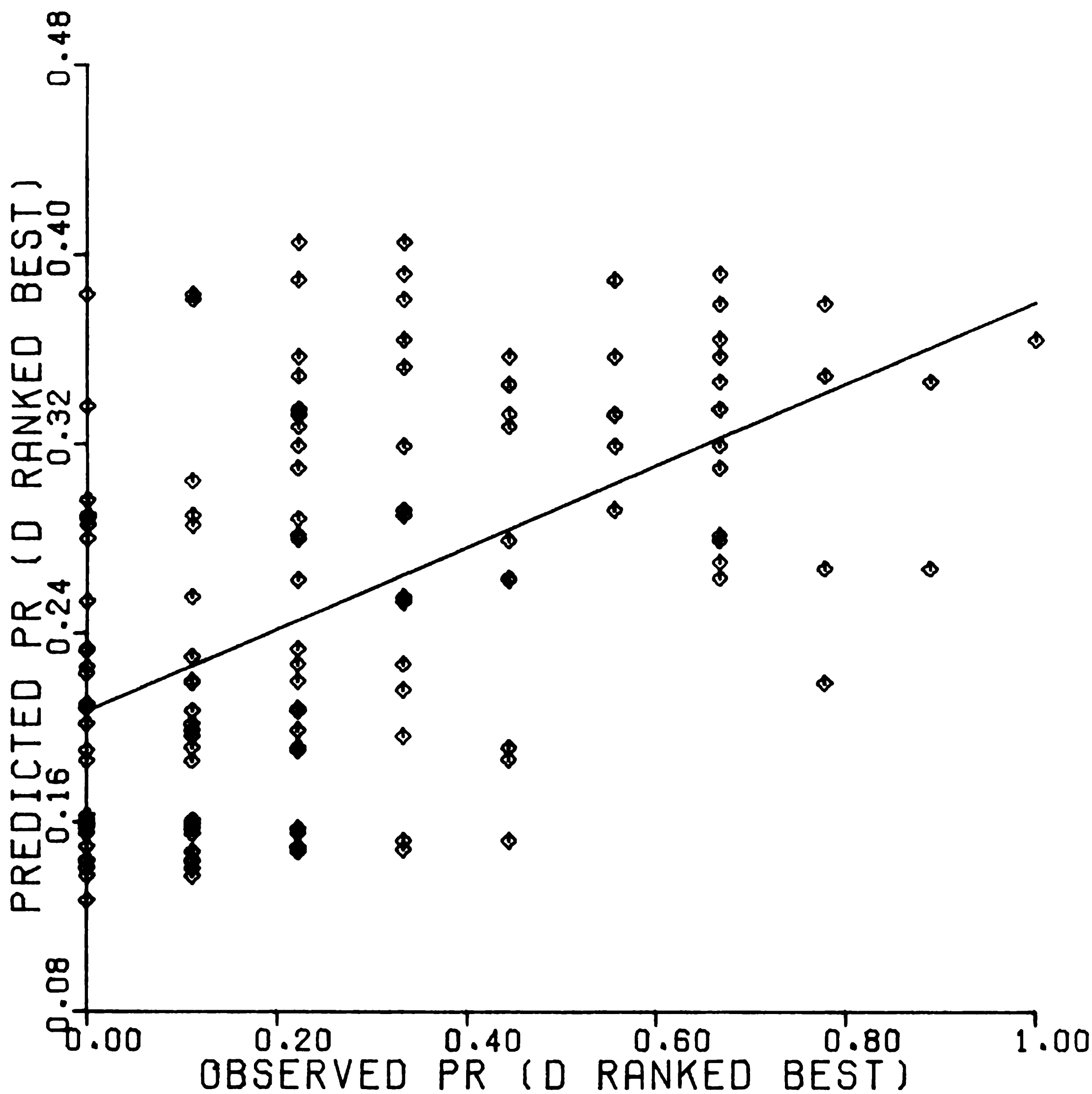

Figure 26. Observed versus predicted proportion of subjects who ranked each alternative as the best solution.

Table 4. Rank order data for musically sophisticated subjects.

RANK DISTANCE OF THE ALTERNATIVE FROM I	SUBJECT-ASSIGNED RANK 1	2	3	4
1	0.406	0.331	0.131	0.131
2	0.325	0.294	0.219	0.162
3	0.213	0.187	0.375	0.225
4	0.056	0.187	0.275	0.481

Table 5. Rank order data for musically untrained subjects.

RANK DISTANCE OF THE ALTERNATIVE FROM I	SUBJECT-ASSIGNED RANK 1	2	3	4
1	0.435	0.280	0.175	0.110
2	0.320	0.275	0.215	0.190
3	0.135	0.330	0.345	0.190
4	0.110	0.115	0.265	0.510

Table 6. Rank order data from analogies of the form A:C::B:D.

RANK DISTANCE OF THE ALTERNATIVE FROM I	SUBJECT-ASSIGNED RANK			
	1	2	3	4
1	0.389	0.344	0.167	0.100
2	0.361	0.283	0.200	0.156
3	0.189	0.239	0.367	0.206
4	0.061	0.133	0.267	0.539

Table 7. Rank order data from analogies of the form A:B::C:D.

RANK DISTANCE OF THE ALTERNATIVE FROM I	SUBJECT-ASSIGNED RANK			
	1	2	3	4
1	0.456	0.261	0.144	0.139
2	0.283	0.283	0.233	0.200
3	0.150	0.294	0.350	0.206
4	0.111	0.161	0.272	0.456

shows that subjects tend to put nearly equal weights on both dimensions, it is not surprising that these two tables are quite similar ($r = 0.92$).

CONCLUSION

The principal dimension of the timbre space which was obtained in the first experiment replicated the results of previous studies. This not only supports the idea that a reasonable solution was obtained, but also increases one's confidence that the energy distribution provides an important cue in the perception of timbre.

The fact that the second dimension was interpreted in terms of the attack transient also agrees with previous research. However, the details of exactly what property (or properties) of the transient provides the perceptual cues are not totally clear. An experiment comparing three variations of stimulus tones could provide further insight about this problem.

This experiment involves three conditions. The first condition would repeat the scaling part of this study with the tones experimentally equalized for pitch, loudness, and duration. The tones scaled in the second condition would be synthesized by replacing the frequency functions used in the first condition with frequency functions which are exactly at the proper harmonic value, thus removing the frequency glides as perceptual cues. The tones in the third condition would be synthesized by replacing the amplitude functions of the original tones with trapezoidal functions that have the same average amplitude as the origninal function. In this condition, the amplitude envelope cannot provide any perceptual cues. A separate dissimilarity matrix and scaling solution would be found for each condition.

By comparing the scaling solutions from the fixed frequency and the trapezoidal amplitude tones to the original tones solution, one will be able to better understand what effect the amplitude variations and the frequency glides have on the perception of timbre. If the fixed frequency scaling solution corresponds more closely to the original solution than the trapezoidal amplitude solution, one will be able to conclude that the perceptual cues provided by the amplitude envelope are more important in the perception of timbre than those cues provided by the frequency glides. This would be consistent with the synchronous-asynchronous interpretation of the second dimension. However, if the trapezoidal amplitude scaling solution is better than the fixed frequency solution, one will conclude that the frequency glides provide important cues in the perception of timbre. This would require that the synchronous-asynchronous interpretation be modified or rejected.

Another method of evaluating the validity of the interpretations that have been given to timbre spaces is to manipulate directly the properties of the tone. If the centroid of the energy distribution and the variability of the onset times in the upper harmonics play the important role which this study suggests, it should be possible to synthesize tones that map onto a predetermined area of the timbre space. Such results would provide additional strong support for this interpretation.

Although there is still room for improvement, the parallelogram model predicts the solutions to timbre analogies more accurately than alternative models. Furthermore, the model accomplishes this even though the particular tones in the present study were only approximately equalized for pitch, loudness, and duration. Although one rarely hears precisely equalized tones in real life situations, a replication with tones that have been equalized experimentally is still desirable.

An even more important step would be to carry out an analogy experiment based on a more orderly timbre space. Assuming the interpretation of the timbre space is correct, it should be possible to construct a space where tones actually occur at the ideal analogy point. Thus it would be possible to get a better test of the parallelogram model.

These results open interesting and challenging avenues for composers and musicians. The concept of timbre analogies suggests that the idea of melodic transposition might now be extended from the domain of pitch to that of timbre.

REFERENCES

REFERENCES

American Standards Association. American standard acoustical terminology (S1.1-1960). New York: American Standards Association, Inc., 1960.

Atkinson, R. D., Bower, G. H., and Crothers, E. J. Introduction to mathematical learning theory. New York: Wiley, 1965.

Bartholomew, W. T. Acoustics of music. New York: Prentice-Hall, Inc., 1945.

Berger, K. W. Some factors in the recognition of timbre. Journal of the Acoustical Society of America, 1964, 36, 1888-1891.

Carroll, J. D., and Chang, J. J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika, 1970, 35, 283-319.

Fletcher, H. Loudness, pitch and the timbre of musical tones and their relation for the intensity, the frequency and the overtone structure. Journal of the Acoustical Society of America, 1934, 6, 59-69.

Grey, J. M. An exploration of musical timbre. Unpublished doctoral dissertation, Stanford University, 1975.

Helmholtz, H. L. F. [On the sensations of tone as a physiological basis for the theory of music] (A. J. Ellis, trans.). New York: Dover, 1954. (Originally published, 1877.)

Henley, N. M. A psychological study of the semantics of animal terms. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 176-184.

Kriz, J. S. A 16-bit A-D-A conversion system for high-fidelity audio research. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1975, 23, 146-149.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 1964a, 29, 1-27.

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. Psychometrika, 1964b, 29, 115-129.

Luce, R. D. Individual choice behavior: a theoretical analysis. New York: Wiley, 1959.

Moorer, J. A. The heterodyne filter as a tool for analysis of transient waveforms (Tech. Rep. CS-379). Stanford University, Computer Science Department, July 1973.

Ohm, G. S. Uber die definition des tones, nebst daran geknupfter theorie der sirene und ahnlicher tonbildender vorrichtungen. Annalen der Physik und Chemie, 1843, 59, 513-565.

Plomp, R., and Steeneken, H. J. M. Effects of phase on the timbre of complex tones. Journal of the Acoustical Society of America, 1969, 46, 409-421.

Rumelhart, D. E., and Abrahamson, A. A. A model for analogical reasoning. Cognitive Psychology, 1973, 5, 1-28.

Saldanha, E. L., and Corso, J. F. Timbre cues for recognition of musical instruments. Journal of the Acoustical Society of America, 1964, 36, 2021-2026.

Shepard, R. N. Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. Psychometrika, 1957, 22, 325-345.

Shepard, R. N. Analysis of proximities: Multidimensional scaling with an unknown distance function I. Psychometrika, 1962 a, 27, 125-140.

Shepard, R. N. Analysis of proximities: Multidimensional scaling with an unknown distance function II. Psychometrika, 1962b, 27, 219-246.

Slawson, A. W. Vowel quality and musical timbre as functions of spectrum envelopes and fundamental frequency. Journal of the Acoustical Society of America, 1968, 43, 87-101.

Strong, W., and Clark, M. Perturbations of synthetic orchestral wind-instrument tones. Journal of the Acoustical Society of America, 1967a, 41, 277-285.

Torgerson, W. S. Theory and methods of scaling. New York: Wiley, 1958.

Wedin, L., and Goude, G. Dimensional analysis of the perception of instrumental timbre. Scandinavian Journal of Psychology, 1972, 13, 228-240.

Wessel, D. L. Psychoacoustics and music. PAGE: Bulletin of the Computer Arts Society, 1973, 30.

Wessel, D. L., and Grey, J. M. Conceptual structures for the representation of musical material. In H. B. Lincoln (Ed.), Composition and sound synthesis with computers. Cornell University Press, in press.

Young, F. W., and Torgerson, W. S. TORSCA: A Fortran-4 program for Shepard-Kruskal multidimensional scaling analysis. Behavioral Science, 1967, 12, 498.

APPENDIX

IADD: An Interpolating Additive Synthesis Program

```
C   PROGRAM "IADD" WRITTEN BY  DAVID EHRESMAN
C                              DEPARTMENT OF PSYCHOLOGY
C                              MICHIGAN STATE UNIVERSITY
C                              EAST LANSING, MI   48824
C
C   THIS PROGRAM, WHICH HAS BEEN IMPLEMENTED ON A PDP-11/40
C   MINI-COMPUTER AT MICHIGAN STATE, DOES ADDITIVE SYNTHESIS
C   USING AN INTERPOLATING OSCILLATOR.  THE CONTROL PARAMETERS
C   ARE SAMPLING RATE, BEATS PER SECOND, INITIAL PHASE, AND
C   SCALE FACTOR.  THE ADDITIVE SYSTHESIS DONE BY THIS PROGRAM
C   IS BASED ON STRAIGHT LINE APPROXIMATIONS OF THE COMPLEX
C   TIME VARYING AMPLITUDE AND FREQUENCY FUNCTIONS OF EACH
C   HARMONIC TO BE INCLUDED IN THE SYNTHESIZED TONE.  THIS
C   PROGRAM USES THE BREAKPOINT INFORMATION FROM THESE
C   FUNCTIONS TO CONTROL THE SYNTHESIS.  THE INPUT FILE MUST
C   BE IN THE FOLLOWING FORMAT:  LINE 1-AMP. LABEL
C   (MAX. = 8 CHAR.); LINE 2-TIME OF FIRST AMP. BREAKPOINT
C   (I5) FOR HARMONIC 1; LINE 3-AMPLITUDE AT FIRST BREAKPOINT
C   (F10.0) FOR HARMONIC 1.  THIS IS REPEATED UNTIL ALL THE
C   AMPLITUDE DATA FOR THE 1ST HARMONIC HAS BEEN ENTERED.
C   THE NEXT LINE MUST CONTAIN 999 WHICH ACTS AS A DELIMITER.
C   THIS IS FOLLOWED BY THE FREQUENCY (HZ) BREAKPOINT DATA
C   FOR HARMONIC 1 USING THE SAME FORMAT AS ABOVE.  DATA FOR
C   EACH OF THE REMAINING HARMONICS MUST BE ENTERED USING THE
C   SAME FORMAT.  THIS PROGRAM CAN PROCESS A MAX. OF 29
C   HARMONICS WITH A MAX. OF 19 BREAKPOINTS/HARMONIC.  THE
C   OUTPUT IS STORED ON A DISK IN A FORMAT SUITABLE FOR
C   READING THRU A DAC.
C
C   THE FOLLOWING SUBROUTINES AND FUNCTIONS ARE NEEDED TO RUN
C   THIS PROGRAM:
C   (1) LINE - COMPUTES SLOPE AND CONSTANT FOR A LINE DEFINED
C              BY TWO POINTS
C   (2) OSCIL - COMPUTES A SAMPLE WHEN GIVEN AMP., SAMPLE
C               INCREMENT, AND PHASE INFORMATION
C   (3) SOUT - PACKS SAMPLES IN A FORMAT SUITABLE FOR USE
C              WITH OUR DAC
C   (4) SYSLIB ROUTINES
```

```
C
C----------------------------------  IADD   ----------------------------------C
C
      DOUBLE PRECISION EXT
      DIMENSION AMP (20,30), ITIME1 (20,30), FREQ (20,30),
     $ I1 (30), I2 (30), ITIME2 (20,30), SLOPE1(30),
     $ SLOPE2(30), CONST1 (30),CONST2 (30), PHS (30),
     $ ISPEC (39)
      COMMON S,SINE (511)
     $ /BUFFER/ ICHAN, INDEX, AMAX, SAMP (256), IBUFF (256)
      DATA INDEX /0/, AMAX /0.0/,EXT /6RDATSND/,
     $ SAMP /256*0.0/, IBEL /"007/
C**    READ CONTROL PARAMETERS
      WRITE (7,60)
60     FORMAT ('$','ENTER SAMPLING RATE (F10.0) - ')
      READ (5,120) SAMRAT
      WRITE (7,70)
70     FORMAT ('$','ENTER BEATS PER SECOND (F10.0) - ')
      READ (5,120) BEAT
      WRITE (7,75)
75     FORMAT ('$','ENTER INITIAL PHASE (F10.0) - ')
      READ (5,120) PHASE
C**    FIGURES SAMPLES PER BEAT
      SPB = SAMRAT/BEAT
      WRITE (7,76)
76     FORMAT ('$','ENTER AMPLITUDE SCALE FACTOR (F10.0) - ')
      READ (5,120) SCALE
C**    READ OUTPUT & INPUT FILE NAMES IN STANDARD CSI
C**    FORMAT AND OPEN FILES FOR I/O
      WRITE (5,79)
79     FORMAT ('$','ENTER COMMAND STRING - ')
      IF (ICSI(ISPEC,EXT,,,0).NE.0) STOP 'INVALID CSI STRING'
      IF (IASIGN(10,ISPEC(16),ISPEC(17),0,32).NE.0)
     $  STOP 'NO CHANNEL FOR INPUT'
      ICHAN = IGETC ()
      IF (ICHAN.LT.0) STOP 'NO CHANNEL FOR OUTPUT'
      IF (IENTER(ICHAN, ISPEC (1),ISPEC (5)).LT.0)
     $  STOP 'NO CHANNEL OR NOT ENOUGH DISK SPACE'
C**    STORE A 512 SAMPLE SINE WAVE
      TEMP = (2. * 3.14159265)/511.
      DO 80 J = 0,511
           SINE (J) =  SIN (TEMP*J)
80    CONTINUE
C**    LOOP FOR MAX. OF 29 HARMONICS
      DO 190 J = 1,30
C**         READ AMP. BREAKPOINT DATA FOR HARMONIC J
           READ (10,100,END=200) LABEL
100         FORMAT (A12)
           DO 130 I = 1,20
                READ (10,110) ITIME1 (I,J)
                IF (ITIME1 (I,J).EQ.999) GO TO 150
                READ (10,120) AMP (I,J)
110              FORMAT (I5)
120              FORMAT (F10.0)
C**               CONVERT BEATS TO SAMPLE NUMBER
                ITIME1 (I,J) = (ITIME1 (I,J)-1) * SPB + 1
130        CONTINUE
```

```
C**          READ FREQ. BREAKPOINT DATA FOR HARMONIC J
150          READ (10,100) LABEL
             DO 175 I = 1,20
                  READ (10,110) ITIME2 (I,J)
                  IF (ITIME2 (I,J).EQ.999) GO TO 190
                  READ (10,120) FREQ (I,J)
C**                CONVERT BEATS TO SAMPLE NUMBER &
C**                HERTZ TO SINE TABLE INCREMENT
                  ITIME2 (I,J) = (ITIME2 (I,J)-1) * SPB + 1
                  FREQ (I,J) = FREQ (I,J) * 511. / SAMRAT
175          CONTINUE
190     CONTINUE
C**      COMPUTE & PRINT NUM. OF HARMONICS & LENGTH OF FILE
200     NHAR = J-1
        ILEN = ITIME2 (I-1,NHAR)
        WRITE (7,225) NHAR,(ILEN-1)/SAMRAT
225      FORMAT ('0',T5,'NUMBER OF HARMONICS = ',I2/
     $ ' ',T5,'LENGTH OF FILE = ',G10.4,' SECONDS')
C**       FIND INITIALIZE SLOPE & Y INTERCEPT FOR AMP. AND
C**       FREQ. FUNCTIONS
        DO 250 J = 1,NHAR
             I1(J) = 1
             I2(J) = 1
             CALL LINE (I1(J),J,AMP,ITIME1,SLOPE1(J),CONST1(J))
             CALL LINE (I2(J),J,FREQ,ITIME2,SLOPE2(J),CONST2(J))
             PHS (J) = PHASE
250     CONTINUE
C**      COMPUTE SAMPLES
        DO 400 K = 1,ILEN
             L = MOD (K-1,256)+1
             DO 300 J = 1,NHAR
C**                UPDATE AMP.SLOPE AND CONSTANT AT EACH
C**                BREAKPOINT & COMPUTE AMP.
                  IF (K.EQ.ITIME1 (I1(J),J) + 1)
     $                 CALL LINE (I1(J),J,AMP,ITIME1,SLOPE1(J),
     $                            CONST1(J))
                  SAMAMP = SLOPE1 (J) * K + CONST1 (J)
C**                UPDATE FREQ. SLOPE AND CONSTANT AT EACK
C**                BREAKPOINT & COMPUTE SINE TABLE INCREMENT
                  IF (K.EQ.ITIME2 (I2(J),J) + 1)
     $                 CALL LINE (I2(J),J,FREQ,ITIME2,SLOPE2(J),
     $                            CONST2(J))
                  SI = SLOPE2 (J) * K + CONST2 (J)
C**                COMPUTE SAMPLE
                  SAMP (L) = SAMP (L) + OSCIL (SAMAMP,SI,PHS(J))
300          CONTINUE
C**           SCALE SAMPLE AND WRITE TO DISK IN BLOCKS OF
C**           256 SAMPLES
             SAMP (L) = SAMP (L) * SCALE
             IF (L.EQ.256) CALL SOUT
400     CONTINUE
        CALL SOUT
```

```
C**    PRINT ENDING MESSAGE, CLOSE I/O CHANNELS & RING
C**    DEC-WRITER BELL
      WRITE (7,500) AMAX
500    FORMAT ('0','MAX AMPLITUDE IS ',1PG20.10/
     $          ' ','LET''S SING')
      CALL IWAIT (ICHAN)
      CALL CLOSEC (ICHAN)
      WRITE (7,600) IBEL,IBEL
600    FORMAT ('+',2A1)
      STOP
      END
C
C**     SUBROUTINE LINE
C
C  LINE RETURNS THE SLOPE & Y INTERCEPT OF A LINE WHEN GIVEN
C  TWO POINTS THAT DEFINE THAT LINE
C
      SUBROUTINE LINE (I,J,YAXIS,ITIME,SLOPE,CONST)
      DIMENSION YAXIS (20,30),ITIME (20,30)
       SLOPE = (YAXIS(I+1,J)-YAXIS(I,J)) /
     $          (ITIME(I+1,J)-ITIME(I,J))
       CONST = YAXIS (I,J) - (SLOPE * ITIME (I,J))
       I = I+1
      RETURN
      END
C
C**     FUNCTION OSCIL
C
C  OSCIL RETURNS ONE SOUND SAMPLE WHEN GIVEN THE AMPLITUDE,
C  THE SAMPLE INCREMENT (FREQ.) AND THE PHASE.  A SINE TABLE
C  MUST ALREADY BE AT MEMORY LOCATIONS SINE(0) THRU
C  SINE(511).  THIS IS AN INTERPOLATING OSCILLATOR.
C
      FUNCTION OSCIL (AMP,SI,PHS)
      COMMON S,SINE (511)
       IPHS1 = IFIX (PHS)
       IPHS2 = MOD (IPHS1+1,511)
       H = PHS - AINT (PHS)
       PHS = AMOD (PHS+SI,511.)
       IF (PHS) 1, 2, 2
1        PHS = PHS + 511.
2        OSCIL = AMP * (SINE (IPHS1) +
     $           (SINE(IPHS2) - SINE(IPHS1)) * H)
      RETURN
      END

 .TITLE  SOUT
 ;       SAMPLE "PACKING" AND OUTPUT ROUTINE FOR PDP-11
 ;       WRITTEN BY LARRY JOHNSON, LAST MODIFIED 18-FEB-76
         .MCALL  ..V2..,  .REGDEF,.WAIT,   .WRITE
         .REGDEF
         .GLOBL  SOUT
```

```
        .CSECT  BUFFER
ICHAN:  .BLKW   1
INDEX:  .BLKW   1
AMAX:   .BLKW   2
SAMP:   .BLKW   1000
OUTBUF: .BLKW   400

        .CSECT
SOUT:   .WAIT   ICHAN
MOV     #100000,R3      ;R3 IS A MASK
        MOV     #SAMP,  R4
        MOV     #OUTBUF,R5
LOOP:   MOV     @R4,    R0
        CLR     (R4)+
        MOV     @R4,    R1
        CLR     (R4)+
;       THE NEXT VALUE OF SAMP HAS BEEN MOVED INTO R0, R1
;       AND IT'S VALUE CLEARED FROM /BUFFER/
        MOV     R0,     R2
        BIC     R3,     R2      ;R2, R1 IS ABS(SAMP)
        CMP     R2,     AMAX
        BLO     INT
        BHI     1$
        CMP     R1,     AMAX+2
        BLOS    INT
;       ABS(SAMP) > AMAX, SO SET AMAX = ABS(SAMP)
1$:     MOV     R2,     AMAX
        MOV     R1,     AMAX+2
INT:    CLRB    R1
        BISB    R0,     R1
        SWAB    R1              ;C IS CLEARED
        BIS     R3,     R1
        ROR     R1      ;R1 HAS 15 BITS OF INT(ABS(SAMP))
        ASL     R0
        BCC     1$
        NEG     R1      ;IF SAMP < 0 THEN R1 = -R1
1$:     XOR     R3,     R0      ;MAKE EXP. 2'S COMP.
        SWAB    R0
        MOVB    R0,     R0      ;R0 IS EXPONENT
        NEG     R0
        BLE     2$
        CLR     R0      ;0 IS LARGE ENOUGH FOR -EXPONENT
2$:     ADD     #17,    R0      ;R0 IS # OF RIGHT SHIFTS
        BLE     NOSHFT
3$:     ASR     R1
        SOB     R0,     3$
NOSHFT: MOV     R1,     (R5)+   ;PUT RESULT INTO OUTBUF
        CMP     R4,     #SAMP+2000
        BLO     LOOP    ;LOOP BACK IF NOT DONE
        .WRITE  #LIST, ICHAN, #OUTBUF, #400, INDEX
        INC     INDEX
        RTS     PC

LIST:   .BLKW   5
        .END
```