Center for Computer Research in Music and Acoustics **August 1982**

Department of Music
Report No. STAN-M-11

INSTRUMENTAL TIMBRE AND RELATED ACOUSTICAL PHENOMENA IN THE PERCEPTION OF MUSIC

Final Report

by

John M. Chowning, John M. Grey, James A. Moorer, Loren Rush

Research sponsored by

National Science Foundation

CCRMA
DEPARTMENT OF MUSIC
Stanford University
Stanford, California 94305

Center for Computer Research in Music and Acoustics August 1982

Department of Music
Report No. STAN-M-11

INSTRUMENTAL TIMBRE AND RELATED ACOUSTICAL PHENOMENA IN THE PERCEPTION OF MUSIC

Final Report

by

John M. Chowning, John M. Grey, James A. Moorer, Loren Rush

Ongoing research at the Stanford Center for Computer Research in Music and Acoustics has been aimed at the investigation of the perception of musical timbre, using advanced developments in the field of digital signal processing for the preparation and presentation of auditory materials. Various interwoven phases of research have been undertaken; they are oriented towards uncovering the distinctive features and dimensions in musical timbre perception. More recently, we have been looking at the perception of timbre in temporal contexts, attempting to more directly relate laboratory findings to the normal activity of musical perception. In the same vein, we have also initiated the cross-cultural study of musical perception, examining the effects of timbre on the way in which musical sounds are combined with one another; we are looking for common underlying perceptual principles that can explain widely varying musical traditions that exist in different cultures. Finally, we are also beginning to relate timbre with our acoustical phenomena in music percpetion, such as dynamic loudness, apparent duration and apparent onset time of time-varying, naturalistic tones.

This research was supported by the National Science Foundation under Contract NSF BNS 77-22305-A1,2. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of Stanford University, any agency of the U. S. Government, or of sponsoring foundations.

1

INTRODUCTION

The primary perceptual attributes of sound most studied in the field of psychoacoustics have bee pitch, loudness and localization. Although there has been much speculation about the perception c *timbre*, there has been far less substantive work on this vastly complex topic of audition. Since timbı is *multidimensional*, corresponding to diverse and interactive attributes of tone, it has been particularl difficult to study. Furthermore, there exists no generally accepted definition of timbre in terms of i constituent elements. Timbre has been most often *negatively defined* as the differences serving t distinguish two signals that are equal in pitch and loudness. This does not delineate the possibl physical *bases* for the phenomenal differences, nor does it discuss the various phenomenal aspects c timbre. Throughout the history of timbre research, different *operational* definitions have bee employed, each suggesting a particular list of physical properties for generating stimuli and a particula subjective measurement of timbre. Unfortunately, little work has utilized stimuli that match sound perceived in normal, every-day settings; most signals used in psychoacoustical research have failed t achieve the dimensional *richness* found in naturalistic sounds.

In much of the classical research on timbre, the definition of timbre has been operationally ove limited to consist only of a *steady-state spectrum*; at that, this spectrum has almost always been limite to components having harmonic frequency ratios. Certainly, this provides a way of formulating manageable research topic, and yet, today we still possess far too little knowledge about the perceptio of timbres more characteristic of music; naturalistic sounds are time-varying in nature and not stead state. Furthermore, we have almost no knowledge about the perception of timbre in such naturalisti listening conditions as *musical contexts*; a majority of research has been limited to *isolated sound* outside of any temporal context. Another notable limitation of past research is that most studies hav been exclusively oriented towards musical timbres of the *non-percussive* and *periodic* variety. We hav no basis for extending research findings to include the different kinds of percussive and inharmoni sounds that occur in our own musical practice, and which in fact serve as the basic timbral foundation in certain other *musical cultures* of the world, such as Indonesian music.

Limitations in past research have been due, in part, to the possibilities of signal generation in th laboratory. Until recently, it has been impossible to generate artificial stimuli which could duplicat the complexities inherent in natural sounds. The equipment available up until the last decade was nc capable of producing signals with a high degree of control or reliability. This was a factor in forcin the more *limited definitions* of timbre employed for research purposes. However, with advances i digital signal processing techniques, many of the former limits of equipment and techniques to contrc the production of signals have been overcome. In principle, a digital computer can generate *any* signa that a loudspeaker can transduce; it also can analyze *any* signal that a microphone can transmit into it memory. Much of our work has been oriented towards perfecting the techniques for these tw complementary processes, *analysis* and *synthesis*, and applying them to more complex musical timbre both harmonic and inharmonic, and to musical contexts. Our aim in this process has been to come t model the essential acoustical elements of musical timbre, those significant aspects of a signal that ar salient in the perception of timbre and closely related musical dimensions.

Ongoing research at Stanford University's Center for Computer Research in Music and Acoustics h been to investigate the perception of timbre, making use of the most recent developments in the fiel of digital signal processing and multidimensional perceptual scaling. Various interwoven phases research have been explored.

A. Analysis by synthesis

A major line of our research program has been in the area of *analysis by synthesis* research. Brief this type of work relates to the investigation of timbre perception through the *digitization* (recordi into the computer) of natural instrumental tones, their *analysis* (using digital signal processi techniques) and subsequent *synthesis* based upon the acoustical information obtained from the analy (this may be altered for the resynthesis). Of central interest in this work is the perceptual salience any alteration of the acoustical information used for resynthesizing a natural timbre. *Discriminatic* then, is the primary perceptual measurement in most studies. The intent of this area of research twofold: 1) to uncover the *perceptually significant features* which are discriminable when eliminate and 2) to *reduce the complexity* of the acoustical representation of timbre to its most perceptually salie attributes. The following specific projects have led to an increased understanding in both of the areas.

1. *Higher-level data reduction.*

The systematic exploration of *simplifications* of the acoustic information used for synthesis has yield some of our most interesting findings in this area. Visiting scientist G. Charbonneau and Gr performed a study that led to very promising findings. The aim of the work was to further simpli the model for timbre that was based on the earlier work of Grey (1975), testing the simplifications l their discriminability with the already successfully simplified tones of an earlier study (Grey ar Moorer, 1977; Appendix IX). The degree of further new simplifications which remained virtual indiscriminable was impressive.

a) Our first successful simplification involved the reduction of the number of *independent* amplitu functions necessary for the resynthesis of a tone. An *average temporal amplitude function* w constructed on the basis of the amplitude functions of all the harmonics (scaled in time such that th all would begin and end together). This *single* time function then was substituted for each of the ma independent harmonic amplitude functions by scaling its peak amplitude to that of the particul harmonic (thus preserving overall spectral shape) and by scaling its time progression such that it beg and ended at the same times as the original function for that harmonic (thus preserving the relati temporal pattern of the onsets and offsets of the harmonics of the tone). The method was found to l successful for a surprisingly large number of timbres, and the results indicate that the most importa aspects of the temporal amplitude functions of the harmonics are their patterns of onset-offset ar their spectral envelope, and not necessarily the fine details and differences of their individual shap The cases where this substitution was the least successful involved tones which had strongly shifti formant regions, the implication being that certain spectrally changing timbres will need more than o single function to control the amplitudes of their harmonics through time.

b) Another significant simplification was made in the frequency domain. Here, it was found that for large number of tones, a *single frequency function* could also be successfully substituted for the man independent temporal frequency functions of the individual harmonics. Past research has shown th the substitution of constant frequency functions for the time-varying functions was highl discriminable, thus indicating that various factors of the temporal changes found in the acoustic analysis of the harmonics' frequency functions were indeed perceptually salient. In the present findin a single time-varying frequency function was successfully substituted for the independent function controlling the harmonics. Usually this function was that of the fundamental harmonic, scaled to th various harmonic frequencies by multiplication.

c) A third area of simplification, given the success of the above two simplifications in many cases, wa a simplified representation of the *timing pattern of onsets and offsets* for the harmonics of a ton Smooth mathematical functions were substituted for the actual pattern of entries and exits of th harmonics. Success was found with best-fitting third-order polynomial representations for the mor highly variant analyzed timing relationships. The differences that were produced were generally on th order of 5 milliseconds; hence the indiscriminability of the operation points out a temporal integratio in hearing that would be expected on the basis of other psychoacoustic research (Dillon, 1979). Th perceptual cues important in the temporal structure of an instrumental timbre in the attack and deca then, may often be represented in a more simplified manner than actually occurs acoustically. This ha been recently supported by A. Benade (1981).

d) A final area of simplification in this work involved simplification of the *shape of the spectra distribution*. Here, we attempted to substitute formant-type representations for the actual distribution o spectral energy. It may be expected that this type of representation would remain indiscriminable fror the original spectral shape because of masking and integration of energy within critical bandwidths i the ear. We found, however, that small changes in the actual levels of the harmonics were often easil detected (similar to the findings of Dillon, 1979).

Based upon the results of this work, a further study of the relationships among these simplification for the 16 different timbres used in earlier work (Grey & Moorer, 1977) was done. The analysis an interpretation of the data was recently published (Charbonneau, 1981). This work, as extended bot into inharmonic tones and musical contexts, is discussed below in the next two areas of research.

2. *The analysis and perception of inharmonic timbres.*

Several inharmonic timbres, both from Western music and also from the Indonesian musical traditior have been recorded into computer memory. We have digitized many of the percussive sounds used i orchestral music of the Western tradition, including a wide variety of bell sounds. Additionally, w have digitized all of the sounds from two different orchestras of metallophones that represent tw important sets of instruments in the Indonesian musical tradition: 1) several sets of Balinese *gende wayang* and 2) one complete Javanese *gamelan*.

Indonesian music is largely based upon timbres that produce *inharmonic* spectra, and we consider i important musical material for perceptual research because it contains many of the most extre differences from Western music in terms of timbral structure. These differences are important, becau they have helped to distinguish learned and innate patterns of pitch perception (see Divenyi, 19 1980 and Houtsma, 1980). We will discuss this more in section C, which looks into the percept relationships between timbre and intervals.

As a first step towards the analysis by synthesis process as applied to inharmonic timbres, we attemp to use the phase vocoder technique (Moorer, 1976) to analyze the acoustical properties of the timb from Balinese gender wayang, an instrument typical of Indonesian music. After much testing, concluded that the phase vocoder presented problems with inharmonic spectra that were not eas solved. The phase vocoder analyzes a signal with a set of evenly-spaced bandpass filters. This is id for harmonic tones, where the energy of the spectrum is evenly-spaced in frequency. However, wh there is inharmonic energy, quite often it may be simultaneously analyzed by two adjacent filters. T recombination of outputs of two adjacent channels was found to be a non-trivial problem, due to n linearities introduced into our system of analysis. Hence, we have temporarily abandoned the ph vocoder in favor of developing alternative methods more appropriate for the analysis of inharmo energy.

Working with J. O. Smith, a graduate student specializing in digital signal processing, we have be developing a method of analysis based upon the output of individual tracking filters with compl output which can be transformed into amplitude and phase (or frequency) information. The filter c be set to any center frequency through time, and thus can track changing frequencies, always keepi the energy centered in its bandwidth, thereby overcoming the problem of energy being analyzed at t edge of a channel in the inharmonic case. The bandwidth of the filter can also be controlled, so tha may be set to a value most appropriate for the particular spectral context; if there are seve components that are close together, narrow bands can be used to analyze them. We expect that t technique will yield far more manageable acoustical data in the case of inharmonic tones than either our previous methods, which have been mainly appropriate to the harmonic case.

3. *The analysis and perception of timbres in musical contexts.*

We have taken first steps towards direct research into the effects of musical *context* on timb perception. The study referred to above (Grey, 1978) shows that musical context has a strong effect the perception of timbre, and this effect is not easily predictable from studies of isolated tones. Strawn, a Ph.D. candidate at our center, is working on the analysis of musical contexts for his docto dissertation.

For Strawn's thesis, a number of simple two and three note groups of notes have been record examples of the most primitive musical contexts. Several different instruments have been digitiz including one brass, string and woodwind instrument. Different musical intervals were recorded, as w as different ways of playing the simple short musical phrases. We are looking at how one note chang into the next, given the conditions of instrument, interval and articulation. Aspects of timbre in conte will be the focus of this work. Mr. Strawn has succeeded in developing an algorithm for automatica modelling this data with line-segment approximations (Strawn, 1980).

B. Multidimensional scaling

A second research area is the analysis of perceptual similarity data using *multidimensional scalin* techniques. Spatial representations of the perceptual relationships between stimuli are constructed : that distances between the stimulus points in the space correspond to their psychological "distances" : measured by the similarity judgment. The similarity judgment is seen to be more neutral and gener in nature than ratings of the stimuli upon more stimulus-specific verbal scales; in that the listeners a: not instructed to attend to specific attributes of tone, there is less inherent experimenter bi: introduced. The spatial representation, or solution, of the similarity data is then interpreted in terms (the physical properties of the stimuli underlying their relational arrangements.

We have explored the perception of *temporal patterns* for timbres as related to the similarity structu: for isolated tones. The formation of *melodies* based not upon pitch patterns, but rather upon patter of timbre, was first proposed by the composer, Schoenberg. Recently, R. Erickson (1973) has reviv interest in this topic with an experiment in *timbre pattern perception*, Grey collaborating in one pha of the experiment. We have looked at this topic more systematically in a pilot study, by producir simple temporal patterns of pitches and timbres, alternating at different rates. For example, thr pitches continually alternating (A, B, C, A, B, C, ...) with four timbres continually alternating (trump horn, bright cello, muted cello, trumpet, ...). Whether the combined melody of pitches played by timbr is heard to perceptually segment into "threes", based on the pitch sequence, or "fours", based on tl timbre sequence, depends upon the relationships of the timbres. In fact, we found that the similari structure for the tones as uncovered through *multidimensional scaling* was useful in predicting tl strength segmentation based upon the timbral pattern perception versus the pitch pattern (Grey, 1977)

We have found in this preliminary research that a temporal context increases the perceptu importance of relationships along the *spectral axis* of the scaling solution, and decreases the strength relationships along the family-related, temporal axes. For instance, in the case above, the timbre patte was dominant, and the segmentation of the sequence was not broken into "fours" but into "twos": tl trumpet and bright cello formed one stream against the counter-stream of the horn and muted cell This was counter-intuitive in that it opposed instrument family membership (which would redu "four" to "two" by joining trumpet-horn and cello-cello). However, it did correspond to links on tl spectral axis of the similarity scaling. Systematic research has shown that the distances of timbres c the spectral axis predicts the strength of the timbral pattern versus pitch pattern, with tones mo separated in the spectral dimension forming more strongly independent timbre streams. This findir would seem to correspond with other results in the area of stream-segregation perception (McAdar and Bregman, 1979).

C. Harmonicity and the perception of intervals

Harmonicity refers to the frequency ratios of the energy within the spectral distribution of a timbre. the spectrum is *harmonic*, then it is composed of partials having integer frequency ratios with tl frequency of the lowest energy (the fundamental). If it is *inharmonic*, then it has partials who frequencies are not in simple integer ratios to the lowest frequency of energy. The former type timbre indeed serves as the foundation for Western music: vibrating strings, reeds, lips, vocal chor are typical means of excitation. The frequencies of vibration in the spectra of such sounds are near harmonic (there are small and perhaps insignificant departures from true harmonicity in most of the sounds). There are other musical traditions, such as Indonesian music, founded upon sounds that a *inharmonic* in nature: vibrating bars free at both ends and gongs provide the sources of excitatio Additionally, many percussive instruments and bells within Western music comprise a class inharmonic sounds.

In the research on musical timbre perception we found that most studies have been exclusive oriented towards Western musical timbres of the *non-percussive* and *periodic* variety. As a result research done at CCRMA, we have seen an increase in research on the perception of inharmon timbres. Moreover, we have established contact with research groups in The People's Republic China who are currently working in this area (Ma, 1980 and Xiang-peng, 1978, 1980). Clearly, tl feeling is international in scope that inharmonic spectra play an important part role in the perceptic of timbre and have heretofore been sorely neglected.

The relationship between spectral distribution (and timbral harmonicity) and the *perception intervals* has become an increasingly important topic. Some contemporary theories of pitch perceptic consider *pitch* may be a *learned* attribute (Terhardt, 1974) and that it is possible that our music intervals are related directly to the harmonic overtone structure that we learn by association. If this the case, then the timbres that we hear from birth provide the materials for forming specific types associations. In our cultural experience, this material is *harmonic* in nature, hence, possibly, tl intervals that we use in our music relates to those intervals found in the *harmonic series*, those havir simple integer ratios, like the octave (2:1), fifth (3:2) and fourth (4:3) (see Divenyi 1979, 1980 ar Houtsma 1980).

A contrasting point of view, which is the more classic view of Helmholtz extended by recent researc (Plomp & Levelt, 1965; Kameoka & Kuriyagawa, 1969), is that the psychophysical *roughness* sinusoidal interactions predicts the composite *consonance* or *dissonance* of specific musical intervals. mathematical model for the dissonance of an interval exists, taking into account the specific frequenci and amplitudes of all the partials of the simultaneous tones. Given *harmonic* spectra, the model predic maximal consonance of those intervals used in Western music: the set of intervals with simple integ ratios (the same ratios as found in the harmonic series). This theory goes on to show that tunings intervals *close to* integer ratios, but *not exactly* integral, will still be relatively high in consonance. Th model accomodates the intervals actually used in modern, equal-temperament tuning. The model al seems to work to predict the relative consonance perceived for *inharmonic* tones (see Piszalski ar Galler, 1979a,b).

Both theoretical viewpoints suggest that the specific set of intervals found in musical scales is related the timbral material out of which such scales are formed (albeit for different reasons). These theori are *indistinguishable* for harmonic tones, hence the increasing importance of understanding wh happens with *inharmonic* spectra. We have begun to look at this issue in two ways. The first w: examines the perception of inharmonic spectra that are synthetically generated by stretching tl frequency ratios of harmonic partials, and looks at various aspects of interval perception for suc spectra. The second way examines the connection between the inharmonicity found in the natur timbres actually used in Indonesian music and the lack of simple-integer frequency ratios in the tuning of intervals; we seek, in this *cross-cultural* study of interval perception, to find underlyi perceptual principles that can explain the widely varying tuning practices that exist in differe musical cultures.

Our research in how harmonicity affects pitch perception and tuning has brought into the forefro the question of how harmonicity relates to fusion in timbre perception. A brief review of past resear in fusion, as well as current work being done in that area, will be discussed in section 3 below.

1. *Stretched partials and intervals*

The basic materials used in this research were first suggested by J. R. Pierce. The spectrum of tl inharmonic stimulus is synthetically generated by stretching the frequency ratios between the partials a harmonic tone by a similar stretch factor. Pierce (1966) was originally interested in a comparison the stretching of the *partials* with a similar stretching of the *intervals* used in musical contexts. He w in essence, interested in assessing the influence of psychoacoustic *consonance* and *dissonance* on tl sense of musical harmony. Traditional Western musical timbres have harmonic spectra, and tl intervals used in tuning are simple-integer ratios. This has been hypothesized to relate to *consonan* and *dissonance*, where consonance relationships in intervals are maximal for harmonic series at simpl integer ratios between fundamentals. An alternate hypothesis, however, is that the *periodicity* betwe fundamentals is the key factor in harmony, where simple-integer ratios are again the most period relationships. Since these two theories cannot be contrasted for normal Western practice, based harmonic spectra, Pierce was interested in examining the effects under the conditions of stretche inharmonic spectra.

Working with M. V. Mathews, various experiments of stretched partials and stretched tunings we performed (Marcus, Mathews & Pierce, 1979). Findings indicate: 1) stretching does not destroy the ke sensing ability of harmonic relationships; 2) stretching does destroy the perception of finality in harmonic cadence; 3) removal of selected dissonant overtones in a cadence using normal timbr materials has little effect on the sense of finality, hence, dissonance relationships have little effect the finality of normal musical cadences; however, 4) augmenting the dissonance relationships in cadence by using special timbres will have an effect on the sense of finality. The findings indicate th there are effects of the consonance relationships in the harmonic case that are preserved in tl inharmonic cases, yet these effects do not necessarily relate to the perception of finality in cadences.

The use of these timbral materials in the study of interval perception was taken up by Dr. E. Cohen, research affiliate at our center. In her research (Cohen, 1979a,b,c,d), a number of experiments we done concerning the perception of intervals using these inharmonic spectra. She had listeners adju the frequecies for one of two simultaneously sounding tones, both of which were composed of stretch partials. The two intervals tuned were the *octave* and the *fifth*, normally having fundamental frequen ratios of 2:1 and 3:2, respectively, in the case of harmonic spectra. When there was more than a fi percent stretch factor in the spectra, there were two distinctive styles of tuning that differe individuals adopted. The first was tuning to *match partials*, hence preserving the relationships consonance found in the harmonic cases for perfect octaves and fifths. In the inharmonic, stretch case, then, the octave and fifth were *stretched* by the same ratios as the partials. The alternate strate found in other individuals was tuning the *fundamental components* only, to a 2:1 frequency rati regardless of stretching and dissonance relationships. These individual differences turned up again the quite different research discussed immediately below, in part (b), of cross-cultural, naturalis perception.

2. *Cross-cultural research on inharmonicity and interval perception*

The timbres that predominate in most of Indonesian music are those having *inharmonic* spectra. It an important musical culture for perceptual research since it embodies many of the most extre differences from Western music in several important musical dimensions. Among the most interesti of these differences are: 1) timbres with *inharmonic spectra*, as opposed to the harmonic spectra that the basis of Western music; 2) tuning systems that do *not* ordinarily contain *simple-integer ratios intervals*, as opposed to the many nearly simple-integer ratios found in the intervals of Western mu (2:1, 3:2, etc.); and 3) the *lack of standardization of intervals* within scales between different sets instruments, as opposed to Western music, in which almost any instrument can play with any other d to standardization of tuning. We feel that there is a connection between *timbre* and *tuning*, specifica between the inharmonicity of the timbre and the non-integer, non-standard nature of the Indonesi tuning system for intervals.

The approach here is slightly different from the above approach to the topic, because we are looki at how an *already developed* musical system, acoustically based upon naturally inharmonic sou structures, constructs intervals in its tuning. In this way, we are examining the *effect* of actu inharmonic spectra on tuning in a fully developed, naturalistic setting. We are here specifica interested in the inharmonic timbres from Balinese gender wayang, a form of metallophone. We ha recorded and analyzed a number of such instruments that have been tuned by various Balinese tune Here, we are attempting to uncover the significant effects of the inharmonic spectral structure on t choices of intervals in their tuning.

Perceptual research carried out in a *cross-cultural* context is attempting to uncover the relationsh between timbre and tuning practices. We are interested here not only in the perceptual underpinnir of widely diverse tuning practices, Western versus Indonesian, but also in the possible differences perceptual styles for the two groups of people, based upon extremely different musical experien (Divenyi 1979, 1980 and Houtsma 1980).

For our initial attack on the question, we have started research on the perception of the *octave*. Western musical practices, the perfect octave is considered to be a 2:1 ratio of fundamental frequenci This can also be expressed as 1200 cents (100 *cents* for each equal-tempered minor second, the distar between two immediately adjacent notes on the piano). In the Balinese instruments we measured, found that the octaves varied between 1130-1300 cents, and were not normally at 1200 cents. T tuning system in Bali actually insures that octaves on many instruments will either be too large or t small.

a) *The ideal perceptual octave*. Several recent psychoacoustic studies have concluded that the b *perceptual octave* for both simple and complex tones is generally slightly wider than a 2:1 ratio frequencies (Sundberg and Lindqvist, 1973; Ward and Martin, 1961; Hood, 1974; Risset, 1978). T has also been put forth as a correlate and possible explanation of the stretched octaves found both piano tuning and in Indonesian tuning (Dowling, 1977). In order to get at this more directly, we eel advisable to determine the frequency ratio for inharmonic tones that corresponded to the ide perceptual octave with experienced Western musicians.

We noted, first, that all of the perceptual research on interval perception that came up with the findin that the perceptual octave was wider than 2:1 was done for *sequential* intervals. In music, th significance of the octave as a *harmonic* interval entails the *simultaneous* sounding of the two notes. N data existed for the best tuned perceptual octaves in the simultaneous condition, so we first set about t find whether the sequential condition had any bearing on the simultaneous case.

We used four spectral conditions in a tuning experiment. Listerers used a method of adjustment t tune up the best perceptual octave. Many of the professional musicians used were experienced tuner One Balinese musician visiting America at the time was used. The four different spectra used were: pure tones (fundamental alone), 2) two-component tones (fundamental plus second harmonic), Balinese tones (fundamental plus inharmonic partials), and 4) modified Balinese tones (fundamenta inharmonic partials and an added second harmonic). The frequencies of Balinese inharmonic ton has been analyzed, and roughly correspond to the spectral ratios of an idealized vibrating bar free both ends: 1.0, 2.7, 5.4, 8.9, and so forth. It is most important to note that normally, there is an *absen of energy* near the frequency of a *second harmonic*. Most theoretical viewpoints on interval perceptio would predict that energy near the ratio of 2.0 (the second harmonic) is used in tuning the octave (th fundamental of the higher tone is in direct proximity with the second harmonic of the lower tone Hence, we created conditions (2) and (4) to test this.

Findings, both for Western and Balinese listeners, was that the best perceptual octave indee corresponded to a frequency ratio of exactly 2:1 in all four cases. This implies serious limitations i applying the findings of a stretched ideal octave, discovered from sequential tuning, to th simultaneous case most common in music. It also implies that if Balinese tuners wanted to tune i perfect octaves of 1200 cents, they would be able to do so. Hence we became interested in what aspe of the inharmonic sounds was important in establishing the octaves they actually use. This was done i study (b) below. (In research for her doctoral dissertation, E. Cohen found both stretched an compressed octaves in tuning simultaneous intervals with harmonic and inharmonic overtones, whic contrasts with our findings. See (Cohen, 1980a,b).)

b) *Octaves in Balinese music.* The next major step in our research project investigated the relationshi between inharmonic timbres and the actual tunings of octaves in Balinese music. There were tw possible hypotheses concerning the tuning practices in Bali. First, that the spectral components of th particular tones that were tuned *determined* the octave size for those tones. This would be based upo a dissonance interaction. The frequency ratios of 2.7 and 5.4 were considered possible determinants a dissonance relationship that was close to 2:1 - in the actual sounds measured, the variance around 2 and 5.4 might have explained the variance found in the tuning of octaves. Contrasting with th hypothesis, our second hypothesis was that the *absence of energy at the second harmonic* eliminated a signicant dissonance relationships with the fundamental of the upper tone, so that octave tunin became a process without the limiting interaction of dissonance. A third hypothesis, based upon th *learning* of spectral structures for particular sounds, leading to intervals in musical practice, mentioned above, was tested. However, we found no correspondence between the average ratios use for intervals and those between the partials of a normal tone. The perceptual learning theory was n pursued. Yet we are interested in more general aspects of perceptual learning, as in the contrast c Western and Balinese listeners, each trained in extremely different musical cultures.

This research involved the retuning of existing Balinese intervals to correspond more to perfect 2: octaves in the simultaneous condition. Actual musical examples were generated on the computer, usin the real tones from the gender wayang instrument. Digital signal processing was used to retune th *fundamental frequencies* of the sounds. Also, processing was employed on the *spectral structure* of th sounds, retuning the upper patials and even adding in energy near the theoretical second harmoni Four experiments were performed, where the stimuli were: 1) Original sounds *versus* fundamentals re tuned to 2:1 octaves where the upper partials of the re-tuned fundamental were unchanged; 2) Origin: sounds *versus* fundamentals re-tuned to 2:1 octaves with the upper partials re-tuned at the same rati of shift with the shifted fundamental; 3) Original sounds with perfect second harmonics added *versu* fundamentals re-tuned to 2:1 octaves, upper patials unchanged and energy added in at a perfect secon harmonic frequency, 2:1 from the fundamental; and 4) Original sounds with stretched secon harmonics added *versus* fundamentals re-tuned to 2:1 octaves, upper patials unchanged and energ added in at a stretched second harmonic frequency, corresponding to the ratio found for the octave i the original tuning.

Tuning preferences among the paired alternatives for the four conditions were determined for Balines musicians and instrument tuners. It was found that there were two individual *styles* of preference, th: came out in conditions (3) and (4) above. With conditions (1) and (2), listeners preferred either th original tunings or the perfect octaves, consistently, showing that there were no strong determinants c octave size in a musical context with the original timbres of Balinese instruments. In conditions (3) an (4), certain listeners maintained a preference for one particular size octave, while other individua adjusted the preferred octave size to maximize the *consonance* relationship between the added parti: (near the second harmonic) and the fundamental of the upper tone, hence a perfect octave wa preferred in the case of the 2:1 harmonic ratio and a stretched (original) octave was preferred for stretched ratio of the partial. These two styles of listening correspond to the individual difference found in Cohen's research in section (1) above.

We conclude from the findings that the inharmonic spectral structure of Indonesian sounds allows fc non-standard tuning practices and non-integer frequency ratios because of the *lack of energy near th second harmonic*. This is in correspondence with a *consonance* and *dissonance* theory of interv: perception: there is no dissonance interaction with the fundamental of the upper tone in octave tunin; so the process is not strictly limited to 2:1 frequency ratios to maximize consonance, as it is in Wester music based upon harmonic series. The importance of consonance relationships between the uppe partials of the original sounds, for instance the approximate 5.4 ratio of the lower tone to the 2.7 of th upper, was suggested not to play a major part in tuning preferences by the results of condition (2 Further experimentation may reveal this to be a possible factor, but for these sounds we found littl evidence for a *deterministic* effect of spectral structure on octave tuning. Rather, Balinese musi appears to be free to tune intervals as an *active parameter* of musical aesthetics, in contrast to Wester music, precisely because there seems to be a *lack of timbral determinism* based on consonance an dissonance.

3. *Spectral fusion*

Recently, Mike McNabb, a composer and researcher at our center, discovered a perceptu phenomenon while experimenting with a vocal synthesis technique. He obtained the Fouri Transform of a soprano tone that was recorded and digitized at the center. He then synthesized tone, using additive synthesis, such that the spectral balance was the same as indicated by the Fouri Transform. At first, the frequency for each harmonic was kept constant; however, the tone did n sound vocal at all. In fact, it didn't even sound natural. When some vibrato was added, such that harmonics were affected synchronously, the percept was strikingly realistic.

John Chowning explored this phenomenon even further. He synthesized a tone such that ea harmonic (a sine tone) began one after another, but remaining sustained. Again, the spectral balan of the partials corresponded to the levels obtained from a Fourier Transform of a recorded soprar tone. With all harmonics playing, it was very easy to hear each harmonic separately, as if there we many sources, or voices (each source being a sine tone). But as soon as a common vibrato was add to all the harmonics, the sound fused into a percept of a single source — that of a sung soprano tone.

These examples show that temporal aspects of a tone are important features of its timbre, even durir its so-called steady-state portion. In other words, spectral balance alone cannot determine timbı because a constant spectrum may not fuse, and one can't really have timbre without fusion. Tl examples also raise this question: What are the characteristics of a sound that cause it to fuse into percept of a single source with a particular timbre?

Elizabeth Cohen has investigated the role of harmonicity (or the lack thereof) in the fusion of compl tones. She has found that fusion depends on temporal envelope, degree of inharmonicity, and spectr content (Cohen 1979a,b,d and 1980a). Stephen McAdams, a graduate student at our center, is al investigating fusion and source identification for his doctoral research. Results of his prelimina work appeared in (McAdams and Bregman, 1979). The work of Cohen and McAdams will be valuable aid in determining the parameters for timbre.

D. Timbre and perceived onset

Finally, in this section we would like to cover some of the interests we have in looking at tl relationships between timbre and other aspects of tone, such as perceived onset time, duration ar loudness. We would expect rather direct and strong relationships between timbre and these other ton attributes because they are effects of similar acoustical dimensions. *Spectral shape* is a determinant timbre as well as of loudness. Similarly, the *temporal envelope* of a signal is a determinant of ons duration and timbre. Combining the two, a spectral shape that changes with time is a comple description of timbre, and also provides the material for making loudness, onset and duratic judgments on *naturalistic* signals, such as actual musical timbres.

In past research, we have found, for example, that a model of *loudness* perception for steady-sta spectral distributions (Zwicker & Scharf, 1965) was useful in modeling the *timbral* dimension relating perceived spectral brightness (Grey & Gordon, 1977). The recent derivation of a model for loudne perception for *time-varying* tones (Zwicker, 1977) presents encouraging possibilities for extending tl model to include various other aspects of the perception of time-varying tones. Already, this model h been useful in analyzing various aspects of the perception of tim

1.*Perceived onset time*

We have long been interested in formal models for the *temporal* properties of tone. One possible ι related to the above, concerns modeling the relationships actually found along temporally-rela dimensions of timbre perception uncoverd in our multidimensional scaling studies. Various *subject* correlates to these temporal dimensions have been noted, one being the "hardness" or "explosiveness" the attack (found with D. L. Wessel, 1977). In the hopes that this attribute may have something to with *perceived onset time* for tones, Gordon and Grey have run an experiment to equalize the on times for the timbres used in the original studies.

The procedure (like that of J. Vos and R. Rasch, 1980) involved the setting of two tones (e.g. A and to be locked into rhythmic phase such that their alternating series (i.e. A B A B A B ...) made perceptually isochronous rhythm and the perceived onset of the B tones perfectly bisected the durati between the A tones. By adjustment, the listener set the temporal delay between A and B, where ı physical delay between all A's and all B's was equal. In looking at the relative delays for the differ stimulus tones, all of which were taken from the set used in the multidimensional scaling studies, independent measure of onset time has been achieved.

Regardless of the success of the experiment in terms of modeling our temporal axes from the scali research, we still have data for the onset relationships among a set of timbres. We hope to be able model these onset relationships taking advantage of the temporal features of the new model : loudness perception mentioned above (Gordon, 1981).

PUBLICATIONS

Charbonneau, G. "Timbre and the perceptual effects of three types of data reduction." Compute Music Journal, vol. 5, no. 2, Summer 1981, forthcoming.

Cohen, E. "The influence of nonharmonic partials on tone perception." Journ. Acous. Soc. Amer. 6 Suppl. 1, Spring, 1979.

Cohen, E. "Fusion and consonance relations for tones with inharmonic partials." Journ. Acous. So Amer. 65, Suppl. 1, Spring, 1979.

Cohen, E. "Stretched tones with only octave partials." Journ. Acous. Soc. Amer. 65, Suppl. 1, Sprin 1979.

Cohen, E. "The effect of envelope on the fusion for tones with nonharmonic partials." Jour Acous. Soc. Amer. 65, Suppl. 1, Fall, 1979.

Cohen, E. "The influence of nonharmonic partials on tone perception." Doctoral Dissertatio Stanford University, 1980a.

Cohen, E. "Pitch processing of nonharmonic tones: A search for an auditory mechanism tha recognizes spectral patterns." Journ. Acous. Soc. Amer. 68, Suppl. 1, S110, 1980b.

Gordon, John W. "Perceptual attack time of orchestral instrument tones." to be published, 1981.

Grey, J. M. and Gordon, J. W. "Perceptual effects of spectral modifications on musical timbres Journ. Acoust. Soc. Amer. 63(5), May 1978, 1493-1500.

Grey, J. M. "Timbre discrimination in musical patterns." Journ. Acoust. Soc. Amer. 64(2), Augu 1978, 467-472.

Grey, J. M. "Multidimensional perceptual scaling of musical timbres." Journal of the Acoustic Society of America, May, 1977.

Grey, J. M. and Moorer, J. A. "Perceptual evaluation of synthesized musical instrument tones Journal of the Acoustical Society of America, August, 1977.

Grey, J. M. "Experiments in the perception of instrumental timbre." invited by the Bulletin of th Council for Research in Music Education (to be published).

Gordon, J. W. and Grey, J. M. "Perceptual of spectral modifications on orchestral instrumen tones." Comp. Mus. Journ. 2(1), July 1978, 24-31.

Moorer, J. A. "The synthesis of complex audio spectra by means of discrete summation formulae Journal of the Audio Engineering Society, Volume 24, #9, November 1976, pp717-727.

Moorer, J. A. "On the coding of high-quality digitized sound". Presented at the 1979 Europea Conference of the Audio Engineering Society, Brussels, Belgium, February 1979, Accepted fo publication in the Audio Engineering Society

Moorer, J. A. "The use of linear prediction of speech in computer music applications". Journal o the Audio Engineering Society, Volume 27, #3, March, 1979, pp134-140.

Moorer, J. A. and Grey, J. M. "Lexicon of analyzed tones, Part 1: a violin tone." ed. J. Snell, Com Mus. Journ. 1(2), April 1977, 39-45.

Moorer, J. A. and Grey, J. M. "Lexicon of analyzed tones, Part 2: clarinet and oboe tones." ed. Strawn, Comp. Mus. Journ. 1(3), June, 1977, 12-29.

Moorer, J. A. and Grey, J. M. "Lexicon of analyzed tones, Part 3: the trumpet." ed. J. Strawn, Com Mus. Journ. 2(2), September 1978, 23-31.

Strawn, John. "Approximation and syntactic analysis of amplitude and frequency functions fo digital sound synthesis." Computer Music Journal, vol. 4, no. 3, Fall 1980, 3-24.

Wessel, D. L. and Grey, J. M. "Conceptual structures for the representation of musical materials I.R.C.A.M. Rpt. No. 14, Paris, 1978.

BIBLIOGRAPHY

Ainsworth, W. A. Duration as a Cue in the Recognition of Synthetic Vowels. Journ. Acoust. Sc Amer. 51, 648-651 (1972).

American Standard Acoustical Terminology. S1.1-1960. American Standards Association, Inc., Ne York (1960).

Arabie, P., and Shepard, R. N. Representations of similarities as additive combinations of discre overlapping properties. presented at Math. Psych. Meeting in Montreal (1973).

Atal, B. S., and Hanauer, S. L. Speech Analysis and Synthesis by Linear Prediction of the Spee Wave. Journ. Acoust. Soc. Amer. 50, 637-655 (1971).

Backhaus, H. Ueber die Bedeutung der Ausgleichsvorgange in der Austrik. Zeit. Tech. Physik 1 31-46 (1932).

Baker, R. F., and Young, F. M. A note on an empirical evaluation of the ISIS procedu Psychometrika, 40, 413-414 (1975).

Bartholomew, W. T. Acoustics of Music. Prentice-Hall, Inc., New Jersey (1945).

Beauchamp, J. W. A Computer System for Time-Variant Harmonic Analysis and Synthesis Musical Tones. in Music by Computers, H. von Foerster & J. W. Beauchamp, ed. Wiley, New Yo (1969).

Beauchamp, J. W. Analysis and Synthesis of Cornet Tones using Nonlinear Interharmon Relationships. J. Audio Eng. Soc. 23, 778-795 (1975).

Benade, A. Spectral Similarities of Tones from 'Especially Useful' Musical Instruments. J. Acou Soc. Amer. 69(S1)S37(1981)

Berger, K. W. Some Factors in the Recognition of Timbre. Journ. Acoust. Soc. Amer. 36, 1888-18 (1964).

v. Bismarck, G. Timbre of Steady Sounds: A Factorial Investigation of its Verbal Attribut Acoustica 30, 146-159 (1974).

v. Bismarck, G. Sharpness as an Attribute of the Timbre of Steady Sounds.. Acoustica 30, 159-1 (1974).

Boring, E. G. Sensation and Perception in the History of Experimental Psychology. Appleto Century Co., Inc., New York (1942).

Burns, E. M. and Ward, W. D. Categorical Perception of Musical Intervals. abstract, Journ. Acou: Soc. Amer. 55 (1974).

Carroll, J. D., and Chang, J. J. Analysis of Individual Differences in Multidimensional Scaling Vi an N-Way Generalization of "Eckart-Young" Decomposition. Psychometrika 35, 283-319 (1970).

Carroll, J. D. and Wish, M. Models and Methods for Three-Way Multidimensional Scaling. i Contemporary Developments in Mathematical Psychology, R. C. Atkinson, D. H. Krantz, R. D. Lu and P. Suppes, eds. W. H. Freeman, San Francisco (1973).

Charbonneau, G. Timbre and the Perceptual Effects of Three Types of Data Reduction. Comput Music Journal, Vol. 5, No. 2, Summer 1981, forthcoming.

Chowning, J. M. The Simulation of Moving Sound Sources. J. Audio Eng. Soc. 2, (1971).

Chowning, J. M. Sabelithe, a computer-generated quadraphonic tape (1971).

Chowning, J. M. Turenas, a computer-generated quadraphonic tape (1972).

Chowning, J. M. The Stanford Computer Music Project, Numus-West, 1, (1972).

Chowning, J. M. The Synthesis of Complex Audio Spectra by Means of Frequency Modulation. Audio Eng. Soc. 21, 526-534 - also enclosed as *Appendix VIII* (1973).

Chowning, J. M. Synthesis of the Singing Voice by Means of Frequency Modulation. Swedis Journal of Musicology (1980).

Chowning, J. M. Phone, a computer-generated qudraphonic tape (1981).

Chowning, J. M., Grey, J. M., Rush, L., and Moorer, J. A. Computer Simulation of Mus Instrument Tones in Reverberant Environments. Stanford Univ. Dept. of Music Tech. Rep. STAI M-1, (1974).

Clark, M., Robertson, P., and Luce, D. A. A Preliminary Experiment on the Perceptual Basis f Musical Instrument Families. J. Audio Eng. Soc. 12, 199-203 (1964).

Clark, M., and Milner, P. Dependencies of Timbre on the Tonal Loudness Produced by Music Instruments. J. Audio Eng. Soc. 12, 28-31 (1964).

Cohen, E. "The influence of nonharmonic partials on tone perception." Journ. Acous. Soc. Amer. 6 Suppl. 1, Spring, 1979.

Cohen, E. "Fusion and consonance relations for tones with inharmonic partials." Journ. Acous. Sc Amer. 65, Suppl. 1, Spring, 1979.

Cohen, E. "Stretched tones with only octave partials." Journ. Acous. Soc. Amer. 65, Suppl. 1, Sprin 1979.

Cohen, E. "The effect of envelope on the fusion for tones with nonharmonic partials." Jou Acous. Soc. Amer. 65, Suppl. 1, Fall, 1979.

Cohen, E. "The influence of nonharmonic partials on tone perception." Doctoral Dissertati Stanford University, 1980a.

Cohen, E. "Pitch processing of nonharmonic tones: A search for an auditory mechanism t recognizes spectral patterns." Journ. Acous. Soc. Amer. 68, Suppl. 1, S110, 1980b.

Coker, C. H. Speech Synthesis with a Parametric Articulatory Model. Speech Symposium, Ky Paper A-4 (1968).

Cooper, F. S., Liberman, A. M., and Borst, J. M. The Interconversion of Audible and Visi Patterns as a Basis for Research in the Perception of Speech. Proc. Natl. Acad. Sci., 37, 318- (1951).

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., and Gerstman, L. J. Some Experime on the Perception of Synthetic Speech Sounds. Journ. Acoust. Soc. Amer. 24, 597-606 (1952).

Cutting, J. E. and Rosner, B. S. Categories and Boundaries in Speech and Music. Perception Psychophysics 16(3), 564-570 (1974).

Dillon, H. A. The Perception of Musical Transients. Ph.D. Thesis, Dept. Elec. Eng., Univ. N.S. Aust., June (1979).

Divenyi, Pierre L. Is Pitch a Learned Attribute of Sounds? Two Points in Support of Terhar Pitch Theory. Journ. Acoust. Soc. Amer. 66, 1210-1213 (1979).

Divenyi, Pierre L. A Note About Terhardt's Pitch Learning Hypothesis: A Reply to Houts Journ. Acoust. Soc. Amer. 68, 1890-1891 (1980)

Dodge, C. Synthetic Speech Music. Disk, Composer's Recordings, Inc., New York, CRI-SD-348 (197

Dowling, W. J., personal communication, (1977).

Dudley, H. The Vocoder. Bell Labs. Record, Vol 17, 122-126 (1939).

Dudley, H., Riesz, and Watkins A Synthetic Speaker. J. of the Franklin Institute, 227, 739-764 (1939

Dudley, H. The Carrier Nature of Speech. Bell System Tech. J. 19, 495-515 (1940).

Ekman, G. Two Methods for the Analysis of Perceptual Dimensionality. Perc. Mot. Skills 20, 5 572 (1965).

Erickson, R. Loops, an Informal Timbre Experiment. unpublished paper, Music Dept., U.C.S (1973).

Fant, C. G. M. Descriptive Analysis of the Acoustic Aspects of Speech. LOGOS 5, 3-17 (1962).

Flanagan, J. L., and Golden, R. M. Phase Vocoder. Bell Systems Tech. J., Vol 45, 1493-1509 (1966).

Fletcher, H. Loudness, Pitch and the Timbre of Musical Tones and their Relation for the Intensi the Frequency and the Overtone Structure. Journ. Acoust. Soc. Amer. 6, 59-69 (1934).

Freedman, M. D. Analysis of Musical Instrument Tones. Journ. Acoust. Soc. Amer. 41, 793-8 (1967).

Freedman, M. D. A Method for Analysing Musical Tones. J. Audio Eng. Soc. 16, 419-425 (1968).

Gjaevenes, K., and Rimstad, E. R. The Influence of Rise Time on Loudness. Journ. Acoust. S Amer. 51, 1233-1239 (1972).

Gordon, John W. "Perceptual attack time of orchestral instrument tones." to be published, 1981.

Grey, J. M. Loops, a Computer Realization. Tape, Music Dept., Stanford University (1973).

Grey, J. M. An Exploration of Musical Timbre. Stanford Univ. Dept. of Music Tech. Rep. STA M-2, (1975).

Grey, J. M., and Moorer, J. A. Perceptual Evaluations of Synthesized Musical Instrument Ton accepted for publication Journ. Acoustical Society Amer., est. Aug, 1977 - also enclosed as *Appendix* (1977).

Grey, J. M. Multidimensional Perceptual Scaling of Musical Timbre. accepted for publicati Journ. Acoust. Soc. Amer., est. May, 1977 - also enclosed as *Appendix X* (1977).

Grey, J. M., and Gordon, J. W. Perceptual Effects of Spectral Modifications on Musical Timbres. preparation for publication - also enclosed as *Appendix XI* (1977).

Grey, J. M. Categorical versus Continuous Perception of Musical Timbre. in preparation f publication - also enclosed as *Appendix XII* (1977b).

Helmholtz, H. L. F. On the Sensations of Tone as a Physiological Basis for the Theory of Music. J. Ellis, trans. Dover, New York (1954).

Holmes, J. N., Mattingly, I. G., Shearme, J. N. Speech Synthesis by Rule. Language and Speech 127-143 (1964). Hood, M. Indonesian Music. In -. Kahler (ed.) Handbuch der Orientalistik, 3rd e Vol. 6. Leiden, Netherlands: E. J. Brill, 1974.

Houtsma, A. Comment on Is Pitch a Learned Attribute of Sounds? Two Points in Support Terhardt's Pitch Theory. Journ. Acoust. Soc. Amer. 68, 1889-1890 (1980).

Itakura, F., Saito, S. Analysis Synthesis Telephony based on the Maximum Likelihood Method. Pr Sixth Intern. Congr. Acoust., Paper C-5-5, C17-20 (1968).

Itakura, F. Minimum Prediction Residual Principle Applied to Speech Recognition. IEEE Tra on Acoustics, Speech, and Signal Processing, ASSP-23, 67-72 (1975).

Johnson, S. C. Hierarchic Clustering Schemes. Psychometrika 32, 241-254 (1967).

Jost, E. Akustische und Psychomettische Untersuchungen an Klavinettenklangen, Arno V Verlag, Koln (1967). reviewed in Webster, et. al. (1970).

Jusczyk, P. W., Cutting, J. W. and Rosner, B. S. Categorical Perception of Non-Speech Sounds the Two-Month-Old Infant. unpublished report, University of Pennsylvania, Philadelphia (1974).

Kameoka, A., & Kuriyagawa, M. Consonance Theory, Parts I & II. Journ. Acous. Soc. Amer. 1451-1469 (1969).

Kruskal, J. B. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmet Hypothesis. Psychometrika 29, 1-27 (1964a).

Kruskal, J. B. Nonmetric Multidimensional Scaling: A Numerical Method. Psychometrika 29, 1 129 (1964b).

Lane, H. The Motor Theory of Speech Perception: A Critical Review. Psych. Review 74, 275- (1965).

Liberman, A. M., Ingemann, F., Lisker, L., Delattre, P. C., and Cooper, F. S. Minimal Rules Synthesizing Speech. J. Acoust. Soc. Am. 31, 1490-1499 (1959).

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Stuodert-Kennedy, M. Perception of Speech Code. Psych. Review 74, 431-461 (1967).

Lichte, W. H. Attributes of Complex Tones. Journ. Exp. Psych. 28, 455-480 (1941).

Lichte, W. H. and Gray, R. F. The Influence of the Overtone Structure on the Pitch of Compl Tones. Journ. Exp. Psych. 49, 431 (1955).

Licklider, J. C. R. Basic Correlates of the Auditory Stimulus. in Handbook of Experimen Psychology, S. S. Stevens, ed. Wiley, New York (1951).

Luce, D. A. Physical Correlates of Nonpercussive Musical Instrument Tones. PhD thesis. M (1963).

Luce, D. A., and Clark, M. Duration of Attack Transients of Nonpercussive Orchest Instruments. J. Audio Eng. Soc. 13, 194-199 (1965).

Ma Chengyuan, Ancient Chinese Two-Pitched Bronze Bells. Chinese Music 3, No. 4, 81-86, 1980.

Marcus, K.T., Mathews, M.V., & Pierce, J.R. A Study of Stretched Harmonics. Journ. Acous. S Amer., Suppl. 1, 66, Fall (1979).

Markel, J. D. Digital Inverse Filtering - a New Tool for Formant Trajectory Estimation. IE Trans. Audio Electroacoust. AU-20, 129-137 (1972).

Markel, J. D. and Gray, A. H., A Linear Prediction Vocoder Simulation Based Upon t Autocorrelation Method. IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol ASSP- 124-134 (1974).

Mathews, M. V. An Acoustic Compiler for Music and Psychological Stimuli. Bell System Te Journ., XL, 677-694 (1961).

Mathews, M. V. The Technology of Computer Music. MIT Press, Mass. (1969).

Mathews, M. V., and Kohout, J. Electronic Simulation of Violin Resonances. Journ. Acoust. S Amer. 53, 1620-1626 (1973).

McAdams, Stephen, and Bregman, A. Hearing Musical Structures. Computer Music Journal 3, No. 26-44 (1979).

McCandless, S. S. An Algorithm for Automatic Formant Extraction Using Linear Predicti Spectra. IEEE Trans. Acoust., Speech, and Signal Processing, ASSP-22, 135-141 (1974).

Miller, J. R., and Carterette, E. C. Perceptual Space for Musical Structures. J. Acoust. Soc. Am. 711-720 (1975).

Moore, F. R. Electronic Music - an Introduction. Carnegie Technical, Nov. (1963).

Moore, F. R. A Theory of Dissonance. Carnegie Technical, May, (1966).

Moore, F. R. Music and Computers. Enciclopedia della Scienza e della Technica/Mondadc Yearbook, pp. 490-498 (1971).

Moore, F. R. Music - Film - Computers, Filmmaker's Newsletter, v.4, Apr. (1971).

Moore, F. R. Computer Controlled Analog Synthesizers, Bell Laboratories Computing Scier Technical Report #10, May, (1973).

Moore, F. R. Real Time Interactive Computer Music Synthesis. Stanford University Report STA M-7 (1977).

Moorer, J. A. The Heterodyne Filter as a Tool for Analysis of Transient Waveforms. Stanfo Artificial Intelligence Laboatory Memo 208, July (1973).

Moorer, J. A. The Optimum Comb Method of Pitch Period Analysis of Continuous Digitiz
Speech. IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-22, 330-338 (1974).

Moorer, J. A. On the Transcription of Musical Sound by Digital Computer, Presented at the Secor
USA-JAPAN Computer Conference, August (1975).

Moorer, J. A. On the Segmentation and Analysis of Continuous Musical Sound by Digit
Computer. Stanford Univ. Dept. of Music Tech. Rep. STAN-M-3 - part extracted as *Appendix*
(1975).

Moorer, J. A. On the Loudness of Complex, Time-Variant Tones. Stanford Univ. Dept. of Mus
Tech. Rep. STAN-M-4 (1975).

Moorer, J. A. The Synthesis of Complex Audio Spectra by Means of Discrete Summatic
Formulae. Stanford Univ. Dept. of Music Tech. Rep. STAN-M-5 (1975) and Journ. of the Audio En
Soc., 24, 717-727 (1976).

Moorer, J. A. The Use of the Phase Vocoder in Computer Music Applications. Presented at the 55
Convention of the Audio Engineering Society available as Preprint number 1146 (E1) - also present
as *Appendix V* (1976).

Moorer, J. A. Signal Processing Aspects of Computer Music - A Survey. accepted for publication
the Proceedings of the IEEE, scheduled for July, (1977).

Moorer, J. A. The Synthesis of Complex Audio Spectra by Means of Discrete Summatic
Formulae, accepted for publication in the Journ. of the Audio Eng. Soc. (1977).

Moorer, J. A., Rush, L, and Loy, D. G. All-digital Sound Recording in preparation for the Journ.
the Audio Eng. Soc. (1977). Morrill, D. Trumpet Algorithms for Computer Composition, Mus
Dept., Colgate University (1976).

Nordenstreng, K. The Perception of Complex Sounds: Semantic Differential Attributes of Spee
and Music. in Contemporary Research in Psychology of Perception, J. Jarvinen, ed. WSOY, Helsin
(1969).

Oetken, G., Parks, T. W., and Schuessler, H. W. New Results in the Design of Digital Interpolato
IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-23, (1975).

Ohm, G. S. Ueber die Definition des Tones, Nebst Daran Geknupfter Theorie der Sirene u
Ahnlicher Tonbildender Vorrichtungen. Ann. Phys. Chem. 59, 513-565 (1843).

Pierce, J. R. Attaining Consonance in Arbitrary Scales. Journ. Acoust. Soc. Amer. 40, 249 (1966).

Pisoni, D. B. and Lazarus, J. H. Categorical and Noncategorical Modes of Speech Perception alon the Voicing Continuum. Journ. Acoust. Soc. Amer. 55(2), 328-333 (1974).

Piszalski, M., and Galler, B. A. Automatic Tone Identification in Continuously Played Musi Journ. Acoust. Soc. Amer. 65(S1), s123, 1979a.

Plomp, R. The Ear as a Frequency Analyzer. Journ. Acoust. Soc. Amer., 36, 1628-1636 (1964).

Plomp, R. Timbre as a Multidimensional Attribute of Complex Tones. in Frequency Analysis an Periodicity Detection in Hearing. R. Plomp and G. F. Smoorenburg, ed. A. W. Sijthoff, Leide (1970).

Plomp, R., and Bouman, M. A. Relation Between Hearing Threshold and Duration for Tone Pulse Journ. Acoust. Soc. Amer. 31, 749 (1959).

Plomp, R., and Levelt, W. J. M. Tonal Consonance and Critical Bandwidth. Journ. Acoust. So Amer. 38, 548 (1965).

Plomp, R. and Mimpen, A. M. The Ear as a Frequency Analyzer, II. Journ. Acoust. Soc. Amer. 4 764-767 (1968).

Plomp, R. and Steeneken, H. J. M. Effects of Phase on the Timbre of Complex Tones. Journ Acoust. Soc. Amer. 46, 409-421 (1969).

Plomp, R. and Steeneken, H. J. M. Pitch vs. Timbre. Seventh Int. Congr. on Acoustics, Budape (1971).

Portnoff, M. R. Implementation of the Digital Phase Vocoder Using the Fast Fourier Transforn IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-24, 243-248 (1976).

Portnoff, M. R. Personnal Communication. (1976).

Rabiner, L. R. Speech Synthesis by Rule: an Acoustic Domain Approach. Bell System Tech. J. 4 17-37 (1968).

Rabiner, L. R. A Model for Synthesizing Speech by Rule. IEEE Trans. Audio and Electroacous AU-17, 7-13 (1969).

Rabiner, L. R. Techniques for Designing Finite-Duration Impulse-Response Digital Filters. IEE Trans. on Communication Technology, COM-19, 188-195 (1971).

Reiser, J. F. SAIL. Stanford Artificial Intelligence Laboratory Memo AIM-289 (1976).

Risset, J. C. Computer Study of Trumpet Tones. Bell Telephone Labs, Murray Hill, New Jers
(1966).

Risset, J. C. An Introductory Catalog of Computer-Synthesized Sounds. Bell Telephone Laboratorie

Risset, J. C. Musical Acoustics. In Carterette and Friedman, eds., Handbook of Perception, Vol. 4, p
521-564, 1978.

Ritsma, R. J. Frequencies Dominant in the Perception of the Pitch of Complex Sounds. Jour
Acoust. Soc. Amer. 42, 191-198 (1967).

Roederer, J. G. Introduction to the Physics and Psychophysics of Music. Springer-Verlag, Ne
York (1973).

Rush, L., Moorer, J. A., and Loy, G. All-Digital Sound Recording and Processing. presented at th
55th Convention of the Audio Eng. Soc., New York (1976).

Rush, L., and Moorer, J. A. Editing, Mixing and Processing Digitized Audio Waveforms. i
preparation for the Journ. of the Audio Eng. Soc. (1977).

Rush, L., and Grey J. M. Relating Digital Techniques: Analysis, Synthesis, and Processing
Recorded Sound. in preparation (1977).

Rush, L. A Little Traveling Music. for amplified piano with computer-generated quadraphonic tap
General Music Pub. Co., in press; recorded for Serenus by Dwight Peltzer, in press (1977).

Saldanha, E. L., and Corso, J. F. Timbre Cues for the Recognition of Musical Instruments. Jour
Acoust. Soc. Amer. 36, 2021-2026 (1964).

Schafer, R. W. Echo Removal by Discrete Generalized Linear Filtering.Ph.D. dissertation, MI
(1969).

Schafer, R. W., and Rabiner, L. R. A Digital Signal Processing Approach to Interpolatio
Proceedings of the IEEE, Vol 61, 692-702 (1973).

Schoenberg, A. Theory of Harmony. Philsophical Library (1948).

Shepard, R. N. The Analysis of Proximities: Multidimensional Scaling with an Unknown Distan
Function. I. Psychometrika 27, 125-140 (1962a).

Shepard, R. N. The Analysis of Proximities: Multidimensional Scaling with an Unknown Distan
Function. II. Psychometrika 27, 125-140 (1962b).

Shepard, R. N. Psychological Representation of Speech Sounds. in Human Communication, E.
Davis and P. B. Denes, eds. McGraw-Hill, New York (1972).

Shepard, R. N. Representations of structure in similarity data: problems and prospec Psychometrika 39, 373-421 (1974).

Slawson, A. W. Vowel Quality and Musical Timbre as Functions of Spectrum Envelope ar Fundamental Frequency. Journ. Acoust. Soc. Amer. 43, 87-101 (1968).

Solomon, L. N. Semantic Approach to the Perception of Complex Sounds. Journ. Acoust. Soc. Am 30, 421-425 (1958).

Stevens, S. S., and Davis, H. Hearing - Its Psychology and Physiology. Wiley, New York (1938).

Strawn, John. "Approximation and syntactic analysis of amplitude and frequency functions f digital sound synthesis." Computer Music Journal, vol. 4, no. 3, Fall 1980, 3-24.

Strong, W., and Clark, M. Synthesis of Wind-Instrument Tones. Journ. Acoust. Soc. Amer. 41, 39- (1967a).

Strong, W., and Clark, M. Perturbations of Synthetic Orchestral Wind-Instrument Tones. Jour Acoust. Soc. Amer. 41, 277-285 (1967b).

Studdert-Kennedy, M., Liberman, A. M., Harris, K. S. and Cooper, F. S. Motor Theory of Spee Perception: A Reply to Lane's Critical Review. Psych. Review 77, 234-249 (1970).

Sundberg, J. E. F., and Lindquist, J. Musical Octaves and Pitch. Journ. Acoust. Soc. Amer. 54, 192 1929, (1973).

Taylor, C. A. The Physics of Musical Sounds. American Elsevier Pub. Co., Inc., New York (1965).

Terhardt, E. Pitch, Consonance, and Harmony. Journ. Acous. Soc. Amer. 55, 1061-1069 (1974).

Vos, Joos, and Rasch, Rudolf. The Perceptual Onset of Musical Tones. Presented at the Thi Workshop on Physical and Neuropsychological Foundations of Music, Osslach/Austria, Aug. 8-1 1980.

Webster, J. C., Carpenter, A., and Woodhead, M. M. Identifying Meaningless Tonal Complex Journ. Acoust. Soc. Amer. 44, 606-609 (1968a).

Webster, J. C., Carpenter, A., and Woodhead, M. M. Identifying Meaningless Tonal Complexes I Journ. Aud. Res. 8, 251-260 (1968b).

Webster, J. C., Carpenter, A., and Woodhead, M. M. Perceptual Constancy in Complex Sour Identification. Brit. J. Psych. 61, 481-489 (1970).

Wedin, L., and Goude, G. Dimension Analysis of the Perception of Instrumental Timbre. Scan Journ. Psych. 13, 228-240 (1972).

Weinstein, C. J., McCandless, S. S., Mondshein, L. F., and Zue, V. W. A System for Acousti Phonetic Analysis of Continuous Speech. IEEE Trans. on Acoustics, Speech, and Signal Processin ASSP-23, 54-67 (1975).

Wessel, D. L. report to Psychometric Society Meeting, San Diego (1973).

Wessel, D. L. report to C.M.E., University of Calif., San Diego (1974).

Winckel, F. Music, Sound and Sensation: A Modern Exposition. Dover, New York (1967).

Xiang-pen, Huang. The Study on the Development on the Chinese Scale Sysem with Know Acoustical Materials of the Neolithic and Bronze Age Cultures. Yunyue Zuncong Vol 1, 184 (197 and Vol. 3, 126 (1980).

Young, F. M., and Cliff, N. Interactive scaling with individual subjects. Psychometrika, 37, 385-4 (1972).

Young, R. W. Musical Acoustics. in McGraw-Hill Encyclopedia of Science and Technology. McGra Hill, New York (1960)..

Zwicker, E., Flottorp, G., and Stevens, S. S. Critical Bandwidth in Loudness Summation. Jour Acoust. Soc. Amer. 29, 548-557 (1957).

Zwicker, E. and Scharf, B. A Model of Loudness Summation. Psych. Review 72(1), 3-26 (1965).

Zwicker, E. Procedure for Calculating the Loudness of Temporally Variable Sounds. Journ. Acou Soc. Amer. 62, 675-682 (1977).