RHYTHM PATTERN PERCEPTION IN MUSIC

RHYTHM PATTERN PERCEPTION IN MUSIC:
THE ROLE OF HARMONIC ACCENTS IN PERCEPTION OF RHYTHMIC STRUCTURE.

By

LLOYD A. DAWE, B.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

1993

DOCTOR OF PHILOSOPHY (1993)
(Psychology)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Rhythm Pattern Perception in Music:
The Role of Harmonic Accents in Perception of Rhythmic Structure.

AUTHOR: Lloyd Andrew Dawe, B.A. (University College of Cape Breton)

SUPERVISORS: Drs. John R. Platt and Ronald J. Racine

NUMBER OF PAGES: xi, 177

ABSTRACT

The application of the label music to complex sound requires structure. Musical or rhythmic structure can be thought of as being due to the interaction of two theoretically distinct structures of phrase and metre. Perception of both metrical and phrase structure is dependent not only on the physical structure of the acoustic presentation but also upon cognitive structure being imposed on the auditory sensations. Early work in the psychology of music focussed on establishing the perceptual cues that determine the parsing of music in time. These perceptual determinants can be categorized on the basis of the theoretical components of music: melody, harmony, rhythm, and timbre. With the exception of accent strength based on stability judgments of tones or chords (i.e., structural accenting), phenomenal accents have been assumed by some theorists to be equally-salient, additive, and categorical. The assumption of equal accent strength not only has been applied to different phenomenal accents within a theoretical component category but also between categories. Three series of experiments were conducted to test the assumption of equal weight and additivity of rhythmic cues.

In the first series, a harmonic and a temporal accent were pitted against each other in such a way as to form different rhythm patterns. As well, two harmonic

conditions which varied in the frequency of chord presentations (i.e., the composite-rhythm) but not the frequency of chord changes (i.e., the harmonic-rhythm) were presented. Musicians and nonmusicians were requested to report perceived rhythm patterns in an attempt to determine the relative salience of the harmonic and temporal accents. In addition, a behavioural measure of the perceived metre was taken. Results indicated that the location of chord changes was the main determinant of subjects' rhythmic perceptions and the perceived onset of a measure. As well, although subjects primarily inferred different metres based on the composite-rhythm, an interaction of metrical and rhythmic choices was found indicating that perception of rhythm patterns and inference of metrical structure may not always be independent.

In the second series of experiments, the contribution of harmonic-temporal and harmonic-structural features to the perception of rhythm patterns was investigated by pitting a harmonic and a temporal accent against each other in such a way as to form 5 possible rhythm patterns. Across the experiments, the chord progressions employed were varied, as was the timing of chord onsets (i.e., the composite-rhythm) and changes (i.e., the harmonic-rhythm). In all experiments, musicians and nonmusicians were requested to report perceived rhythm patterns in an attempt to determine the relative salience of the various accents. Results indicated that changes in the composite- and harmonic-rhythm led to a predictable change in an inferred metrical structure, and that all diatonic chord progressions lead to similar patterns of responses in which coincidences of harmonic, temporal, and inferred metrical accents were perceptually salient events. When a nondiatonic chord progression was employed however, there was neither evidence of an inferred metre, nor of responses on the basis of accent coincidence. Overall, musicians were found to primarily report rhythm patterns defined by the location of harmonic accents, while nonmusicians reported rhythm patterns defined by an inferred metrical structure.

In the third series of experiments, the relative contribution of cues for metre inference was determined. In many theories of metre inference, the cues which serve as markers for major metrical accent locations are the basis from which one infers or determines a metre. However, phrase and metrical structure often support one another with phrase boundaries coinciding with metrically important locations. Thus, it becomes difficult to determine which cues, if any, are used exclusively, or predominantly as the basis for metre inference. Three experiments were conducted in which different time-spans defined by harmonic, melodic, and temporal accents, and their coincidences were systematically pitted against one another. Musicians and nonmusicians were requested to identify the metre of the stimuli as belonging to a category of either a triple (e.g., 6/8 or 3/4 time), or a duple metre (e.g., 2/4 or 4/4 time). It was found that musicians use harmonic information much more often and reliably than do nonmusicians who also use the temporal accent to define a metrical structure. Nevertheless, across the experiments, when a harmonic accent was present, subjects used that accent to define the metre. Furthermore, the coincidence of melodic accents was used more often than a temporal accent to determine a metrical structure.

Together the three series of experiments highlight the significant role of harmonic accents in the perception of rhythm patterns in music.

ACKNOWLEDGEMENTS

The research described in this thesis has been reported in three manuscripts submitted for publication in collaboration with Dr. John Platt, and Dr. Ron Racine: the research from Chapters 2 and 3 has been submitted to the journal Perception and Psychophysics, and the research from Chapter 4 has been submitted to the journal Music Perception.

I wish to express sincere gratitude to my supervisors, John Platt and Ron Racine for their guidance and support. In addition, I wish to thank the other members of my supervisory committee, Lorraine Allan and Lee Brooks, who helped me with their comments and encouragement.

Special thanks must also be extended to all my fellow graduate students, in particular, my friends James Gire, and Mike Woloszyn -- who has not only been a great friend but also a fantastic lab partner and roommate. Thank-you.

Finally, it is to my wife Kate, my daughter Denyse, and my son Graeme, that I dedicate this thesis. We have become so much a part of one another that this dissertation is just as much yours as it is mine. The continual love and support which you have given me allowed for the completion of this work.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE PAGE

LIST OF TABLES

CHAPTER I

Introduction

Music can be thought of as a series of simultaneously presented and rapidly changing sounds which must be both interpreted and encoded by a listener if it is to be remembered. Webster (1984) defined music as the art and science of combining sounds or tones in varying melody, harmony, rhythm, and timbre, especially so as to form structurally complete and emotionally expressive compositions. This definition of music explicitly states the four theoretical components of music (i.e., melody, harmony, rhythm, timbre) and alludes to the theory of musical form.

Of the various theoretical components that constitute music, timbre is probably the easiest to define, and yet, is the least investigated. Timbre is that subjective quality of a complex tone which depends primarily upon the physical sound. For example, it is the quality that distinguishes a clarinet from a violin tone of the same pitch and loudness.

In contrast to timbre, melody is one of the most difficult components to define. Generally, it refers to the sequential or serial aspect of music. A very liberal definition of melody would involve variations in both the duration and pitch of successively presented tones. Preliminary work conducted at McMaster University by Mike Woloszyn on the determinants of melody perception indicates that melody at the perceptual level is much more complicated than this definition implies. However, for the purpose of this thesis, such a definition will suffice. In contrast to melody, if a number of tones are presented simultaneously, we are dealing with harmony. While melody refers to the horizontal element of music along the time continuum, harmony is music in its vertical or simultaneous aspect.

At first glance, the distinction between melody and harmony seems obvious. However, perceptually the distinction is not so clear. Harmony refers not only to the simultaneous presentation of tones in isolation (i.e., a chord), but also to the principles governing the sequential combinations of such presentations. As such, the melody-harmony distinction becomes more difficult to make. This is because one of the notes that contributes to a chord's structure, usually the upper-most voice or stream, is customarily part of the melody. This intimate connection between harmony and melody has lead some researchers to operationally define melody as that aspect of the music to which one would sing along. The harmony would be that which accompanies and supports the melody.

The theoretical component known as rhythm has been used to refer to various aspects of music. Several researchers employ the term <u>rhythm</u> in reference to different phenomena. While it is true that all studies of rhythmic structure can be thought of as an analysis of the pattern of time durations that results from various events that occur in a musical flow, researchers often focus on more specific aspects of music. In the most general musical sense, rhythm refers to any form of reiteration. More specific use of the term has been in reference to the regular recurrence of grouped strong and weak beats, or heavily and lightly accented tones (i.e., metre).

Cooper and Meyer (1960) defined rhythm as the way in which one or two unaccented events are grouped in relation to an accented event; and to many, rhythm is used to refer to the temporal aspects of individual notes' and rests' durations.

Some theorists (Agmon, 1990; Boretz, 1971; Lerdahl & Jackendoff, 1983) have thought of rhythm as being multi-leveled, with the structure at each rhythmic level being dependent upon the pattern of time-spans at lower rhythmic levels. Events which differ in perceptual salience create the hierarchy. Highly salient events serve as defining boundaries for long durations high in the hierarchy, while less salient events serve as defining boundaries of shorter time-spans lower in the hierarchy.

This rhythm hierarchy has been discussed by several researchers and theorists (Agmon, 1990; Boretz, 1971; Lerdahl & Jackendoff, 1983). For example, Boretz (1971) discussed rhythm in terms of it being both an insignificant by-product of musical activity (like Agmon, 1990) at the lowest level of the hierarchy and an exalted all-subsuming musical dimension at the highest level. He states:

> And so rhythm is seen as the secondary creation of other aspects of musical perception, an automatic effect of their significant activity. As such, it seems hardly to qualify as a significant activity in itself, except in its most superficial and immediate manifestation...Thus, in the denial of the independence of rhythm, its transcendence is affirmed: the rhythmic structure of a piece is, in the current view, simply all of its musical structure, subsuming every dimensional and inter-dimensional substructure, including as a more or less significant aspect the foreground structure of attack durations. The theory of rhythm, then, is nothing more or less than the theory of musical structure in its most comprehensive form. (p.153)

While a hierarchical view of rhythm established by the parsing of time durations within larger time periods may have some utility, it creates some confusion regarding the use of the term 'rhythm'. One is never quite sure at which level a discussion of rhythm is occurring unless qualification is made.

Thus, for the purpose of this thesis and clarity of meaning, a distinction will be made among the various levels of the rhythmic hierarchy. Starting with the lowest level, the surface duration pattern that results from accents associated with the onset and offset of individual notes and rests is referred to as the temporal aspect of the musical presentation. In keeping with traditional music-theoretical notions, two intermediate-level duration patterns are of importance for the perception of rhythmic structure (Lee, 1985; Lerdahl & Jackendoff, 1983; Palmer & Krumhansl, 1990). The regular reiteration of strong and weak beats or pulses, and the hierarchical pattern of time-spans that results from this regular reiteration is referred to as the metrical structure. The pattern of durations based on salient timbre-based, temporal, harmonic, and melodic phenomenal events (e.g., chord changes, relatively large interval jumps, and rests) which serves as the basis of themes, motives, sections, and phrases is called phrase structure. Finally, the overall musical structure that results from the interaction of phrase and metrical structure is referred to as rhythm.

As such, rhythm can be defined as the complex pattern of successive and simultaneous time-spans that result from the interaction of a phrase and metrical structure, both of which are themselves, to varying degrees, dependent on different timbre-based, melodic, harmonic and temporal events. An adequate delineation of rhythmic structure will depend, therefore, on identification of the perceptually salient temporal, melodic, and harmonic events upon which the complex pattern of time-spans is built, as well as the relationship between phrase and metre.

There has been a substantial amount of research already conducted to determine the perceptually salient accents in music. Of the four theoretical components that constitute music, three have been investigated extensively and will be dealt with in this thesis: the temporal, melodic, and harmonic. Although changes in timbre undoubtedly have a role to play in the determination of rhythmic structure, for the purposes of this thesis, all experiments were conducted using only one timbre: a synthesized piano. In order to investigate the complex domain of music on a scientific level it is imperative that as many variables as possible be controlled. Of the four theoretical components, timbre is the only one that can be eliminated without sacrificing an adequate delineation of rhythmic structure.

The Perceptual Determinants of Rhythmic Structure

Before reviewing the research on the determinants of rhythmic structure, it is necessary to make a few theoretical distinctions which will serve as a basis for categorization of the various perceptual cues.

Lerdahl and Jackendoff (1983) made a useful distinction between 'phenomenal' and 'structural' accents. The distinction is founded upon the view that perception of structure in the world is based on both data-driven (i.e., phenomenal accents) and conceptual-driven (i.e., structural accents) processing. Phenomenal accents have been defined as any event at the musical surface that gives emphasis or stress to a moment in the musical flow (Lerdahl & Jackendoff, 1983). Examples of phenomenal accents include relatively large interval jumps, contour changes, changes in dynamics, relatively long-duration notes, chord changes, and so forth. Structural accents result from more abstract properties and cognitive principles associated with tonal and diatonic organization. Although not explicitly linked to the accent identifications, the work of Krumhansl and her colleagues offer a general basis from which one might identify structural relationships between various notes, chords, and tonal centres or keys (Bharucha & Krumhansl, 1983; Cuddy, Cohen, & Miller, 1979; Krumhansl, 1979, 1990; Krumhansl, Bharucha & Kessler, 1982; Krumhansl & Castellano, 1983; Krumhansl & Shepard, 1979).

It should be noted that even though the phenomenal and structural distinction may be useful for categorization and discussion of various perceptual determinants of rhythmic structure, the distinction, may not be a perceptually veridical construct. The extent to which tonality and diatonicity can be explained via similarities in the temporal periodicity of frequencies, blurs the distinction. If structural accenting has its locus, not in cognitive organization, but rather in the relationship of component sound waves of the tones that make up a diatonic set, then it becomes necessary only to speak of events in the music flow. Although this fact is acknowledged, this thesis makes use of the distinction as an organizational aid.

Phenomenal and structural accents can further be categorized according to the theoretical components of music. The distinction between melodic, harmonic, and temporal components serves to organize the accents into useful categories for purposes of research and discussion.

Melodic Structural Accenting.

The tones used to compose traditional Western music are usually complex harmonic waveforms. With a harmonic complex tone, all of the frequency components (i.e., partials) that make up the tone are simple integer multiples of a not necessarily present fundamental frequency. The corresponding psychological dimension of frequency is pitch. With a complex harmonic waveform the perceived pitch will usually correspond to the fundamental frequency of the presented partials, even if the fundamental is not present. Because frequency is a continuous dimension, there are an infinite number of pitches. However, in traditional Western music pitch is usually perceived according to theoretically-based categories and thus only a relative few fundamental frequencies are used.

In traditional Western music, there are 12 discrete categories of pitch within the span of an octave. Two tones are said to be an octave apart when the interval between them is such that their frequencies are in the ratio of 2:1. The difference between two frequencies an octave apart can be divided into 12 equally-spaced logarithmic intervals. Musically, each of these intervals is called a semitone. Thus, a semitone can be defined as the smallest interval traditionally used in Western music. Each of the 12 semitones spanning an octave are represented on a keyboard by either one of 5 black or one of 7 white keys. The white keys are labelled with letter names ranging from A through to G, inclusive. Each black key is labelled according to its neighboring white keys and a sharp or flat sign. A sharp sign (i.e., #) indicates that the tone is one semitone higher than the letter name indicates. Thus, the black key on a piano that sits directly above a white C key, is identified as C#. The black keys can also be identified through the use of a flat sign (i.e., b) which indicates that the tone is one semitone lower than the letter name indicates. Using the same example, the black key on a piano that sits directly above a white C key ('C#') is also sitting directly below the white D key and therefore, can also be referred to as a 'Db'.

The twelve tones within an octave make up a chromatic scale. The individual notes are called chromatic tones. Each chromatic tone differs in terms of its structural significance within a traditional composition. The melodic structural accent strength of any chromatic tone is dependent upon the position of that note in a tonal hierarchy. The basic notion behind the existence of a tonal hierarchy is that when we listen to music, we hear the sounded elements (i.e., the notes) not as separate units, but rather in relation to one another. In particular, the tonal hierarchy reflects the relationship various tones have to one another in a given key and mode. In Western music, the key is identified by the tonic and by the mode of the scale, which is usually diatonic major or minor.

Major and minor scales make use of 7 of the 12 possible chromatic tones within an octave. These 7 tones are referred to as being the diatonic notes of a scale. Traditionally, the position of each of the seven diatonic notes in a scale is identified by either a label or a Roman numeral. Using Figure 1 as an example, each of the 7 diatonic tones that make up C Major scale can be identified by their appropriate Roman numeral and name: I or tonic is C, II or supertonic is D, III or mediant is E, IV or subdominant is F, V or dominant is G, VI or submediant is A, and VII or leading tone is B.

The mode of a scale is determined by the interval relationship of the diatonic components. If the interval pattern between the seven diatonic notes is tone-tone-semitone-tone-tone-tone-semitone (where a tone is equivalent to two semitones), then we have a major scale that is identified by the letter name of the first note (e.g., the C Major scale represented in Figure 1 is identified by the tonic C). A major scale can be built on any of the twelve chromatic tones of an octave. For example, if the tonic tone was G, following the structure of tone-tone-semitone-tone-tone-tone-semitone would result in the use of an F# rather than an F (Figure 1). In such a case, one would have a G Major scale.

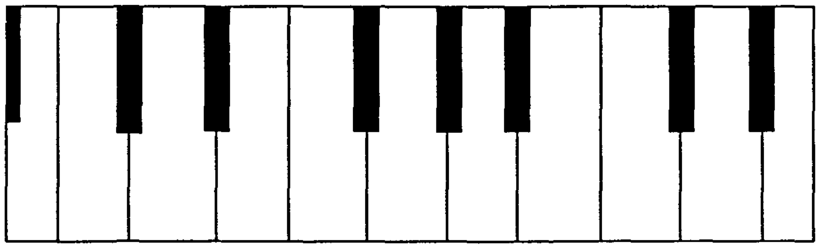
Similarly, a minor scale can be built on each of the twelve chromatic tones, except that with a minor scale the interval pattern of the diatonic tones is tone-semitone-tone-tone-semitone-tone-tone. As with major scales, minor scales are identified by the tonic note (e.g., C Minor Scale). All major and minor scales are related to one another in differing degrees depending upon the number of common tones they share. For example, in Figure 1, C Major and G Major scales have in common 6 of their 7 diatonic notes. The tonal hierarchy captures many of the relationships between the various tones in the context of a major or minor key.

The traditional experimental method that has been used to measure the perceived tonal hierarchy is called the probe-tone or tone profile technique (Krumhansl & Shepard, 1979). With this methodology, each trial of an experiment begins with the establishment of a particular tonal centre or key. This has been done in several ways, including the playing of the major scale (Krumhansl, 1979; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979), a tonic chord (Krumhansl, 1979; Krumhansl & Kessler, 1982), or a chord cadence (Krumhansl & Kessler, 1982). After the key has been established, a single tone from among the twelve of the chromatic scale is presented (i.e., the probe tone) and subjects are asked to judge on a seven point scale how well the probe tone fits within the established key.

Figure 1. The C Major and G Major diatonic scales.

C MAJOR SCALE

NONDIATONIC TONES C# Db D# Eb F# Gb G# Ab A# Bb

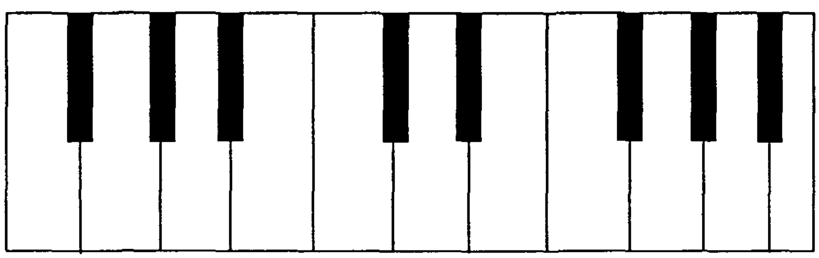


DIATONIC TONES C D E F G A B C

STRUCTURE T T S T T T S

G MAJOR SCALE

NONDIATONIC TONES G# Ab A# Bb C# Db D# Eb F



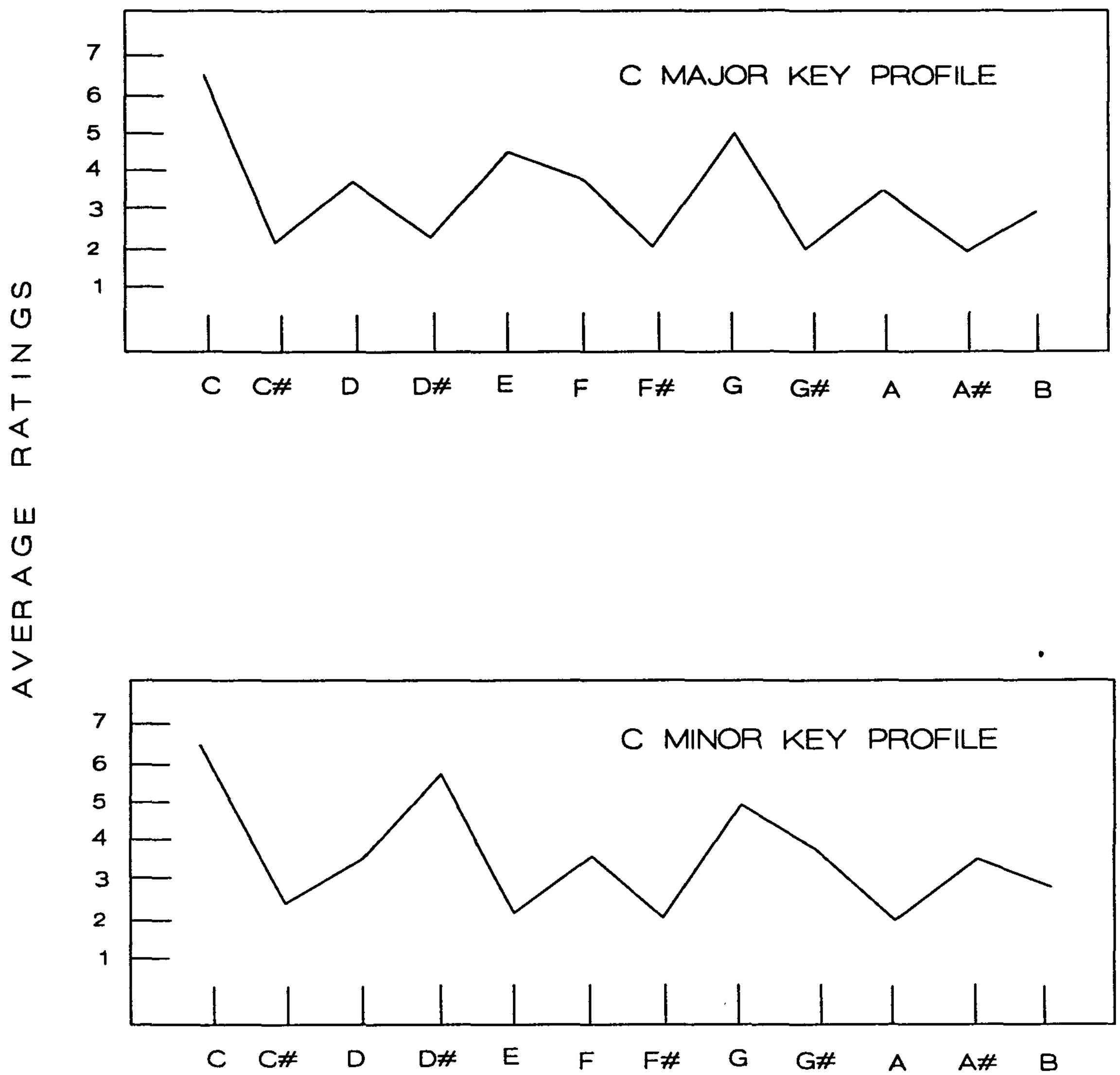
DIATONIC TONES G A B C D E F# G

STRUCTURE T T S T T T S

The typical pattern of results from this probe-tone procedure can be seen in Figure 2. For both major and minor keys, the highest ratings are given to the tonic, followed by the dominant and mediant notes of the scale. Next highest ratings are given to the other diatonic tones and lowest ratings are given to those tones outside the realm of the key (i.e., the non-diatonic tones). The work of Krumhansl and others supports several contentions made by Meyer (1956), who was one of the first theorists to describe this hierarchy of stability. Conventionally, theorists explain stability as an effect of the psychological distance of a tone from the tonic in a prevailing tonality or key (Cook, 1978). Meyer (1956) described the tonic as "a point of gravity." According to Meyer, all other tones gravitate toward the tonic. Movement through the hierarchy towards the tonic creates stability or a sense of completion, whereas movement away results in tension. He also noted that the third and fifth of the scale serve as the higher of two intermediate levels in the hierarchy. The lower intermediate level consists of the other diatonic tones which serve as structural focal points relative to the nondiatonic chromatic tones.

Meyer (1956) further argued that musical styles are complex systems of probability relationships, and that the tonal hierarchy may be related to the statistical properties of the music. Specifically, Meyer thought that the stable tones of a composition would occur more frequently and for longer durations. This would serve the purpose of initially establishing and then maintaining a listener's sense of the key. Krumhansl (1985) described support for Meyer's proposal. She reported that the tonal hierarchy, as determined by the probe-tone technique, had been correlated with a

Figure 2. The tonal hierarchy for Major and Minor scales as determined by the probe-tone technique.



AVERAGE RATINGS
C MAJOR KEY PROFILE
7
6
5
4
3
2
1
C C# D D# E F F# G G# A A# B
C MINOR KEY PROFILE
7
6
5
4
3
2
1
C C# D D# E F F# G G# A A# B

number of statistical analyses of compositions in the Western tonal tradition (Hughes, 1977; Knoppoff & Hutchinson, 1983; Youngblood, 1958). These analyses tabulated the frequency of occurrence or total duration of each tone of the chromatic scale in pieces of Schubert, Mozart, Strauss, Mendelssohn, Hasse, and Schumann. The total duration and frequency of occurrence for each chromatic tone resulted in a profile that strongly resembles the tonal hierarchy for the specific key in which the piece was written. The tonic occurred most often and for the longest durations, followed by the mediant and dominant tones. The average correlation between the tonal hierarchy, as established by the probe-tone technique, and the total durations of the chromatic tones was 0.89. Krumhansl argued that the convergence between the distribution of tones in these compositions and the probe-tone ratings suggests that, in support of Meyer's proposal, listeners have internalized the statistical properties of traditional Western tonal music.

Other researchers, most notably Butler (1983; 1989), disagree with Krumhansl on the latter point, arguing that the similarity between statistical analyses of compositions and the probe-tone's pattern of results indicate that her probe-tone technique is only tapping the statistical properties of a musical presentation and not the tonal hierarchy. Butler and his colleagues have argued that rare intervals (e.g., tritone) are more important for the unambiguous establishment of keys than are tones high in the tonal hierarchy (Brown & Butler, 1981; Butler, 1989). In addition, it has been argued that in Krumhansl's probe-tone technique, the distribution of the key-establishing units gives rise to the pattern of results. Despite the controversy regarding the methodology, there is little doubt that some notes, most notably the tonic, mediant and dominant tones, are psychologically more stable or complete than others within a particular key.

Melodic Phenomenal Accents.

Contour changes. One of the earliest melodic phenomenal accents to be investigated was melodic contour (i.e., the ordinal relations of successive 'ups' and 'downs' in pitch). Pitch contour may yield contour-based melodic accents by segmenting a melody on the basis of rising or falling pitch trajectories. Beginnings or endings of rising or falling trajectories can function as events that determine time-spans (Boltz & Jones, 1986; Thomassen, 1982). Dowling and Fujitani (1971) have shown that exact transpositions of slightly different novel melodies that share the same contour tend to sound like transpositions of the same melody. Others have demonstrated that contour information can aid melody identification (Idson & Massaro, 1978; Kallman & Massaro, 1979; Massaro, Kallman & Kelly, 1980), although this is limited by the fact that different melodies often share the same contour (Watkins & Dyson, 1985). It has been argued that contour information is easily encoded in memory, and that subjects perform very well when tested immediately for contour information (Bartlett & Dowling, 1980; Dowling, 1978; Dowling & Fujitani, 1971). As well, Dowling (1982) cited evidence that the identity of a melody was dependent in part upon melodic contour and others have suggested that contour is important for the communication of structure in music (Jones, 1981; Rosner & Meyer, 1982).

Relative interval sizes. In comparison to melodic contour accents, the interval appears to be equally important for melody recognition. Dowling (1982) argued that contour is easy to extract from a melody but is not as easy to remember as intervals. Contour information seems useful as an "indexical" device to access melodies in long-term storage, but recognition of such melodies seems critically dependent upon scale-step information (Dowling, 1982). In terms of music perception, relatively small successive or melodic intervals (i.e., 1 or 2 semitones) are common in Western music and so when a theme introduces large pitch intervals, it is noticeable (Jones, 1987). Dowling and Harwood (1986) stated that melodies the world over appear to use a great many narrow intervals and few skips, and that this fact may be reflective of a musical and perceptual universal. In studies which have analyzed the frequency distribution of intervals in melodies from numerous cultures (Dowling, 1968, 1978; Merriam, 1964) there is general agreement that intervals larger than 4 or 5 semitones are avoided. When they do occur, they usually do so at a phrase boundary (Dowling & Harwood, 1986).

Dynamic accents. A sudden change in the intensity of a tone will be perceived as a change in the loudness or dynamics. Sloboda (1983) suggested that metrical information may be most unambiguously conveyed by means of dynamic differences. It is also important to realize that dynamics are extremely effective in altering the position of group boundaries and the sense of directed motion within rhythmic groups (Clarke, 1985). Dynamics play a vital role in the perception of rhythmic structure (Schachter, 1980; Yeston, 1976).

Temporal Phenomenal Accents.

Rests and Pauses. A number of workers emphasize the role of timing and pauses for the perception of organized groupings or phrases. Restle (1972) investigated the role of pauses in phrasing of light patterns. In Restle's paradigm, learning of sequential light patterns was facilitated most when long pauses corresponded to major divisions and short pauses to minor divisions in a hierarchical tree description. He argued that phrasing in both speech and music, as well as in light patterns, serves to facilitate learning of hierarchical structure. Seashore (1937) performed an analysis of note durations in singers' performances and found that pauses between phrases were found to be on average four times longer than pauses within phrases. The relative length of a rest or pause may carry important information regarding the structure of the musical sequence. For example, Povel and Okkerman (1982) found that sounds followed by a long silent interval were perceived as accented with respect to sounds followed by a short interval.

Royer and Garner (1966, 1970) attempted to show that timing characteristics were the most important perceptual cues for phrasing in comparison to other forms of cues. They presented subjects with repeating sequences of two elements (e.g., a high pitched buzz and a low pitched buzz) and asked subjects to report the way sequences were perceived. They found that subjects preferred starting points that began with a run of identical elements (e.g., HHHLHLHL) or that produced a pattern ending with identical elements (e.g., LHLHLHHH). If the pattern structure was made more complex by inserting a temporal pause in a nonpreferred starting location, the pause organization was dominant -- subjects would report a pattern structure based upon where the temporal pause was located. Handel (1973) has also argued that if repeating patterns are segmented by temporal pauses, the pattern perception will be based on the structure of the temporal grouping rather than on the structure of the pattern elements.

Relatively long-held notes. Many researchers have noted the effects of temporal structure on perception and memory of phrases. Dowling (1973) performed an interesting experiment which was similar to the paradigm presented in the Royer and Garner studies (1966, 1970). In this study, Dowling was interested in memory for brief melodic phrases. Using a short-term recognition memory paradigm, he presented subjects with short twenty-note melodies made up of four, five-note phrases. Each phrase consisted of four short duration notes followed by a long duration note (i.e., 'SSSSL'). He presented subjects with a test phrase of five notes, asking them whether the test sequence was contained in the immediately preceding melody. He found that memory was enhanced and items were easier to remember if the five notes presented were those within a phrase as opposed to those falling across a phrase boundary.

Perceptually, Boltz (1989) found that "good" endings of phrases will usually be marked by a tonic tone of prolonged duration. She went on to suggest that there may be a common scheme in the auditory environment for using certain structural invariants, such as relatively long durations, for grouping of information. Indeed, there is evidence that this is the case. In monophonic music, notes with relatively lengthened durations are perceived to end a rhythmic group (Garner, 1974; Vos, 1977; Woodrow, 1909). Similar results have also been observed in speech. In declarative statements, the duration of the final syllable is prolonged relative to the preceding context (Cooper, Paccia, & LaPointe, 1978; Lehiste, 1975; Oller, 1973; Scott, 1982).

A demonstration of the perceptual salience of relatively long duration notes was conducted in the laboratory at McMaster University in October, 1989. In this demonstration, the location of a long duration note was systematically adjusted in relation to four short duration notes, resulting in five possible rhythm patterns used as stimuli (i.e., 'LSSSS', 'SLSSS', 'SSLSS', 'SSSLS', and 'SSSSL'). An alternately ascending and descending C Major scale was then presented using each of the rhythms. Each stimulus was made up of the same rhythm pattern repeated eight times to form a 40-note melody. Subjects were asked to choose which of the five rhythm patterns they had heard. There were a total of eight subjects; four were musicians with a mean of 8.5-years experience and four were non-musicians (i.e., they had no formal musical training). The results showed that all eight subjects reported hearing the same rhythm pattern for all trials (the reported rhythm pattern was 'SSSSL'), regardless of where the long duration notes fell in relation to the short duration notes. In addition, there were no differences in the reported rhythmic pattern between musicians and non-musicians; both reported the same rhythmic pattern 100% of the time. The saliency of a relatively long-held note to act as a boundary cue was thus demonstrated.

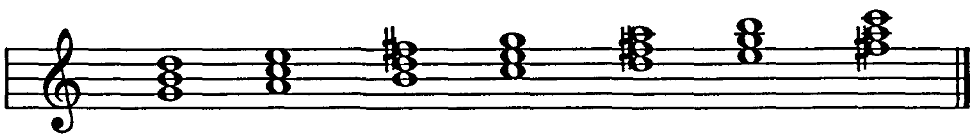
Harmonic Structural Accenting.

Each of the 12 chromatic tones comprising an octave in traditional Western music can be used as a base (or root) upon which chords can be built. The simplest type of chord is a triad. As its name implies, a triad is made up of three tones which are played simultaneously. There are many types of triads, each of which can be defined in terms of the three tones which make up its structure. If the three tones which constitute the triad are diatonic within a particular key, then the triad is considered a diatonic chord of that key. If however, one of the tones is nondiatonic, the triad will be a nondiatonic chord.

Diatonic chords are assigned the same Roman numeral and specialized name as the scale tone which serves as its root. Diatonic triads are built by adding tones of a third and fifth above each scale tone. For example, Figure 3 shows the diatonic triads of a G Major scale. The tonic chord (I) in this example is made up of the tonic (G), mediant (B) and dominant (D) notes of the G Major scale. The dominant chord (V) is made up of the dominant note (D) which serves as the root, a third above this root or the leading tone (F#), and a fifth above the root which is the supertonic (A).

Just as there is a tonal hierarchy for melodic tones, there is a hierarchy based upon the sequential presentation of chords. Harmonic structure plays a central role in traditional Western music, as is evident from the large body of music theory that addresses this aspect of the musical style. If similar principles of psychological organization apply to the perception of musical tones and chords, one would expect similar findings to that which has been found for musical tones when diatonic and nondiatonic chords are presented. By analogy to the experimental results for tones, one would first expect to find a hierarchical ordering of the chords on the basis of stability. Second, there should be a preference for chord sequences (i.e., cadences) ending on chords higher in the tonal hierarchy for that particular composition. Finally, the relations perceived between chords should depend on the tonal context in which they occur.

Figure 3. The diatonic triads for the key of G Major.



G	A	B	C	D	E	F#
I	II	III	IV	V	VI	VII

In chord studies reported by Krumhansl and others, each trial began with a strong key-defining unit such as a scale or a chord cadence (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha, & Castellano, 1982; Krumhansl, Bharucha, & Kessler, 1982). After this unit, two chords were sounded, and the listener's task was to rate on a seven-point scale how well the second chord followed the first in the context provided. In such studies, the resulting tonal hierarchy supports that which would be expected based on the theory of elementary classical harmony (Krumhansl, Bharucha, & Kessler, 1982).

Work by Krumhansl and her colleagues has shown that there are a number of structural features that remain invariant across contexts. Using Roman numerals I to VII to designate the scale step upon which the triad was built, Krumhansl and her colleagues showed that there is a central core of chords (i.e., the I, IV, and V chords). Together these three triads firmly establish a sense of key because, taken as a unit, they contain all the notes of the major diatonic scale from which they are built (Stark, 1988). Furthermore, this central core of chords captures the similarity of three closely related keys.

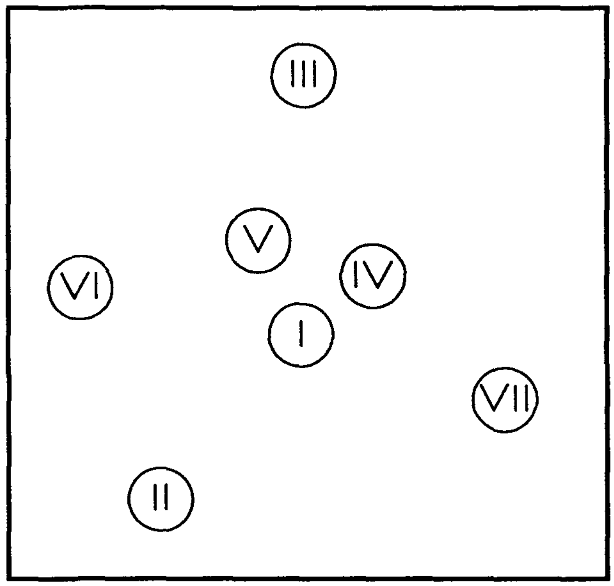
A specific example may help to make this point clear. In the key of C Major, the tonic (I) triad is the C Major chord, the subdominant (IV) triad is the F Major chord, and the dominant triad (V) is the G Major chord. These three chords sound relatively stable in the key of C Major. However, the F and G Major chords not only serve as the subdominant and dominant chords respectively in the key of C Major but also as the tonic triads in their respective keys of F and G Major. The two keys of F Major and G Major are closely related to the key of C Major. Recall that a major scale is established by playing a series of interval patterns represented by the scheme tone-tone-semitone-tone-tone-tone-semitone. If we do this starting on the note G to establish a G Major scale (see Figure 1, p.11), or an F to establish an F Major scale, we would find that both scales have 6 out of 7 diatonic notes in common with the C Major scale. Indeed, F and G Major scales are the two most closely related Major keys to C Major.

Figure 4 shows both the multidimensional scaling solution and the hierarchical clustering of the relatedness ratings, from Krumhansl, Bharucha, and Kessler (1982). The multidimensional scaling showed that there was a core of structurally significant chords - the I, IV, and V - which was surrounded by less stable chords: the II, III, VI, and VII. Similarly, the hierarchical clustering method applied to the same data showed that the highest ratings were received by the tonic and the dominant chords (i.e., the I and V chord respectively). This cluster was successively joined by the IV, VI, II, III, and VII chords. This pattern of results was found to be largely invariant across different tonal contexts (Krumhansl, Bharucha & Castellano, 1982).

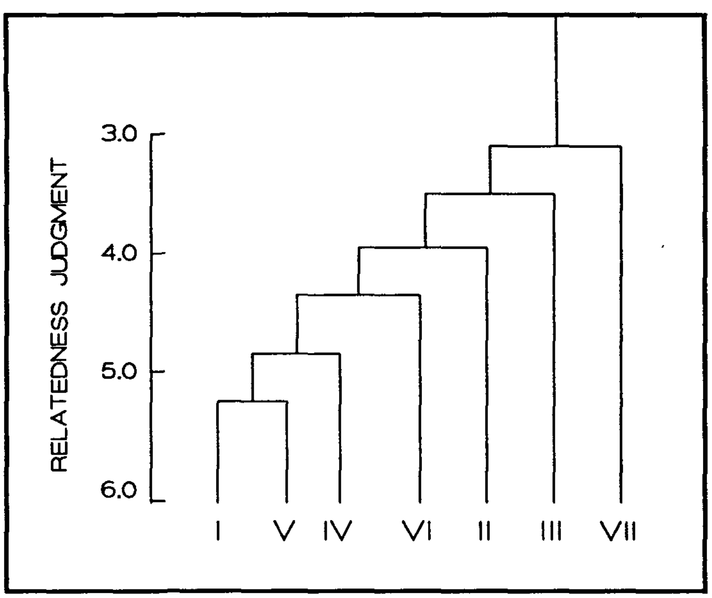
Nondiatonic chords have also been found to cluster into groups close to the diatonic chords to which they are most closely associated (e.g., through similarity of scale tones for keys in which the chords are representative). For example, the keys of C Major, G Major and A Minor are closely related and thus have several diatonic chords in common. The overlapping chords have been found to be close to one another in a multidimensional scaling solution while the clustering of diatonic chords for each of the keys was maintained (Krumhansl, Bharucha & Kessler, 1982). As

Figure 4. The multidimensional scaling solution and the hierarchical clustering solution of the relatedness ratings of chords from a key (from Krumhansl, Bharucha, & Kessler, 1982).

MULTIDIMENSIONAL SCALING



HIERARCHICAL CLUSTERING



well, when chords are drawn from two maximally distant major keys, analyses of the relatedness ratings reflected a clear separation of the chords of these two nonoverlapping keys (Bharucha & Krumhansl, 1983).

Within a particular key, movement towards the more stable chords has been rated as being more stable than movement away from them (Krumhansl, Bharucha & Castellano, 1982). Chord progressions that move towards the tonal centre are generally perceived as being the most stable. For example, movements from secondary chords (i.e., VII, VI, III, II) to primary chords (V, IV, I) conventionally have been seen as being satisfactory, while those that move in an opposite direction have been seen as being less stable. It is the harmonic principles based upon such findings that support the music theoretical notion that phrasing is based at least in part upon harmonic progression. A section of music ending on the I, IV, or V will be viewed as being more stable or complete than those that do not. Presentation of less stable chords will create tension that must be resolved by playing the more stable chords. Thus in a musical composition, chord progressions create a sense of movement and finality to varying degrees throughout the piece.

Harmonic Phenomenal Accents.

A simple chord change can serve as a phenomenal accent. Traditionally, researchers have emphasized the structural facet of a harmonic presentation (e.g., the hierarchical ordering of harmonic stability) rather than the phenomenal aspects of chord presentations. With structural accenting of a chord presentation, the emphasis is placed on how the harmonic event relates to other chord presentations either in the future or the past within a tonal context (e.g., chord progressions). In contrast, concern with the phenomenal accenting of a chord change is with the event itself, apart from any tonal relationships. While most researchers are undoubtedly in agreement that chord changes serve as phenomenal accents, research on the phenomenal accenting of chord changes is lacking. This lack of research is unfortunate, especially since music theory describes a role for the durational pattern of time-spans defined by the phenomenal event of a chord change. Music theorists refer to the harmonic-based durational pattern as harmonic-rhythm. Theorists believe the choice of metrical structure is related to the timing of harmonic phenomenal accents although the exact relationship is still unknown (Lerdahl & Jackendoff, 1983). The relationship of harmonic-rhythm to grouping in general and, more specifically, to metrical structure remains an important empirical question.

This treatise deals with two types of harmonic phenomenal accents: that associated with the presentation of a chord, and that associated with a change of chord. Both types of events are based on a chord's onset, with the distinction resting on whether the chord being presented is the same or different from that previously presented. The distinction between these two types of harmonic phenomenal events, although subtle, is important not only theoretically, but also functionally. While it is true that all phenomenal accenting associated with a change of chord is also associated with the accenting due to chord onset, the reverse is not necessarily true. That is, one can have multiple presentations of the same chord. This leads to a consideration of two important constructs in most music-theoretical accounts of rhythm: harmonic- and composite-rhythm.

Harmonic- and Composite-Rhythm.

Harmonic-rhythm is the sequential pattern of durations provided by either chord changes, or implied triadic patterns within a melody (Smith & Cuddy, 1989; Piston, 1948). Composite-rhythm is the sequential pattern of durations associated with the onset of individual note and chord presentations in a polyphony. To determine the composite-rhythm one need only initiate a new timed duration whenever a note in the polyphony begins or ends. In short, the composite-rhythm will consist of a series of the shortest duration notes across the lines of music.

As with the distinction between the two harmonic phenomenal accents discussed above, harmonic-rhythm contributes to and is part of the composite-rhythm, but the composite-rhythm can display serial time patterns not apparent in the harmonic-rhythm. Thus, while it is theoretically possible for the harmonic- and composite-rhythms in a musical presentation to be identical (e.g., when every note's onset creates a new chord), it is usually the case that the harmonic-rhythm represents higher-order or longer temporal relationships.

Several theorists and researchers have focussed on the role of harmonic-rhythm in the perception of a metrical structure (Cooper & Meyer, 1960; Piston, 1948; Smith & Cuddy, 1989). The consensus is that in Western tonal music, harmonic changes coincide with important metrical locations, especially the first beat of the measure. Far less empirical research has been conducted on the role of composite-rhythm in perception of metre, phrase, and rhythm. According to several theorists, however, the serial durations created by individual note onsets, in addition to the tempo (i.e., the speed of presentation) will determine the perception of iterative pulses

(Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1982, 1984; Longuet-Higgins & Steedman, 1971; Steedman, 1977), from which one can infer a metrical structure. In addition, some theorists believe the choice of tactus and metre is related to the timing of harmonic accents. The tactus corresponds to the metrical level at which a listener taps his foot or a conductor waves his baton (Lerdahl & Jackendoff, 1983). Occasionally, the tactus corresponds with the notated beat of the music (e.g., in 3/4 time a listener will tap his/her foot three times per measure), but this is not always the case (e.g., in 6/8 time, a listener usually taps his/her foot only two times per measure). According to Lerdahl and Jackendoff, a piece of music with frequent harmonic changes (viz., the harmonic-rhythm) is heard with a faster tactus than a piece with equal note values (i.e., the same composite-rhythm) but less frequent harmonic changes. However, the specific relationship between harmonic- and composite-rhythm has yet to be investigated. The parsing of time by note and chord onset events within longer durations created by chord changes may be an important factor in the perception of rhythmic structure.

Summary and Conclusions

The auditory system is amazingly good at detecting changes in a complex sound such as speech or music (Moore, 1989). Any change in our physical environments can be defined as an event, and events differ in terms of their perceptual salience. For example, a door slamming could be viewed as a perceptually more salient event than the drone of an air conditioner as it is turned on. However, the perceptual salience of an event is defined not only by the amount of physical change but also in terms of the context in which that change occurs (i.e., it is a relative phenomenon). The expectations, goals, and cognitions of the perceiver play an important role in the perceived salience of an event. Thus, it is not only the picking-up of salient physical changes in the musical flow (i.e., phenomenal accents) but also an active interpretation of the meaning of these changes within a perceiver's cognitive framework (i.e., structural accenting) that will determine the perceived rhythmic structure of music.

Melodic, harmonic and temporal events combine to form higher-order time-spans that constitute rhythmic structure. Theorists and researchers differ in opinions regarding how phenomenal and structural accents combine to determine phrase and metrical structure. As well, there is considerable debate as to how phrase and metrical structure relate to one another in the determination of the overall rhythmic structure of a musical presentation. A thorough discussion of these views will be presented in the following sections.

The Perception of Phrase Structure

The definition of music presented at the beginning of this thesis not only mentioned the theoretical components of music, but also alluded to the theory of musical form or structure. Looking at the theory of musical form we find that music can be broken down into component parts in much the same way as poetry or literature. A musical idea is presented in a 'sentence' that is ended with a 'period', while the phrases that make up the sentence are separated by 'commas'. A composition generally presents a number of musical 'sentences' or 'ideas' repetitively, with each repetition involving some changes in how the theme is presented. Early work in the psychology of music was directed towards discovery of the perceptual cues that serve as 'periods' and 'commas' (i.e., the phenomenal accents). These cues have often been referred to as 'phrase boundary cues'.

While the analogy between the structures of literature and music is useful, it is not exact. With literature, there may be presentation of ideas, and occasionally, certain central motives will be repeated, but literature is not as dependent upon the principle of reiteration as is music. With music it is through 'repetition' and 'contrast' that a composer is able to communicate his musical idea or theme. Furthermore, it is important for a composer to provide a judicious mixture of both these principles if his music is to be successful. Either of these principles, if followed exclusively, would be artistically unsuccessful (Bray, Snell & Peters, 1967). In traditional Western music, this goal has been accomplished by repeatedly presenting the same idea or theme with subtle variations of the melodic, temporal, and harmonic features with each repetition.

The idea of theme and variation or invariance and transformation is an important and useful concept both for psychology and for the study of musical pattern structure (Jones, 1981a, 1981b, 1982; Welker, 1982). Jones and her colleagues (Jones, 1981a, 1981b, 1982, 1984, 1987; Jones, Boltz & Kidd, 1982; Jones, Kidd & Wetzel, 1981) have developed a model of music perception called 'the theory of dynamic pattern structure' that shows how the various phenomenal accents reviewed thus far may combine to determine not only phrase boundaries, but also the structure of the music within phrases and the structure of a theme. It must be noted that the theory of dynamic pattern structure is not definitive nor complete, but rather was presented as a plausible initial description (Jones, 1987).

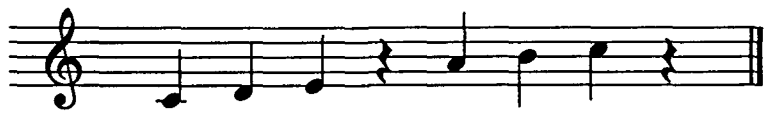
Jones (1987) argued that both theme recognition by a listener and the functional pattern similarities in theme variations as presented by a composer could usefully be conceived in terms of the dynamic pattern structure. The dynamic pattern structure refers to how the on-going musical flow is broken into phrases and time-units on the basis of the pattern of accents and accent couplings or coincidences.

This theory has so far only been applied to the interplay of melodic and temporal accents in an attempt to understand how the various accents combine to form a higher-order dynamic structure that Jones has called a 'joint-accent structure'. The joint-accent structure refers to a particular combination of accents, accent strengths, and time symmetries that result when melodic and temporal accent patterns combine (Jones, 1987). Figure 5 shows an example of a joint-accent structure analysis. Variations in accent strength occur because of coincidences of accents in time. The notion of time symmetries and asymmetries in the joint-accent structure refers to the hierarchical temporal regularities and irregularities that result from the phase relationship between temporal and melodic accents.

In Figure 5 it should be noted that melodic accents (i.e., 'm' accents) occur on the low 'C' which is the first note of the piece of music and on the 'A' because of the relatively large interval jump from the previous note 'E'. The temporal accents (i.e., 't' accents) for two different rhythms are also given in Figure 5. In the first example, the one beat pause creates two 3-tone groups and results in a 't' accent falling on the first and fourth tones. The melodic and temporal structures combine to determine the joint-accent structure. The coincidences of accents ('m' plus 't') yield

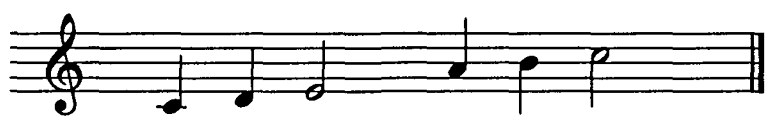
Figure 5. Two examples of a joint-accent structure.

a). Example 1: Rhythm is SSS REST SSS



Notes:	C	D	E		A	B	C'	
Beats:	1	2	3	4	5	6	7	8
Melodic Accents :	m	.	.	.	m	.	.	.
Temporal Accents :	t	.	.	.	t	.	.	.
Joint Accent Structure:	a'''	a'	a'	a	a'''	a'	a'	a

b). Example 2: Rhythm is SSL SSL ...



Notes :	C	D	E		A	B	C'	
Beats :	1	2	3	4	5	6	7	8
Melodic Accents :	m	.	.	.	m	.	.	.
Temporal Accents :	t	.	t	.	.	.	t	.
Joint Accent Structure :	a'''	a'	a''	a	a''	a'	a''	a

stronger accents than would an 'm' or 't' accent alone. The accent strength in this example is represented by the number of primes (a''' > a'', etc.). In the second example, the temporal and melodic accents do not coincide, resulting in a much more complex joint-accent structure. The resultant accent couplings and differences in accent strength define or mark higher order time spans and phrase boundaries within a musical composition.

Jones (1987) argued that it is the joint-accent structure that determines not only phrase boundaries but also ease of melody or theme recognition. She states:

> Theme recognition is determined by the extent to which recurrent patterns preserve psychologically important structural invariances of the melody. These functional similarities, in turn, will depend on what initially captures a listener's attention and on how these relations are reinstated later during some recognition test... Specifically, if people target their attention over distinctive higher order time periods, associated with combined accents, then commonalities in joint accent structures may provide a framework for understanding dynamic pattern similarity. (p. 626).

Jones, Summerell and Marshburn (1987) found some support for Jones' theory in this regard. They found that the joint relation of melodic <u>contour</u> accents and temporal accents combined to specify a tune's dynamic shape. The concept of dynamic shape refers to that part of the joint accent structure that involves the relative timing of a melody's ups and downs in pitch. The specific hypothesis being tested by Jones et al. (1987) was that people recognize musical themes by using the dynamics of melodic contour accents. Their research found that confusion of themes occurs as a function of similarity of the joint-accent structure. A series of experiments used, as standard melodies, tunes that had strong accents based on couplings of melodic contour and temporal accents. When comparison melodies maintained the placement of melodic contour accents and changed the location of temporal accents, subjects were more likely to confuse the experimental melodies with the standards, compared to when both temporal and melodic contour accents were changed. Jones explained this result in terms of the joint-accent structure. The more similar two melodies are in terms of their joint-accent structure, the more they will be confused with one another.

Jones' theory is useful for studying not only dynamic pattern similarity or theme recognition, but also dynamic pattern simplicity or complexity (Boltz & Jones, 1986; Jones, 1987). The complexity of a piece of music is determined in Jones' model by the phase relationship between temporal and melodic accents and their couplings in the joint-accent structure. Jones (1987) argued that melodies that have a large number of accent couplings should be easier to remember than melodies with complex joint-accent structures. Thus, the second melody in Figure 5, which is an example of a more complex dynamic pattern than is the first melody, should be more difficult to remember.

It must be noted that Jones' theory is not without its difficulties. In particular, all accents are equally-weighted. She defines accents as any salient attention-getting event, yet fails to allow for the possibility that some accents could be phenomenally more salient. Jones' notion of accent strength is tied more to the coincidences of various equally-weighted accents (i.e., all melodic and temporal accents contribute the same amount of accent strength to the joint accent structure). There is no reason to assume that events occurring in a musical flow will be equally salient. For example, the Gestalt principle of figure-ground would predict that accents corresponding to a figure (e.g., melodic accents) may be perceptually more salient than accents occurring in the ground (e.g., harmonic accents). Alternatively, traditional Western music emphasizes the fundamental role of harmony in establishing rhythmic structure. From this perspective, one would expect harmonic accents to be perceptually more salient.

Jones' theory is Gibsonean in the sense that accents interact in the physical stimulus and it is the resulting accent structure that we experience. In this way, the notion of a phenomenal accent strength belongs at the level of the joint-accent structure. Jones' notion of an accent is categorical at the level of the individual melodic, temporal or harmonic accent. That is, because there is no gradation of accent strength at this level, Jones must present individual units (i.e., notes) as being either an accent or not. Such a state of affairs leads to conceptual problems when she includes tonal accents (i.e., structural accents) as examples of melodic phenomenal accents. It is evident in her writing that she is also aware of this difficulty since in her example (see Figure 5), she ignores tonal accent strength, noting that "we assume no tonal accenting, a dubious assumption in this case because notes E and C could, arguably, function as tonal [melodic] accents" (p.626, Jones, 1987). Since the example presented was in the key of C Major, the reader is left to assume that what Jones means is that only notes high in the structural accent hierarchy would be considered as melodic accents. The problem now arises as to how structural accents, which by definition differ in terms of accent strength, are incorporated into the realm of a categorical accent model. If a cutoff is required, a question can be raised as to where that cutoff point is placed and why. From Jones' discussion, it is implied that the first two levels of the structural accent hierarchy (i.e., the 1st, 3rd, and 5th) should be included within the categorical boundaries of melodic accents, leading to the further difficulty of accounting for the structural distinction between the tonic and the mediant or dominant tones.

A slightly different but related problem with Jones' model is that she places the locus of melodic accenting within the physical stimulus. If structural accenting is to be included in Jones' theory as a form of melodic accenting, then the tonal hierarchy must be assumed to be communicated via the physical parameters of a sound. This places Jones' theory within the realm of direct realism . The tonal hierarchy was developed within the cognitive-structural tradition (see Cross, West & Howell, 1991). Although most researchers think of structural accenting as a cognitive or top-down phenomenon, it may be possible to develop a theory based on direct realism principles in which the tonal hierarchy is based on bottom-up processing of stimuli. However, such an endeavor is yet to be done successfully.

Some of these difficulties have led Jones to modify her original model of joint accent structure. Recently, Jones (1992) has modified her views to include the notion of accent weights - an acknowledgement that phenomenal accents can differ in their perceptual salience. However, the notion of equally-weighted accents which combine in an additive fashion is not limited to Jones' (1987) theory. In fact, the idea of accent strength being determined by accent coincidences is consistent with the views of several other researchers (e.g., Benjamin, 1984; Berry, 1976). There are two levels at which one can think of differences in phenomenal accent strength: between and within the theoretical components of music. An example of the former would be that harmonic events are perceptually more salient than temporal or melodic accents.

As an example of the latter, one could question whether a melodic contour accent is as perceptually salient as a relatively large interval jump. Jones' theory views all phenomenal accents, both within and between the theoretical component categories, as equally salient. Other theorists (e.g., Palmer & Krumhansl, 1987a; 1987b) acknowledge a distinction of perceptual salience within a theoretical component category but argue for an equal and additive model between categories.

Palmer and Krumhansl (1987a, 1987b) presented evidence that temporal and melodic features of a melody combine both additively and independently to determine the perceived phrase structure. They presented subjects with stimuli derived from four measures of J.S. Bach's Fugue XX in A Minor (Palmer & Krumhansl, 1987a) or the eight measures of the opening theme of Mozart's A Major piano sonata (Palmer & Krumhansl, 1987b). Three conditions were created by presenting only the pitch pattern using equitemporal durations, only the temporal pattern using equitonal notes, and both pitch and temporal patterns together (i.e., the melody). Trials were generated by altering the length of the various conditions so that each trial ended on a different pitch, or temporal event, or both. Subjects were requested to rate on a seven point scale how good or how complete a phrase each of the segments made. Within both the temporal and pitch conditions, each location received different mean ratings indicating that subjects perceived differences in the relative stability of each note. Thus, in terms of stability at least, there is some evidence that different phenomenal events will differ within a theoretical component category.

The authors also reported that the two conditions contributed equally and additively to the perception of the melody. Simply adding the two mean ratings for each location of the temporal and pitch conditions predicted the ratings which were found for the melody condition. It should be noted however, that the perceptual independence of temporal and melodic features which Palmer and Krumhansl reported is questionable on the grounds that the stimuli they used may have allowed subjects to temper their ratings for each location relative to the other locations in each of the three categories. In both studies, the musical presentation consisted of a repeated musical phrase. In the first study (Palmer & Krumhansl, 1987a), the stimuli were made from four measures of a Bach fugue with the first and last two measures similar in terms of the temporal and pitch components. In the second study (Palmer & Krumhansl, 1987b) the stimuli were made from eight measures of a Mozart sonata in which the first and last four measures were practically identical. In both cases, the trials differ in terms of their length, allowing for subjects to rate each note. However, even the shortest trials in each study included more than the first half of the segment from which the stimuli were derived. This may have allowed subjects to become aware of the structure of the repeated phrase. If this occurred, subjects may have reported ratings based not on how well the final note of that trial would serve as an ending, but rather on how well the final note of the trial served as an ending relative to the other possibilities. This would undoubtedly not only influence the pattern of ratings for each location, but also leave any conclusions based on the relationship between the patterns of results for various conditions in doubt.

One of the main difficulties with these theories has been the exclusion of harmonic phenomenal events in both theoretical discussions and investigations of phrase structure. It is possible that harmonic events are perceptually more salient than events within the other theoretical component categories (i.e., melodic, timbre, and temporal). Theorists have long emphasized the fundamental role of harmonic functions in traditional Western music. Indeed it was undoubtedly this fact that lead to abundant research directed at the structural relationships of harmonic events (e.g., Bharucha & Krumhansl, 1983; Krumhansl, 1979; Krumhansl, Bharucha & Kessler, 1982; Krumhansl & Castellano, 1983; Krumhansl & Shepard, 1979). As well, it is possible that the two types of harmonic phenomenal accents associated with chord onset and a change of chord, discussed above, are both perceptually and functionally distinct.

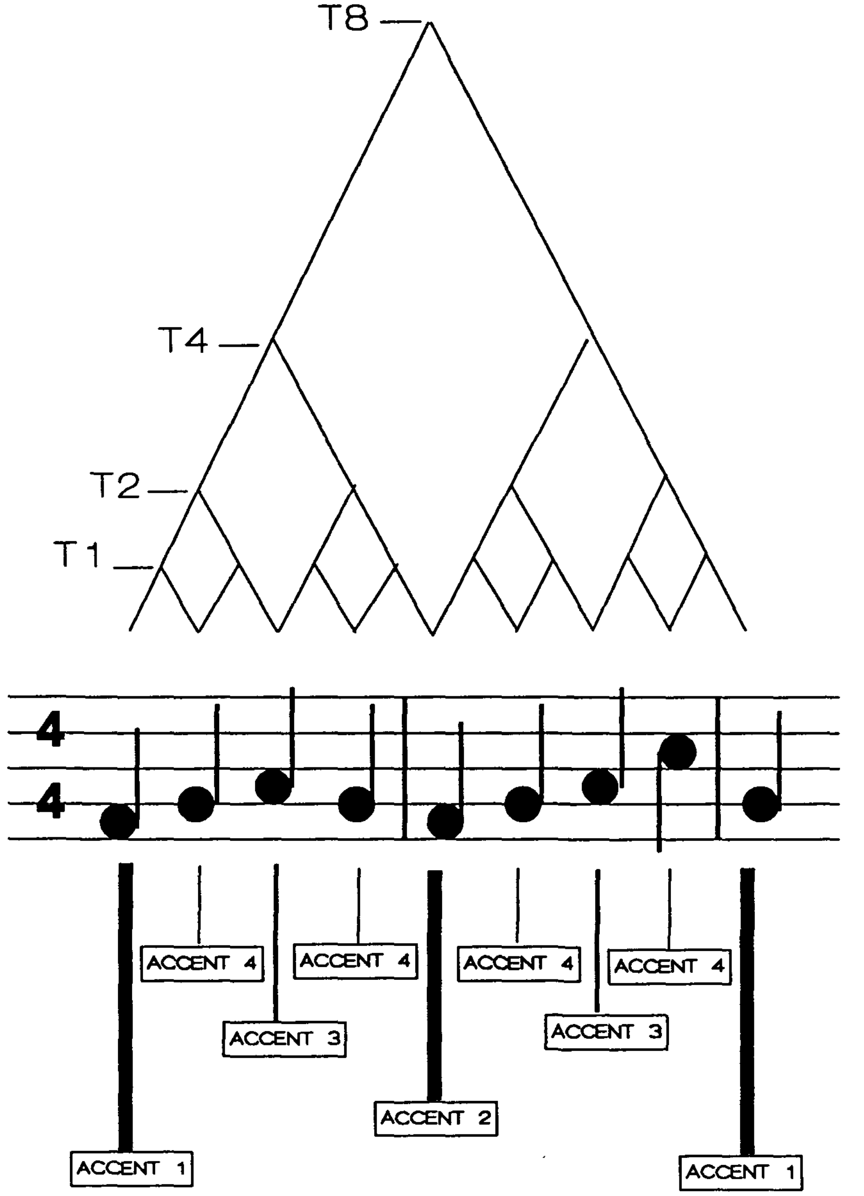
The Perception of Metrical Structure

All music is perceived within a metrical context. Metre is the division of regular pulses into groupings of strong and weak beats. Metre is represented in written music by the time signature which lets the musician know where the strong and weak beats usually occur. For example, 3/4 time means the music is broken up into groups of three beats (i.e., the '3') with the quarter note receiving one beat (i.e., the '4') whereas 6/8 time means the music is broken up into groups of six beats with the eighth note receiving one beat. Perceptually, the first beat of each measure is the most strongly accented beat within that measure. Usually, the strength of accents alternate between strong and weak. In 4/4 time, for example, the first beat is the strongest followed by a weak second beat. The third beat is strong, but not as strong as the first, and the fourth beat is the weakest. However, there are exceptions to this alternation of strong and weak beats. For example, 6/8 time has its strongest beat in the first position followed by its next strongest beat in the fourth position with the second, third, fifth, and sixth beats being weak. In this way it sounds very similar to 3/4 time, with an accent occurring every third beat, except that with 6/8 time the fourth beat is not quite as strong as the first.

Metrical structure can be represented visually in the form of a hierarchical tree. To illustrate, imagine two measures of 4/4 time, as shown in Figure 6. The span of time between beats can be thought of as a unit time T_1 (Jones, 1981). The span of time that constitutes a measure in 4/4 time is represented in Figure 6 by T_4. Similarly, half of one measure (i.e., the time span of a half note in 4/4 time) is represented by T_2. By representing metrical structure in the form of a hierarchy, one emphasizes the fact that every fourth beat (i.e., the first beat in each measure) in the figure is accented, forcing into prominence higher-order periods which Jones (1981) calls 'accent subsequences'. In addition, higher level metrical structure (i.e., metrical structure at the hypermeasure level) is represented by T8. Tonal music often has from one to three levels of metrical structure that are larger than the level notated by barlines, corresponding to regularities of two, four, and sometimes even eight measures (Lerdahl & Jackendoff, 1983).

Metrical accent strength is represented by the point of origin in the tree. Stronger accents originate higher in the hierarchy. Jones (1981) pointed out that a structure such as a 4/4 time metre has an inherent perceptual ambiguity. A sequence of notes in a 4/4 time metre can be expressed either in lower order periods by T_1 or in

Figure 6. A metrical time hierarchical tree for 4/4 time.

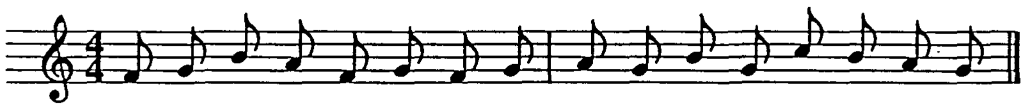

T8
T4
T2
T1
4
4
ACCENT 4
ACCENT 4
ACCENT 4
ACCENT 4
ACCENT 3
ACCENT 3
ACCENT 2
ACCENT 1
ACCENT 1

terms of higher order periods by T_4. As well, hypermeasure time units may be used (e.g., T_8). From a theoretical perspective, this ambiguity illustrates flexible attention focusing. Jones (1976, 1981) argued that attention is fundamentally rhythmic. It can be focused upon relationships between neighboring events by synchronizing appropriate subjective time rules with lower level temporal periods within a pattern, such as T_1. Conversely, attention may lock into a higher order period such as T_8, T_4, or T_2. Because these time periods are multiples of one another, Jones (1981) argued that the ambiguous time structure affords the shifting of attention between lower and higher periods. A bottom-up attentional strategy involves heightening subjective rhythms corresponding to T_1 initially and then expanding to T_4 or T_8. A top-down attentional strategy involves locking into a higher period such as T_8 or T_4 first and then contracting to pick up finer details at the level of T_1.

This can easily be experienced when listening to music. If a melody is presented in 4/4 time, attention can be focused on higher order periods, in which case the subject's foot may tap at the beginning of each measure. Alternatively, attention could be focused to lower order periods such that the foot would be tapping on every beat or second beat. If a metre has such a doubling structure (e.g., 4/4 time), the duration of the perceptually salient unit of time -- which has been referred to as the 'tactus' by Lerdahl and Jackendoff (1983) -- can easily shift from one unit-time to another. Jones' (1981) principles of top-down and bottom-up attentional strategies allow one to attend to either smaller or larger unit-times respectively. For example, one could imagine tapping one's foot 2 or even 4 times per beat.

Not all metres have such a doubling structure however. In some cases, such as in 3/4 time where there is an accent every third beat, if the tactus changes, so also does the metre. The theoretical work of Lerdahl and Jackendoff can illustrate this. Lerdahl and Jackendoff (1983) tried to capture the hierarchical nature of metrical structure through the principle of 'time-span'. They preferred to use dots as opposed to the lines in a hierarchical tree to represent beats, arguing that beats, as such, do not have duration. Dots are a spatial analogy to the auditory phenomenon of beats. However, the interval of time between successive beats (i.e., the time span) does have a duration. Lerdahl and Jackendoff argued that in order for a time-span to be perceived as being metrical, the time duration between successive beats would have to be equally spaced. This is because one of the functions of metre is to mark off the musical flow into equal time-spans.

To capture the hierarchical structure of metre, Lerdahl and Jackendoff constructed 'metrical grids' that consisted of time-spans at various levels of the hierarchy. At each level, the time-spans have to be equally spaced, with higher levels representing larger time-spans and stronger beats. An example of a metrical grid for 4/4 time is given in Figure 7. In this notation, a stronger beat appears with more dots under it. Thus in Figure 7 it can be seen that in 4/4 time the first and third beats are stronger than the second and fourth beats. In addition, the first beat of every measure is stronger than the third beat. Although not illustrated in Figure 7, lower levels within a measure and hypermeasure metrical levels are also possible. In fact, an infinite number of metrical grids could be generated for any time signature. One could add smaller and smaller levels provided that the time-span between the 'beats' is evenly spaced. In addition one could add 'hypermeasure' levels to the metrical grid.

Figure 7. A 4/4 time metrical grid (after Lerdahl & Jackendoff, 1983).

Beats: 1 2 3 4 1 2 3 4

Metrical Grid: . . . .

. . . .

. . . .

. .

Lerdahl and Jackendoff (1983) pointed out that at very low levels, metrical structure becomes meaningless. Since the purpose of metrical structure is to mark off the musical flow of notes into regular durations, time-spans shorter than the shortest duration note become useless. It is for this reason that Lerdahl and Jackendoff stated as a general rule, that the shortest duration note usually serves to define the lowest level of the metrical grid. At very large hypermeasure levels, metrical structure loses its regularity. This is because phrase structure has a tendency to define time-spans at the larger hypermeasure levels. In terms of the highest level, Lerdahl and Jackendoff suggested two or three levels above the tactus as adequate.

Lerdahl and Jackendoff (1983) made note of an important limitation on metrical grids for classical Western tonal music. In all metres of traditional Western music, the time-span between beats at any level can only be two or three times longer than the time-spans at the next lower level. For example, in 4/4 time (see Figure 7), the time-spans consistently multiply by 2 from level to level as you go higher in the metrical hierarchy. In 3/4 time however, the hierarchical pattern of time-spans is such that the time-span doubles from the lowest level of the hierarchy to the intermediate level and then triples from the intermediate level to the highest level, as represented in Figure 8. In 6/8 time the pattern is reversed. The time-span first triples and then doubles (Figure 8).

It is important to note that in Figure 8 the beats as defined by the time signature (i.e., Jones' T units) are represented above the metrical grids by Arabic numerals and that the same sequence of notes are presented in the two examples. In Jones' notation, an attentional shift down one T unit in 3/4 time will result in the

Figure 8. A contrast between a 3/4 time metrical grid and a 6/8 time metrical grid (after Lerdahl & Jackendoff, 1983).

a). A 3/4 time metrical grid.

Beats:	1		2		3		1		2		3		1
Metrical Grid:	.	.	.	.	.	.	.	.	.	.	.	.	.
	.		.		.		.		.		.		.
	.						.						.

b). A 6/8 time metrical grid.

Beats:	1	2	3	4	5	6	1	2	3	4	5	6	1
Metrical Grid:	.	.	.	.	.	.	.	.	.	.	.	.	.
	.			.			.			.			.
	.						.						.

perception of a different metrical structure (i.e., 6/8 time). Similarly, an attentional shift up one T unit in 6/8 time will result in the perception of a 3/4 time metre. This is represented in Figure 8 in the level of the metrical grid at which the beats (the Arabic numerals) for the two examples correspond. In 3/4 time, the beats correspond to the second level of the metrical grid, while in 6/8 time, the beats correspond to the lowest level of the metrical grid.

As a final note, it is important to realize that the beat, as defined by the time signature, need not be the tactus (Lerdahl & Jackendoff, 1983). The tactus is the level at which a listener will tap his foot or a conductor will wave his baton. The tactus almost invariably will correspond to an intermediate level in which the beats pass by at a moderate rate (Lerdahl & Jackendoff, 1983). In 3/4 time the tactus (and intermediate level of the metrical grid) corresponds to the beat given in the time signature, and thus a listener will usually tap his foot three times per measure. In 6/8 time however the tactus does not usually correspond to the beat given in the time signature, rather a listener will more often than not tap his foot only two times per measure (i.e., he will use the next metrical level above that given by the time signature).

The Nature of Perceived Metrical Structure and its Relationship to Phrase Structure

The structuring of music according to themes, sections and phrases is theoretically different than the structuring of music according to metre. The distinction is apparently known to both musicians and non-musicians; that is, most listeners, both sophisticated (i.e., musicians, music theorists and researchers) and unsophisticated (i.e., non-musicians), speak of musical structure as being dependent upon both 'phrase structure' and 'metrical structure'. For example, an informal survey was conducted at McMaster University in which graduate students and faculty members of the Psychology department were asked to report on the defining principles of music. Specifically, the participants were asked what they thought constituted 'music.' Responses were varied but most respondents identified two important structural components. One structure incorporates the underlying ideas of the composer. This component was commonly referred to as the 'theme' or 'idea' behind the composition. The other structure involves a 'rhythmic component' commonly referred to as the 'beat.' It was obvious through the course of discussion with the respondents that they were using the terms 'beat' or 'rhythm' to refer to what has been called 'metrical structure' in this thesis. The use of the terms 'themes' or 'ideas' closely matched the type of musical structure referred to as phrase structure in the theory of musical form (i.e., themes, motives, sections, phrases).

The distinction between metrical and phrase structure is not new. Other researchers have made similar note of this distinction. For example, Lee (1985) stated that one of the most obvious facts about the experience of listening to music is that one perceives a sequence of note durations according to some sort of temporal structure. He stated that there seem to be two ways in which notes may be temporally organized: they may be grouped into what musicians call phrases, and they may also be grouped into metrical units. Lee distinguishes between the two by stating that they obey different restrictions. Metrical units are normally all of the same time-span, whereas the same is not true for phrases (Lee, 1985). While there is little doubt that this distinction is sometimes true, it is certainly not always the case. Indeed, many theorists argue that phrase and metrical structure usually reinforce one another (Palmer & Krumhansl, 1990).

Lerdahl and Jackendoff (1983) also identified a theoretical distinction between metre and phrase structure, except they referred to phrase structure in terms of 'groupings'. For example, they explicitly state:

> The first rhythmic distinction that must be made is between grouping and meter. When hearing a piece, the listener naturally organizes the sound signals into units such as motives, themes, phrases, periods, theme-groups, sections, and the piece itself. Performers try to breathe (or phrase) between rather than within units. Our generic term for these units is group. At the same time, the listener instinctively infers a regular pattern of strong and weak beats to which he relates the actual musical sounds. The conductor waves his baton and the listener taps his foot at a particular level of beats. Generalizing conventional usage, our term for these patterns of beats is meter. (p.12)

While most people will agree that phrase structure and metrical structure are two <u>theoretically</u> distinct structural components of music, it is questionable whether or not the two structures are perceptually independent. Postulating a theoretical distinction between metrical and phrase structure in no way allows us to draw conclusions regarding the basis upon which they are perceived. The relationship between metrical and phrase structure is a matter of considerable debate among theorists. Most theorists agree that phrase structure is based upon phenomenal accents in the musical sequence. The difficulty lies in determining the nature of metrical structure. Music theorists disagree as to whether metre is completely defined by the phenomenal accents perceived in music or primarily by an inferred mental structure (Palmer & Krumhansl, 1990).

Direct perception of metre. Proponents of the position that metrical structure is perceived directly from the music argue that the same accent pattern determines metrical and phrase structure (Cooper & Meyer, 1960; Palmer & Krumhansl, 1990). For example, Berry (1976) and Cooper and Meyer (1960) defined metre on the basis of phenomenal accents that occur in the music itself. Support for this position comes from two lines of evidence. The first line of evidence comes from research on the performance of musical compositions which has shown that performers accent metrically important beats in a number of ways. They may slightly elongate the duration of notes corresponding to strong beats and shorten the duration of notes that fall on weak beats (Clarke, 1985; Gabrielsson, 1974, 1987). For example, Clarke (1985) investigated piano performance deviations and found that notes in strong metrical positions tended to be lengthened and notes in weak metrical positions were shortened. A performer may also use melodic features, such as a subtle change in dynamics with notes corresponding to stronger beats being a little louder and those corresponding to the weaker beats being a little softer (Schachter, 1980; Sloboda, 1983; Yeston, 1976).

The second line of evidence comes from statistical analyses of different musical styles and the relationship of event regularities to metre. For example, Palmer and Krumhansl (1990) investigated the hypothesis that composers may reinforce the perception of metre by manipulating the frequency of events at each metrical level. To test this hypothesis, they tabulated the frequencies of note onset for each of 16 equal subdivisions within a measure for 2/4 and 4/4 metres, and 12 equal subdivisions in 3/4 and 6/8 metres. In addition they used different compositional styles by Bach, Mozart, Brahms and Shostakovich. They argued that if frequency of event occurrence is an important cue to metre, then frequency distributions of musical events should differentiate one metre from another, both within and across musical periods. They found that differences in note onset events across temporal locations corresponded to the predictions from music-theoretic notions of metrical accent strength. The correlations between predictions of metrical accent strength for each metrical structure and the frequency of note onset events were significant (2/4, r=.98; 3/4, r=.92; 4/4, r=.96; and 6/8, r=.87). Subjects may use this information to help establish the metre of a musical presentation.

There are difficulties associated with both lines of evidence. In terms of the performance research, one could reasonably argue that subtle variations of notes corresponding to metrically significant locations detract from accurate perception of a metrical structure. That is, insofar as a performance fails to produce the precise time ratios upon which a metrical structure is built, then perceived metre may differ from the scored metre (Jones, 1990). Arguments that subtle performance cues mark important metrical locations would be strengthened if there was evidence that performers consistently employed them, but the evidence suggests the contrary (Jones, 1990; Sloboda, 1983; Vos & Handel, 1987). Many of the cues found to occasionally mark metrically important locations are also used to mark phrase boundaries. Palmer and Krumhansl (1990) acknowledged this difficulty when they concluded that it seems unlikely sensory cues alone determine metre.

One difficulty with the statistical analysis approach is that several measures must pass before subjects could unambiguously perceive a metre if statistical frequency data were being used. While it is possible that this occurs, intuitions and personal experiences with music suggest that perception of a metrical structure is much faster. In addition one must keep in mind the possibility that the statistical frequency data are based on coincidence of phrase and metrical structure, which often support one another. Phrase boundaries frequently occur at strong metrical locations (i.e., the first beat of a measure). This may reflect a sufficient, but not necessary condition for the perception of a particular metrical structure.

Inference of metre. Some theorists believe that metre is independent of phrase structure and is an inferred mental structure (Benjamin, 1984; Lerdahl & Jackendoff, 1983). Evidence in favor of this point of view is the fact that exact time-ratios are rarely, if ever, produced by performers (Jones, 1990; Gabrielsson, 1986, 1987; Palmer, 1989; Palmer & Krumhansl, 1990). Therefore it seems unlikely that metre is directly communicated.

The basis for inference of metrical structure is rarely agreed upon. Lerdahl and Jackendoff (1983) believed that metrical accents arise from the existence of a periodic mental framework or schema. To establish a particular metrical structure, a listener must infer the metrical accents from the location of phenomenal and structural accents in the music. However, metrical accents themselves are defined as qualitatively distinct types of accents in Lerdahl and Jackendoff's thesis. Benjamin (1984) similarly claimed that metre's role is to give time-points an identity independent of the tonal, motivic, or harmonic accents present.

The difficulty with assuming that metre is an inferred structure becomes clear when one notes that although most musical sequences are perceived as metrically unambiguous, all sequences have, in principle, an indefinite number of possible metrical interpretations (Lee, 1985; Longuet-Higgins & Lee, 1984). In addition, it is not clear why there is unambiguous selection of a metrical structure from two closely structured alternatives (e.g., 2/4 and 4/4 time). The pivotal question is how does a listener determine the correct metrical impression of a musical sequence if it is not presented unambiguously via the use of phenomenal cues?

Several researchers and theorists (Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1982, 1984; Longuet-Higgins & Steedman, 1971; Steedman, 1977) have suggested various rules that a listener might use to infer metre. In spite of differences among these accounts, it is possible to discern a common underlying assumption. They all agree that the temporal phenomenal accent of a relatively long-duration note serves as a cue for the perception of metrical structure. Long-duration notes tend to be heard as initiating major metrical units. As well, music-theoretical views point to the role of harmonic- and composite-rhythms, although to date there has been very little empirical research conducted to verify the psychological importance of these factors.

There are valid arguments for both the position that metre is defined by the phenomenal accents perceived in music and the position that it is an inferred mental structure. Despite the apparent differences, there are several points on which both views concur. For example, both accounts agree that the perception of metrical structure involves two processes: one is the perception of an iterative pulse or beat, and the other is the structuring of these pulses into a hierarchy of time-spans. Furthermore, there is little argument that the perception of iterative pulses or beats is dependent upon the temporal aspects of the musical presentation (i.e., the speed of note onsets and offsets). Tempo and note durations play an important role in this regard. The difference between the two positions lies in the basis upon which the structuring of the perceived pulses is believed to occur.

Regardless of the stance one takes in this debate, there is general agreement that regular time-spans based on metrical structure combine with durations based on the phenomenal events which define phrases and sections to form a perceived rhythmic structure. In order to hear a series of notes as music rather than as a succession of individual elements or units, the notes have to be linked together to form a group. That is, music by definition requires structure. This 'grouping' may be based upon phenomenal events which occur in the musical flow resulting in a perceived metrical and phrase structure, or it may be based upon a combination of these perceived durations with an abstract mental framework (i.e., an inferred metrical structure). This is, in many regards, the way Jones (1990) has thought of musical structure. According to Jones (1990) the occurrence of events in the music-flow create expectations via the establishment of temporal invariants which are abstracted. These temporal invariants result in a metric to which the occurrence of events are judged. Thus, it is the combination of expectations being fulfilled and unfulfilled (i.e., the combination of the phrase and metre) that results in the perceived structure of the music.

Summary and General Purpose

The perception of rhythmic structure is a complex phenomenon. As with any form of structure perception, the physical salience of a musical sequence (i.e., the occurrence of phenomenal accents) will be mediated by cognitive factors (i.e., structural accent strength). The occurrence of phenomenal and structural accents serve as events which determine a complex pattern of successive and simultaneous time-spans. Furthermore, some theorists believe that some of the time-spans which contribute to a perceived rhythm are themselves based on an inferred structure (i.e., metre) and may therefore not correspond to a direct phenomenal event in the musical sequence.

The review of the literature presented in this chapter should have highlighted for the reader some of the current debates and difficulties that exist in the literature on rhythmic structure. Of note, there is considerable debate regarding the nature of phenomenal and structural accents in the perception of phrase structure (Jones, 1987; Palmer & Krumhansl, 1987a, 1987b). Some theorists (e.g., Jones, 1987) believe that accent strength (i.e., perceptual salience) is dependent on mere coincidence of melodic, harmonic, and temporal features. Others (e.g., Palmer & Krumhansl, 1987a, 1987b) acknowledge subtle differences in the perceptual salience of various melodic or temporal events, but insist that the two theoretical component categories combine in an additive fashion to determine the perceived structure.

There is also the debate regarding the nature of metrical structure. Some theorists (e.g., Benjamin, 1984; Lerdahl & Jackendoff, 1983) believe that metre is predominantly an inferred mental structure while others (e.g., Cooper & Meyer, 1960) believe it is, like phrase structure, dependent upon the occurrence of direct phenomenal accents in the music. The relationship between phrase structure and metrical structure will, therefore, depend upon the position one takes in this debate.

This thesis deals with many of the issues which have been raised in this introductory chapter. Together, the series of experiments which shall be reported in the chapters that follow address a number of specific questions in an attempt to shed light on these areas of debate. First, are the theoretically distinct categories of phenomenal accents equally salient in the determination of phrase boundaries or can they be ordered according to a hierarchy of cognitive or perceptual significance in the same way that tones and chords are ordered in terms of structural accent strength? Second, are harmonic phenomenal accents (i.e., chord changes) perceptually more salient than temporal phenomenal accents (e.g., a relatively long-held note) in the perception of phrase structure? Third, if it is found that harmonic and temporal phenomenal accents are not perceptually equivalent, is the relative salience of a harmonic phenomenal accent mediated by the structural accent strength of the chords which are used? Fourth, is metre an inferred structure or is it defined primarily on the basis of phenomenal accents? Fifth, if metre is an inferred structure how does it relate to phrase structure in the determination of rhythm perception? Finally, if metre is an inferred structure, upon which phenomenal events is it primarily based: melodic, harmonic or temporal?

Three series of experiments are reported in this thesis. Each series deals with the problems which were outlined in the review of the literature. In the first

Experiment (Chapter 2) a harmonic and a temporal accent were pitted against each other in such a way as to form different rhythm patterns. As well, two harmonic conditions which varied in the frequency of chord presentations (i.e., the composite-rhythm) but not the frequency of chord changes (i.e., the harmonic-rhythm) were presented. Musicians and nonmusicians were requested to report perceived rhythm patterns in an attempt to determine the relative salience of the harmonic and temporal accents. In addition, a behavioral measure of the perceived metre was taken. Results indicated that the location of chord changes was the main determinant of subjects' rhythmic perceptions, although nonmusicians additionally showed evidence of selecting rhythms on the basis of the temporal accent. For all subjects, the onset of a measure was typically inferred to coincide with a chord change. As well, although subjects primarily inferred different metres based on the composite-rhythm, an interaction of metrical and rhythmic choices was found. It was concluded that musical passages with frequent chord presentations are heard with a faster beat, and thus a different metre, even when the harmonic-rhythm remains constant, and that perception of rhythm patterns and inference of metrical structure are not always independent.

In the second series of experiments (Chapter 3) the contribution of harmonic-temporal features and the role of musical expertise for the perception of rhythm patterns in music was more closely investigated. Three experiments were conducted in which a harmonic and a temporal accent were pitted against each other in such a way as to form 5 possible rhythm patterns. In Experiment 1, the composite-rhythm was controlled by presenting common diatonic chord progressions, with each chord being presented six times, each time for a short duration, before changing. In

Experiment 2, the same chord progressions which were employed in the first experiment were used, but the composite-rhythm was varied by changing the number of times chords were presented while maintaining the overall duration between chord changes. In Experiment 3, the composite-rhythm was maintained, but the type of chord progressions employed were varied. In all experiments, musicians and nonmusicians were requested to report perceived rhythm patterns in an attempt to determine the relative salience of the various accents. Results indicated that changes in the composite-rhythm led to a predictable change in an inferred metrical structure for both musicians and nonmusicians. All common diatonic chord progressions were found to lead to similar patterns of responses in which coincidences of harmonic, temporal, and inferred metrical accents were perceptually salient events. When a nondiatonic chord progression was employed however, there was neither evidence of an inferred metre, nor of responses on the basis of accent coincidence. Overall, musicians were found to primarily report rhythm patterns defined by harmonic accents, while nonmusicians reported rhythm patterns defined by an inferred metrical structure.

The final series of experiments (Chapter 4) investigated the basis upon which an inferred metrical structure is made. The relative roles of temporal, harmonic and melodic phenomenal accents in this inferential process were also investigated. The three series together give an indication of how phrase and metre combine to determine perceived rhythmic structure.

CHAPTER II

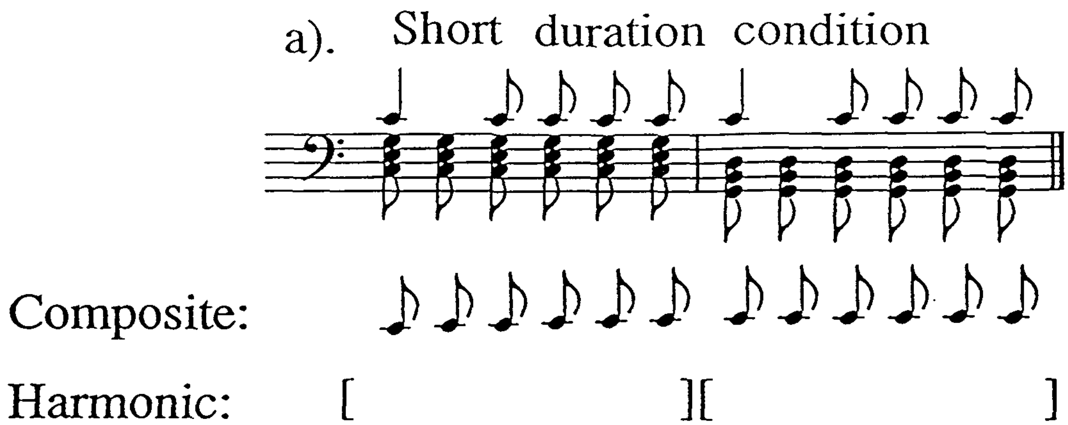
Series 1: The role of harmonic accents in inference of metrical structure and perception of rhythm patterns.

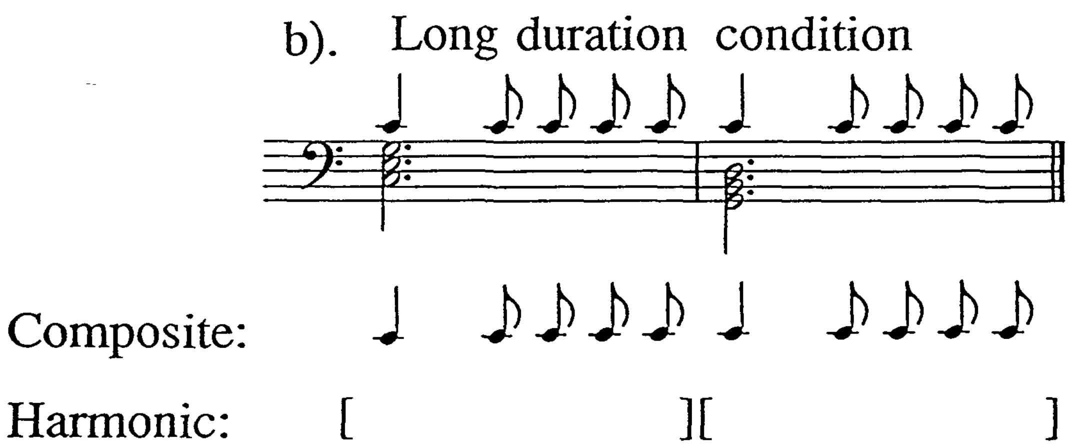
Introduction

Many of the issues discussed in Chapter 1 were addressed in the present study. In respect to the phrase structure literature, the relative accent weight of temporal and harmonic accents, and the possible role of accent coincidence in defining perceptually salient events in the music were investigated. In order to explore these issues, a harmonic accent (i.e., a chord change) was systematically pitted against a temporal accent. The temporal accent was created by playing a relatively long duration note (L) followed by four short duration notes (S). Chord change locations were systematically varied relative to the location of the temporal accent resulting in 5 possible harmonically defined rhythm patterns (LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL). Subjects were requested to report the perceived rhythms. Based on subjects' responses, the relative phenomenal accent weight for the temporal and harmonic accents in the phrasing of music could be determined. It was assumed that if the two phenomenal events differ in their perceptual salience, the more prominent event would define the phrase boundary and thus the perceived rhythm pattern. A temporal accent rather than a melodic or timbre-based accent was selected to be pitted against the harmonic accent because of its known role as a boundary cue. With metre, long duration notes tend to be heard as initiating major metrical units of time (Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1982, 1984; Longuet-Higgins & Steedman, 1971; Steedman, 1977). In contrast, long durations tend to mark the end of phrases in music (Boltz, 1989; Dowling, 1973; Garner, 1974; Royer & Garner, 1966, 1970; Vos, 1977; Woodrow, 1909), and speech (Cooper, Paccia, & LaPointe, 1978; Lehiste, 1975; Oller, 1973; Scott, 1982).

The relationship between metre and phrase structure was also of interest. Specifically, it was necessary to determine: (a) if subjects would consistently and unambiguously infer a metrical structure in the absence of performance cues in the microstructure of the music; and (b) upon what basis such an inference would be made. Of particular interest was the relative contribution of composite- and harmonic-rhythms to the inference of a metrical structure. In order to address this issue, two harmonic conditions which varied in the frequency of chord presentations (i.e., the composite-rhythm) but not the frequency of chord changes (i.e., the harmonic-rhythm) were presented (see Figure 9). In Figure 9, the composite-rhythm patterns are immediately below the staves. The harmonic-rhythm is represented by the grouping below the composite-rhythm. While only one example is given in each of the long- and short-duration harmonic conditions, the experiment reported in this chapter employed five stimuli which differed in the relationship between the harmonic-rhythm and the melodic line (see Table 1). The interesting question that arises is, in what metre will subjects hear these presentations? With a quarter note followed by 4 eighth notes, a multitude of metres is theoretically possible, but two metres are probable: 3/4 and 6/8.

Figure 9. The Composite- and Harmonic-Rhythm patterns of stimuli.


a). Short duration condition
Composite:
Harmonic:
[ ][ ]


b). Long duration condition
Composite:
Harmonic:
[ ][ ]

If a particular metre was inferred it could also bias subjects to perceive and report rhythm patterns consistent with that metre. Figure 10 illustrates that the onset of some rhythm patterns, as defined by the location of a chord change, would coincide with accented locations in a 3/4 and 6/8 metrical structure if one assumes that major metrical time-spans were inferred to begin with the onset of the temporal accent. The Figure also shows an important constraint of metrical structure in traditional Western music, as well as the relationship between 3/4 and 6/8 time. In traditional Western music, metre involves a hierarchy of simultaneous time-spans that either double or triple from one lower level to the next higher one. This is evident in the time-span reduction (after Lerdahl & Jackendoff, 1983) in which dots represent beat events in time.

Movement from the lowest metrical level to the intermediate level can involve the doubling of the duration between beats, which is followed by the tripling of it as one moves from the intermediate to the next highest metrical level. In this case one would hear a metre of 3/4 time in which there is a strong accent (indicated in this figure by the number of dots below a specific location) associated with the onset of the first note (see Example A, Figure 10). Coincidence of the chord change with this metrical accent could bias subjects to report the pattern LSSSS. Two slightly weaker metrical accents would fall on the first and third eighth notes. Chord changes coinciding with these salient metrical locations could bias subjects to report SSSSL (Example B) and SSLSS (Example C), respectively. It should be noted that these accents correspond to each notated beat in 3/4 time (i.e., 3 beats per measure, with a quarter note receiving a beat).

Figure 10. Coincidence of chord-change based rhythm patterns and metrical accents when major metrical time-spans are inferred to begin with the onset of the relatively long duration note. In Examples A, B, C, the inferred metrical structure is 3/4 time. In Examples D and E, the metre is 6/8 time.

Example A: Chord-change based rhythm of: L S S S S

Chord Changes:	↓						↓						↓
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩ *level*	•		•		•		•		•		•		•
𝅗𝅥. *level*	•						•						•

Example B: Chord-change based rhythm of: S S S S L

Chord Changes:			↓						↓				
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩ *level*	•		•		•		•		•		•		•
𝅗𝅥. *level*	•						•						•

Example C: Chord-change based rhythm of: S S L S S

Chord Changes:					↓						↓		
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩ *level*	•		•		•		•		•		•		•
𝅗𝅥. *level*	•						•						•

Example D: Chord-change based rhythm of: L S S S S

Chord Changes:	↓						↓						↓
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩. *level*	•			•			•			•			•
𝅗𝅥. *level*	•						•						•

Example E: Chord-change based rhythm of: S S S L S

Chord Changes:				↓						↓			
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩. *level*	•			•			•			•			•
𝅗𝅥. *level*	•						•						•

Alternatively, movement from the lowest level to the intermediate level can involve the tripling of the time-span followed by the doubling of it (see Example D, Figure 10). In this case one would hear a metre of 6/8 time in which there is a strong accent associated with the onset of the quarter note and a slightly weaker accent falling on the second eighth note. Chord-change based rhythm patterns of LSSSS and SSSLS (Example E) would coincide with these metrical accents. It is important to note that these two metrical accents do not correspond to each notated beat in 6/8 time (i.e., in 6/8 time there are 6 beats per measure, with the eighth note receiving one beat). This is because the tactus of a musical sequence invariably corresponds to an intermediate level of the metrical hierarchy in which the beats pass by at a moderate rate. Due to the fast rate at which the notated beat passes by in a 6/8 time, subjects will more often than not adopt as the tactus the next highest metrical level where the beats proceed at a slower rate (i.e., only 2 beats per measure; Lerdahl & Jackendoff, 1983).

As stated above, these predictions are based on the assumption that the measure would be perceived to begin with the onset of the temporal accent. However, one of the reasons for this investigation was to establish whether harmonic and temporal phenomenal accents differ in perceptual salience. This assumption may not be valid. Subjects may perceive the onset of the measure to coincide with the chord change instead. Nevertheless, the inferred metre could still bias subjects to perceive and report rhythm patterns consistent with that metre. That is, subjects could be biased to report rhythm patterns in which the temporal accent, rather than the chord change, coincides with a metrical accent.

Figure 11 shows that if this were to occur, the same rhythm patterns found to coincide with the locations of metrical accents in Figure 10, would be found to coincide with metrical accents, but for a different reason. That is, a bias may be found for the rhythm patterns of LSSSS (Example A, Figure 11), SSSSL (Example B), and SSLSS (Example C) with a 3/4 metre because of the coincidence of the temporal and metrical accents. Similarly, a bias for the rhythm patterns LSSSS (Example D), and SSSLS (Example E) may be found with a 6/8 metre.

Finally, it would be of interest to determine if musical training would result in a difference in the inferred meter and/or a difference in the weighting assigned to the phenomenal accents under investigation. Musical experience, and in particular, formal musical instruction may change one's explicit knowledge and vocabulary of music. Such a knowledge-base could alter the way a subject perceives the musical structure through conceptually-driven processing of incoming information. Thus, the participants in this experiment were divided into two groups on the basis of whether or not they had ever received formal musical training. The results from the two groups were compared to determine the effects of musical instruction on rhythm perception.

Summary

This experiment focussed on the role of two functionally distinct types of harmonic phenomenal accents; those resulting from chord onset, and those resulting from chord change, in the perception of rhythmic structure. These accents are

Figure 11. Coincidence of a temporal and metrical accents when major metrical time-spans are inferred to begin with the location of a chord change. In Examples A, B, and C, the inferred metrical structure is 3/4 time. In Examples D and E, the metre is 6/8 time.

Example A: Chord-change based rhythm of: L S S S S

Chord Changes:	↓						↓						↓
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩ *level*	•		•		•		•		•		•		•
𝅗𝅥. *level*	•						•						•

Example B: Chord-change based rhythm of: S S S S L

Chord Changes:			↓						↓				
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩ *level*	•		•		•		•		•		•		•
𝅗𝅥. *level*			•						•				

Example C: Chord-change based rhythm of: S S L S S

Chord Changes:					↓						↓		
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩ *level*	•		•		•		•		•		•		•
𝅗𝅥. *level*					•						•		

Example D: Chord-change based rhythm of: L S S S S

Chord Changes:	↓						↓						↓
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩. *level*	•			•			•			•			•
𝅗𝅥. *level*	•						•						•

Example E: Chord-change based rhythm of: S S S L S

Chord Changes:				↓						↓			
Temporal Line:	L		S	S	S	S	L		S	S	S	S	L
Metrical Grid:													
♪ *level*	•	•	•	•	•	•	•	•	•	•	•	•	•
♩. *level*	•			•			•			•			•
𝅗𝅥. *level*				•						•			

associated with the music-theoretical notions of composite- and harmonic-rhythm, respectively. With the emphasis on the phenomenal aspect of chord presentations, this study departed from traditional approaches to harmonic investigations that have focussed on the structural relationship among chord presentations within a tonal context. By systematically varying the location of a chord change relative to a long duration note, and asking subjects to report the perceived structure and metre, several current and relevant theoretical issues were addressed. Of note is the relative accent weighting of temporal and harmonic accents, as well as the role of accent coincidence in phrase structure investigations. The effects of formal music training on perceived accent weights were also investigated. Finally, the issue of whether or not subjects could unambiguously infer a metrical structure from the composite- and harmonic-rhythm in the absence of performance cues directly addressed the relationship of metre to phrase structure.

Method

Subjects

Twenty-four experimentally naive volunteers participated in this study. They were tentatively divided into two groups on the basis of musical training. There were 12 musicians with a range of 4-11 years and a mean of 6.7 years of formal musical training. Their age ranged from 18-22 with a mean of 20.1 years and there were 5 males, 7 females. Twelve subjects were nonmusicians with an age range of 19-21 and a mean of 19 years. Four of these subjects were male and 8 were female. Subjects were recruited from introductory psychology classes, and were given course credit for participating in the experiment.

Apparatus

The stimuli were generated by a Yamaha CX5M music computer and presented via a NAD 3020e integrated amplifier to two speakers placed approximately 2-3 feet on either side of a video monitor. Each subject sat at a table in the laboratory and responded to the stimuli by moving a joystick that communicated the subject's choices directly to the computer. A Hitachi 1481 Color Monitor was situated at eye-level directly in front of the subject. By moving the joystick, the subject controlled the movements of the cursor visible on the monitor screen. The cursor had to be moved into one of five appropriately labelled boxes to indicate the response choice for each trial. Parameters and data for each trial were automatically recorded and stored by the computer on disk. A table and chair were also set up for the experimenter who sat directly behind the subject.

Stimuli

The stimuli for this experiment consisted of a temporal presentation and a corresponding harmonic component. On any given trial one basic rhythm pattern of a long duration tone (L) followed by four short tones (S) was presented ten times (LSSSSLSSSSLSSSS...). In musical notation, the long-held note was a quarter note (i.e., it received one beat in common time) and the short-duration notes were equal to eighth notes (i.e., they received one half beat in common time). In real time, the long duration notes (L) were 444 msecs and the short duration notes were 222 msecs. There were no pauses between notes. The fifty tones presented for each trial were all Middle C. The reason for presenting all the tones at the same frequency was to eliminate melodic phenomenal (e.g., contour changes, large interval jumps) and structural accents. Different harmonic presentations accompanied the temporal line. Chords were presented in such a way as to break up the rhythm pattern in relation to the long-duration tone, resulting in one of five possible rhythm patterns (i.e., LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL). The harmonic accompaniment in all conditions consisted of a diatonic cadence (I-V-I in the key of C Major).

Two types of harmonic accompaniments were presented: a long- and a short-duration condition (see Table 1). In the long-duration harmonic condition, each chord was presented for a duration of 3 beats before changing (i.e., 1332 msecs). In the short-duration harmonic condition, each chord was presented six times before changing (222 msecs each), with each presentation lasting the same duration as a short-duration note in the temporal line. The only exception to this general description of the harmonic conditions was at the beginning and end of trials, where the duration between trial onset and chord change (beginning), and chord change and trial termination (end) was systematically varied to create the five rhythm patterns (see Table 1). It is important to note that the harmonic-rhythm patterns are the same in both harmonic conditions while the composite-rhythm patterns differ slightly (see Figure 9). The stimuli were presented through speakers at approximately 72 dB (SPL) at a tempo of 135 beats per minute, and the timbre was that of an FM-synthesized piano. The temporal line was presented at a slightly greater intensity (approximately 4 dB difference) than the harmonic accompaniment to facilitate perception of the temporal accent.

Table 1

A Schematic Representation of Stimuli

```
Temporal line:  L   S S S S L   S S S S L   S S S S L  ...

Long Condition
 Harmony #1:    C-----------G-----------C-----------G ...
 Harmony #2:    C---------G-----------C-----------G-- ...
 Harmony #3:    C-------G-----------C-----------G---- ...
 Harmony #4:    C-----G-----------C-----------G------ ...
 Harmony #5:    C---G-----------C-----------G-------- ...
Short Condition
 Harmony #1:    C C C C C C G G G G G C C C C C C G G...
 Harmony #2:    C C C C C G G G G G G C C C C C C G G G...
 Harmony #3:    C C C C G G G G G G C C C C C C G G G G...
 Harmony #4:    C C C G G G G G G C C C C C C G G G G G...
 Harmony #5:    C C G G G G G G C C C C C C G G G G G G...
```

Note. This Table is read by following the location of a chord change upward to see where it intersects the temporal line. The five harmonies create five rhythm patterns on the basis of a chord change: LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL.

Procedure

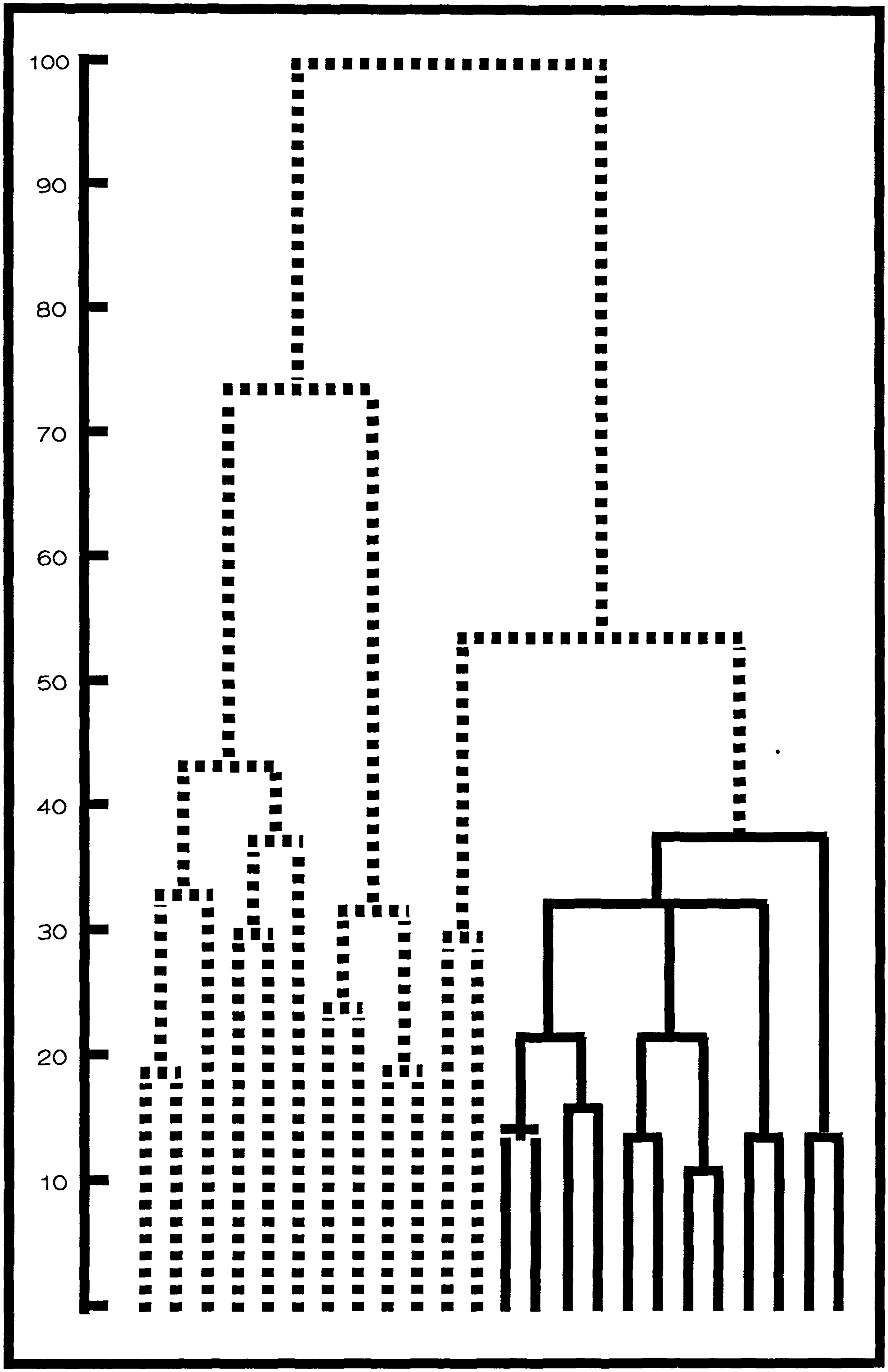
There were a total of 60 trials, 30 in each of the two harmonic conditions. Each harmonic condition was presented in a blocked format. Half of the subjects received the short-duration harmonic condition first while the other half received the long-duration harmony first. The 5 different stimuli based on location of the chord change were presented in random order 6 times. The subjects were to listen to the stimuli and then choose which rhythm pattern they had heard by placing the cursor in one of five appropriately labelled boxes (i.e., LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL). Subjects were under no time restraints and started each trial by placing the cursor into a start box. The experimenter ensured that each subject understood what each schemata represented by having them tap each of the five rhythms prior to the presentation of the stimuli. Subjects were also told that there were no correct answers and that it was possible that all trials would sound as if they had identical rhythm patterns. Subjects were further instructed to tap their foot during the presentation of the stimuli. An experimenter sat immediately behind the subject and noted how many times per iteration of the rhythmic pattern the subject tapped his/her foot. Subjects had no difficulty doing this task. Finally, subjects were required to fill out a standard questionnaire regarding their musical experiences and formal training.

Results and Discussion

The first result to be noted was that subjects invariably tapped their foot only two or three times per iteration. It was assumed a priori that if subjects tapped their foot two times per iteration, then they were operating in a 6/8 time metre. If subjects tapped their foot three times, it was assumed that they were operating in a 3/4 time metre. This behavioral consistency was believed to be due to the grouping of iterative pulses established by the harmonic-rhythm. While tapping of one's foot 4 or 5 times per iteration was theoretically possible, such behavior would not be consistent with the pattern set by the composite- and harmonic-rhythms.

The first analysis conducted was to establish objective support for our arbitrary musical experience criterion which resulted in subjects being placed into one of two subject groups. Response profiles consisting of percentage counts for the two dependent variables of rhythmic choice (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL), and metrical choice (3/4 and 6/8 time) were computed for each subject. These profiles were specified for the five rhythms defined by chord change locations (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL) in each of the two harmonic conditions (long- and short-duration), resulting in a 10 X 10 matrix. The 24 matrices (one for each subject) were then analyzed by means of the hierarchical clustering module in the software package "Statistica." The solution, which is shown in Figure 12, supported our original subjective classification. From the figure one can see that our original group of musically experienced subjects, here marked by a solid line, forms a relatively tight cluster on the right side of the solution. Similarly, the nonmusicians, here represented by the dotted line, form a cluster on the left side of the solution. The higher level connections among the nonmusicians relative to that which is seen in the musician's cluster is indicative of greater variability in the pattern of responses for nonmusicians compared to musicians. As well, the pattern of responses for two of the nonmusicians were found to be more similar to that which was found for musicians than their fellow

Figure 12. Hierarchical clustering (Ward's Method) of the 24 subjects based on a comparison of individual response profiles. Solid lines represent individuals with formal musical training. Broken lines represent nonmusicians. The ordinate represents a proportional measure of dissimilarity relative to the maximum Euclidean distance.



100
90
80
70
60
50
40
30
20
10

nonmusicians. This is seen in Figure 12 by noting that the two nonmusicians' cases closest to the musicians' cluster join with it, rather than with the nonmusicians' cluster. An investigation of the musical questionnaires completed by these two subjects indicated that they were the only "nonmusicians" with some notable musical experiences. They both had some experience with guitars, although they did not play often, nor had they any formal musical training.

Because the measures used in this study were discrete and categorical, a loglinear analysis was performed on the data, with four design and two response variables. The design variables were the musical Training of the subject (inexperienced or experienced), the Location of the chord change relative to the long duration note (resulting in the rhythm patterns of LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL), the Duration of the harmonic condition (short or long), and the Order of presentation for the two durations. The response variables were the subject's Choice of rhythm (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL), and the Metre in which they were operating (3/4 or 6/8) based on their foot tapping.

The loglinear analysis was implemented with the loglinear module in Statistica. An appropriate model was determined by first conducting tests of partial and marginal association. These tests examine all possible effects and interactions. The test of partial association involved the comparison of a model that included all equivalent and lower-order interactions including the interaction under consideration with a model of the equivalent and lower-order interactions except the one under consideration. If the difference between the two models was statistically significant, the effect was retained; if not, it was excluded. The test of marginal association involved the comparison of a model without any interactions with a model that included the interaction under consideration. Again, if the difference between the two models was statistically significant, it was retained. The results of these tests for association indicated the significant effects were: Choice (Partial $\chi^2(4) = 13.19$, $p < .01$; Marginal $\chi^2(4) = 13.19$, $p < .01$), Choice X Training (Partial $\chi^2(4) = 23.82$, $p < .001$; Marginal $\chi^2(4) = 22.16$, $p < .001$), Choice X Location (Partial $\chi^2(16) = 462.16$, $p < .001$; Marginal $\chi^2(16) = 472.71$, $p < .001$), Choice X Duration (Partial $\chi^2(4) = 32.56$, $p < .001$; Marginal $\chi^2(4) = 13.68$, $p < .01$), Metre X Duration (Partial $\chi^2(1) = 856.81$, $p < .001$; Marginal $\chi^2(1) = 829.89$, $p < .001$), Choice X Metre (Partial $\chi^2(4) = 87.80$, $p < .001$; Marginal $\chi^2(4) = 77.90$, $p < .001$), Choice X Location X Training (Partial $\chi^2(16) = 106.23$, $p < .001$; Marginal $\chi^2(16) = 107.18$, $p < .001$), Choice X Duration X Training (Partial $\chi^2(4) = 15.10$, $p < .005$; Marginal $\chi^2(4) = 11.38$, $p < .05$), Metre X Duration X Training (Partial $\chi^2(1) = 9.71$, $p < .005$; Marginal $\chi^2(1) = 15.43$, $p < .001$), and finally, Choice X Location X Duration (Partial $\chi^2(16) = 40.30$, $p < .001$; Marginal $\chi^2(16) = 51.26$, $p < .001$).

The second stage of the analysis first fitted a model that contained only the significant effects from the first stage. The best fitting model was determined by systematically removing effects and seeing if a significant reduction in the fit occurred. If so, the effect was retained. All possible combinations of effects were systematically fitted by this backward elimination procedure until the best model was obtained. This model contains 4 three-way interactions which account for all of the observed variance: Choice X Location X Duration; Choice X Location X Training; Metre X Duration X Training; and Choice X Metre X Location; $\chi^2(296, N = 1440) =$

202.11, $p = 1.00$).

The significant interaction of rhythmic Choice with Location of a chord change and Training, presented in Table 2, shows that subjects predominantly chose rhythm patterns on the basis of the location of the chord change. With both musicians and nonmusicians, the pattern of results, with most rhythmic choices along the negative diagonal, indicates several things. First, when the location of the chord change coincided with the onset of the temporal accent (i.e., LSSSS), subjects predominantly chose that rhythm pattern (51.4% for nonmusicians, and 63.9% for musicians), supporting theories of musical structure based on accent coincidence (Benjamin, 1984; Berry, 1976; Jones, 1987). Second, when the rhythm pattern on the basis of the chord change was SLSSS, SSLSS, SSSLS and SSSSL, subjects chose these rhythms most often: 29.2%, 34.7%, 29.2%, and 22.2%, respectively for nonmusicians; and 50.%, 59.7%, 55.6%, and 62.5%, respectively for musicians. In addition, when subjects chose a particular rhythm pattern (e.g., LSSSS), they did so most often when the chord change supported that rhythm pattern. This means that subjects principally used the location of the chord change as a determinant for the perceived rhythm pattern. Finally, musicians and nonmusicians differ in two distinct ways. First, it is apparent from the greater percentage of responses along the negative diagonal for musicians compared to nonmusicians, that the harmonic accent associated with a chord change is perceptually more salient for musicians than for nonmusicians. Second, nonmusicians show a greater bias to select a rhythm pattern on the basis of the temporal accent alone, compared to musicians. This becomes apparent when one

Table 2

Effect of Rhythmic Choice by Location of the Chord Change and Training

	Rhythm on the basis of the chord change				
Rhythmic choice	LSSSS	SLSSS	SSLSS	SSSLS	SSSSL
			Nonmusicians		
LSSSS	51.4%	26.4%	18.1%	16.7%	27.8%
SLSSS	13.9%	29.2%	13.9%	19.4%	18.1%
SSLSS	13.9%	9.7%	34.7%	22.2%	16.7%
SSSLS	11.1%	19.4%	15.3%	29.2%	15.3%
SSSSL	9.7%	15.3%	18.1%	12.5%	22.2%
			Musicians		
LSSSS	63.9%	8.3%	11.1%	9.7%	6.9%
SLSSS	11.1%	50.0%	2.8%	12.5%	15.3%
SSLSS	8.3%	6.9%	59.7%	9.7%	4.2%
SSSLS	1.4%	13.9%	15.3%	55.6%	11.1%
SSSSL	15.3%	20.8%	11.1%	12.5%	62.5%

compares the percentage of responses for the rhythm pattern LSSSS when the chord change location supports the rhythm patterns SLSSS, SSLSS, SSSLS, and SSSSL, for musicians and nonmusicians. In all instances, responses are higher for nonmusicians than for musicians: 26.4% > 8.3% (SLSSS), 18.1% > 11.1% (SSLSS), 16.7% > 9.7% (SSSLS), and 27.8% > 6.9% (SSSSL), respectively. In fact, for nonmusicians, grouping on the basis of the temporal accent (LSSSS) was slightly more salient than was the rhythm defined by a chord change location when the chord supported rhythm was SSSSL (27.8% compared to 22.2%).

The significant interaction of metrical choice by the duration-condition of chord presentation and training indicated that subjects were for the most part operating in a 3/4 metre when they were presented with the long-duration harmonic accompaniment and a 6/8 metre when they were presented with a short-duration harmonic accompaniment. Furthermore, this pattern of metrical responses was more strongly evident in the data of the musicians compared to the nonmusicians. When the chord presentation was long, subjects more often than not tapped their feet 3 times per iteration (87.2% and 92.2% for nonmusicians and musicians, respectively) in support of a 3/4 metre. In contrast, when the chord presentation was short, subjects were observed to tap their feet more often in support of a 6/8 metre (81.7% and 93.3% for nonmusicians and musicians, respectively) than in support of a 3/4 metre.

Because the psychologically meaningful patterns of responses for the two three-way interactions of Choice X Location X Duration, and Choice X Metre X Location are nicely evident in the four-way interaction of Choice X Metre X Duration X Location, the results can be completely specified by examining the four-way interaction presented in Table 3. When the duration of the chord presentation was long, subjects predominantly perceived a 3/4 metrical structure (89.7% of the time). When the perceived metrical structure was 3/4 time under the long-duration harmonic condition, the pattern of rhythmic choices on the basis of the location of a chord change was as expected. Subjects chose rhythm patterns predominantly on the basis of a chord change with a bias for the rhythm patterns of LSSSS, SSLSS, and SSSSL (61.0%, 55.6%, and 45.7%, respectively). Occasionally (only 10.3% of the time) subjects showed evidence of operating in a 6/8 time when the chord presentation was of a long duration. Under such conditions, subjects only chose three different rhythm patterns. It is noteworthy that the rhythm patterns of SSLSS and SSSSL were never chosen, and that the pattern LSSSS was chosen only when it was supported by a chord change. Subjects predominantly chose the rhythm patterns of SLSSS and SSSLS under such conditions. Under these conditions, subjects chose the rhythm pattern of SLSSS every time the location of the chord change supported the rhythm pattern of SSSSL. This pattern makes sense, given that subjects were operating in a 6/8 time metre. In a 6/8 metre, 3 eighth notes are grouped at the intermediate level of the metrical grid (refer to Figures 10 & 11). If subjects used the location of the chord change as a metrical cue, then the two rhythm patterns supported by the metrical structure would be SSS-SL and SL-SSS when the chord-change based rhythm pattern was SSSSL. Subjects chose the latter of these two alternatives. These findings are interpreted as evidence that the perceived rhythm pattern biased subjects to infer a metre consistent with that rhythm.

Table 3

Interaction of Rhythmic and Metrical Choice by the Location of the Chord Change and Harmonic Condition

	Rhythm on the basis of the chord change				
Rhythmic choice	LSSSS	SLSSS	SSLSS	SSSLS	SSSSL
	Long duration harmonic condition				
3/4 Metre					
LSSSS	61.0%	13.9%	8.3%	8.3%	13.9%
SLSSS	4.2%	37.5%	8.3%	13.9%	19.5%
SSLSS	9.7%	5.6%	55.6%	19.5%	11.1%
SSSLS	2.8%	6.9%	9.7%	26.4%	5.6%
SSSSL	13.9%	20.8%	13.9%	12.5%	45.7%
6/8 Metre					
LSSSS	4.2%	--	--	--	--
SLSSS	1.4%	4.2%	1.4%	8.3%	4.2%
SSSLS	2.8%	11.1%	2.8%	11.1%	--

(table continues)

	Rhythm on the basis of the chord change				
Rhythmic choice	LSSSS	SLSSS	SSLSS	SSSLS	SSSSL
	Short duration harmonic condition				
3/4 Metre					
LSSSS	8.3%	5.6%	5.6%	2.8%	4.2%
SLSSS	--	5.6%	1.4%	--	--
SSLSS	6.9%	2.8%	4.2%	1.4%	2.8%
SSSSL	1.4%	2.8%	1.4%	2.8%	2.8%
6/8 Metre					
LSSSS	41.7%	15.3%	15.3%	15.3%	16.7%
SLSSS	19.5%	31.8%	5.6%	9.7%	9.7%
SSLSS	5.6%	8.3%	34.6%	11.1%	6.9%
SSSLS	6.9%	15.3%	18.0%	47.2%	20.8%
SSSSL	9.7%	12.5%	13.9%	9.7%	36.1%

Note. Dashes and omitted rhythms mean subjects did not select them under those specific conditions.

When the duration of the chord presentation was short, subjects predominantly perceived a 6/8 metrical structure (87.5% of the time). When the perceived metrical structure was 6/8 time and the duration of the harmonic condition was short, the pattern of rhythmic choices on the basis of the location of a chord change was as expected. Subjects chose rhythm patterns predominantly on the basis of a chord change. The rhythm patterns that coincided with the location of accents in the metrical structure (i.e., LSSSS, and SSSLS) were chosen on the basis of a chord change most often (41.7% and 47.2%, respectively). Occasionally (only 12.5% of the time) subjects showed evidence of operating in a 3/4 time when the chord presentation was of a short duration. Under such conditions, subjects only chose four different rhythm patterns. It is noteworthy that the rhythm pattern of SSSLS was never chosen under such conditions. Subjects predominantly chose the rhythm patterns LSSSS, SSLSS, and SSSSL under such conditions. Again, this is interpreted in terms of a perceived rhythm pattern biasing subjects to select a metrical structure.

These findings are of some importance, particularly for theories of metrical structure. The data clearly support the conclusion that metre need not be communicated through performance cues. That is, metrically important locations were not marked by phenomenal accents such as a slight elongation of a note's duration or a subtle change in dynamics (Clarke, 1985; Gabrielsson, 1974, 1987). Nevertheless, subjects inferred a metrical structure and then interpreted the musical sequence in terms of that structure. This raises a number of questions regarding the role of performance cues in the perception of metrical structure. This is a notable concern, given that the cues used by performers to communicate the metre are often ambiguous, interactive, or simply absent -- as they were in this experiment -- yet listeners still determined a metrical structure (Palmer & Krumhansl, 1990).

In order to fully address these questions further research is necessary. It is clear that subjects could have used different means to arrive at the inferred metre and perceived rhythm. Specifically, the stimuli contained a number of temporal invariances which subjects could have used. According to Jones' influential work on attentional rhythmicities in music perception, a listener can pick-up temporal invariances in the musical flow (i.e., regularly occurring accents) which can then be used to guide attention at future intervals of time. The abstraction of a time interval from unfolding events is a dynamic process based in part on both fulfillment and violation of temporal expectations (Jones, 1990). From this dynamic process emerges a metric in which events occurring in the presentation are compared to the reiteration of directed attentional resources. Thus the occurrence of accents in the musical flow are perceived relative to a temporal invariance.

In this study temporal invariances could have been established by the regular occurrence of the temporal accent and/or the harmonic accent of a chord change. The onset of the relatively long duration note occurred every 1332 msecs. If this accent was used to establish an internal rhythmicity of attention then the occurrence of harmonic accents would have been judged relative to the temporal accent. In other words, it may be how the duration of time established by the regular recurrence of the temporal accent is parsed into smaller time-spans by the occurrence of other events, such as chord onsets or chord changes, that is of psychological significance in establishing a metre or rhythm (refer to Figure 10). If the occurrence of a chord change was perceived to coincide with a regular metrical dividing point (Smith & Cuddy, 1989), the range of probable inferred metres is reduced. However, the regular occurrence of a chord change also created a temporal invariance equal to a duration of 1332 msecs. In contrast to the previous consideration, it may be how the time-spans created by the chord change are parsed by the occurrence of other events such as chord onsets and the temporal accent that is important (refer to Figure 11). If the occurrence of a temporal accent was perceived to coincide with a regular metrical dividing point, the range of probable metres is also reduced.

Which of these two interpretations is supported by the data? Subjects primarily used the location of the chord change to determine the perceived rhythm pattern, implying that this harmonic phenomenal accent was more salient than the temporal accent for determining the phrase structure. In terms of the metrical structure, the primary basis for determining an inferred metre was the two harmonic-accompaniment conditions which differed in their composite-rhythm patterns. It was based on the frequency of chord presentations within the temporal or harmonic time-span that subjects primarily inferred a 6/8 or a 3/4 metrical structure. In those few instances where selection of the metrical structure was not based on the composite-rhythm, it appeared to be based on the rhythm pattern as defined by the location of the chord changes, implying that the chord change was also more salient than the temporal accent for metre determination.

Finally, evidence in favor of the chord change serving as the more prominent accent comes from an observation made during the experimental procedure and reports made by subjects upon completion of the experimental session. In both harmonic conditions, the duration between chord changes was typically 1332 msecs. The only exception was at the beginning and end of trials, where the duration between trial onset and chord change was systematically varied to create the five rhythm patterns (refer to Table 1). The result of this manipulation was that some rhythm patterns as defined by the location of a chord change would not be completed at the end of a trial. Interestingly, the experimenter noted that with most trials, subjects continued to tap their feet to the completion of the rhythm pattern, had it continued. In addition, subjects reported that some trials appeared to end prematurely (i.e., before they were expected to end). Jones and her colleagues have reported that how a duration is temporally patterned can evoke expectations of when a stimulus should end (Jones & Boltz, 1989). Thus, it was concluded that the chord change was perceptually more salient than a temporal accent in both phrase and metrical structure determination.

There are several possible reasons for this finding. In this experiment only one chord progression was employed (I.e., the I-V-I progression). This was because of the interest in the timing of phenomenal accents such as chord changes in the determination of rhythmic structure, excluding structural factors. However, the structural features of a chord progression may affect the perceptual significance of the chord change location. Structural factors could be of particular concern in this experiment because of the dissonance created by the suspension of Middle C in the temporal line while the dominant chord is being played as a harmonic accompaniment. The contribution of structural features to rhythm pattern perception is more closely investigated in the experiments reported in Chapter 3.

Once a metrical structure was determined, the timing of metrical accents could serve to heighten attention at regular intervals to events in the musical presentation, in a manner similar to the way accent coincidences serve to create expectations at particular locations in the joint-accent structure of Jones' (1987, 1992) theory. This is evident in the data from the increased salience of phenomenal events at metrically accented locations. The coincidence of the temporal accent, chord change, and metrical accent resulted in a higher percentage of responses for a rhythm pattern consistent with this coincidence (e.g., LSSSS in both 6/8 and 3/4 time), compared to the rhythm pattern defined by the coincidence of a temporal accent and chord change alone (i.e., SSSSL in 6/8 time). Similar higher response percentages were found for metrical accent coincidences with (a) chord changes only, and (b) the temporal accent only; compared to instances when there was no inferred metrical accent.

Conclusions

The results of this experiment support a number of conclusions. The first conclusion to be made is that the assumption of equal salience of phenomenal accents in phrase structure determination was not supported. There was evidence that the theoretical component categories of harmonic and temporal features are not perceptually equivalent. When a temporal phenomenal accent was pitted against a harmonic phenomenal accent, the harmonic accent won out. The location of a harmonic accent was fundamental in establishing a phrase boundary, and thus the reported rhythm pattern. This supports a recent proposal put forth by Jones (1992) regarding the notion of accent weights.

The second conclusion is that the relative salience of different accents can be mediated by musical training. For the most part, both musicians and nonmusicians were qualitatively similar in their responses, except musicians showed far less variance in their tendency to employ a chord change location as a cue for the rhythm pattern, and in their inference of a metrical structure. Nonmusicians did show a slight tendency to rely upon and thus select rhythm patterns based on the temporal accent (i.e., the rhythm pattern LSSSS), however, the variability in responses for the nonmusicians make us cautious regarding a firm conclusion in this regard. Several explanations for the general findings regarding the mediating effects of musical training are possible and Chapter 3 discusses them more thoroughly.

A third conclusion is that the coincidence of temporal and harmonic accents resulted in a perceptually significant event in music. This supports those models of phrase structure that rely on the coincidence of phenomenal accents to serve as psychologically important events (e.g., Benjamin, 1984; Berry, 1976; and Jones, 1987). Overall, whenever both the relatively long-held note and the chord change supported the same rhythm pattern, subjects showed a higher percentage of choice for the rhythm pattern based on the accent coincidence (LSSSS) than was evident for rhythm patterns based on either the chord change or temporal accent alone.

A fourth conclusion is that in the absence of phenomenal cues in the microstructure of a musical presentation, subjects will infer a metrical structure unambiguously, supporting conclusions made by others regarding metre determination (Essens & Povel, 1985; Povel, 1981; Povel & Essens, 1985; and Povel & Okkerman, 1981). In this study, the metrical structure was not experimentally controlled. Nevertheless, subjects inferred a metrical structure and then interpreted the musical sequence in terms of that structure. Furthermore, the inferred metrical structure appeared to be based primarily on harmonic information. The harmonic-rhythm pattern was the same for the two experimental conditions, which probably set the stage for selection of one of two closely associated metrical structures: 3/4 and 6/8 time. Unambiguous selection of the metre however was determined primarily by the frequency of chord presentation within the time-spans defined by either the temporal accent, or the chord changes. This conclusion can be seen as an extension to Lerdahl and Jackendoff's (1983) observation that a piece of music with frequent harmonic changes is heard with a faster tactus than a piece with equal note values (i.e., the same composite-rhythm) but less frequent harmonic changes. That is, we conclude that music with frequent chord presentations is also heard with a faster beat than a piece with the same pattern of harmonic changes (i.e., harmonic-rhythm) and fewer chord presentations. This perception of a faster beat can lead to the adoption of a slower tactus and thus the perception of a different metrical structure (e.g., a 3/4 and a 6/8 time metre).

A fifth conclusion is that once the metre was determined, inferred metrical accents resulted in a perceptually more salient location in the music. That is, the data are interpreted in terms of inferred metrical accents serving to heighten attention to events at those locations. In this way there is agreement with models such as Jones (1987; 1992), which argue that the occurrence and/or coincidence of accents may serve to heighten expectations and attention for specific locations. A chord change, the onset of the long duration note, and the offset of the long duration note, all appeared more salient when they coincided with a metrical accent than when they did not.

Finally, perception of rhythm patterns and a metrical structure were not always independent. Subjects had a tendency to select rhythm patterns on the basis of harmonic information (i.e., the location of a chord change) and showed a bias for those patterns consistent with the inferred metre. However, in those cases where the inferred metre was not consistent with the composite-rhythm, it was consistent with rhythm patterns defined by the location of a chord change. Put another way, subjects sometimes inferred a metre on the basis of the rhythm pattern that occurred within the time-span defined by a chord change rather than the composite-rhythm defined by that chord change. This suggests that subjects principally used information within the time-span defined by the chord change rather than information within the time-spans defined by the temporal accent to infer the metrical structure.

CHAPTER III

Series 2: Rhythm perception and differences in accent weights for musicians and nonmusicians.

The series of experiments reported in this chapter serves to replicate and thus establish the reliability of the results from the previous experiment. As in Experiment 1, the three studies reported in this chapter focus on the timing of chord presentations and changes as cues for the perception of a phrase boundary, and the interaction of this process with the inference of a metrical structure in the reporting of rhythm patterns. As well, mediation of the perceptual salience of a chord change by the structural accenting of a progression and the listener's musical background was also investigated. Several areas of concern and debate in the literature on rhythm-pattern perception were addressed.

The first concern deals with phrase structure investigations, and in particular the differential weighting of accents in the perception of a rhythm pattern. Work on phrase structure in the psychology of music has been concerned with the principle of reiteration, since it is through repetition and contrast that a composer is able to communicate his musical idea or theme (Bray, Snell & Peters, 1967). The aesthetic success of a musical composition is determined in part by the judicious mixture of these two principles and is accomplished by repeatedly presenting the same idea or theme with subtle variations of the timbre-based, melodic, temporal, and harmonic features with each repetition.

The presentation and transformation of themes has been a prominent notion in psychology and the study of musical structure (Jones, 1981a, 1981b, 1982, 1987; Welker, 1982). Of note is the theory of dynamic-pattern structure which has been developed by Jones and her colleagues (Jones, 1981a, 1981b, 1982, 1984, 1987; Jones, Boltz & Kidd, 1982; Jones, Kidd & Wetzel, 1981) that shows how various phenomenal accents may combine to determine not only phrase boundaries, but also the structure of the music within phrases and thus, the structure of a theme.

Another purpose of this investigation was an attempt to delineate stable and reliable differences in the perceptual salience of two theoretically distinct categories of phenomenal accents. Specifically, confirmation of the pattern of results reported in Chapter 1 was sought in terms of the perceptual salience of harmonic phenomenal accents (i.e., chord changes) compared to temporal phenomenal accents (e.g., a relatively long-held note). In order to investigate this question, the same basic paradigm employed in the first Experiment was used here. A harmonic phenomenal accent was systematically pitted against a temporal accent (i.e., a relatively long duration note -- L) resulting in 5 possible rhythm patterns (LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL). Subjects were requested to report the perceived rhythm. It was assumed that if the two phenomenal accents differ in their salience then the more prominent perceptual event would define a phrase boundary and thus determine the perceived rhythm pattern.

Another concern which was addressed in this study was the mediation of a harmonic phenomenal accent by the structural accent strength of a chord progression. In the first study, the structural accenting of different chord progressions was not manipulated. Only one chord progression was employed as a harmonic accompaniment (i.e., the I-V-I progression in the key of C Major). This led to the possibility that structural factors could have contributed to the perceptual salience of the chord change since the combination of the dominant chord and Middle C would have resulted in a relatively dissonant suspension. To determine the extent to which this dissonance could have contributed to the results found in the first experiment, three chord progressions were employed in each of the experiments reported in this chapter. These chord progressions varied in the relative consonance and dissonance of the constituent triads and Middle C of the temporal line (e.g., I, IV, and V chords in C Major).

A third concern addressed was that the perceptual salience of temporal sequences that coincided with metrical accents in Experiment 1 could have been inflated due to the behavioral measure used (i.e., tapping to the "beat" of the presentation). By having subjects give a behavioral response at specific locations in each sequence, a bias for certain temporal sequences could have been established, that in the absence of tapping would not have existed. In order to investigate this possibility, the same composite- and harmonic-rhythms used in Experiment 1 were used in this series of experiments, but subjects were not asked to indicate the inferred metrical structure. If similar patterns of results are found for the two harmonic conditions as that which was found in Experiment 1, then this would support the interpretation that the increased salience of certain temporal sequences was due to the directing of attentional resources (i.e., metrical accents). Following the procedure of Experiment 1, the harmonic accompaniment in this series of experiments was systematically varied to encourage the adoption of a particular inferred metrical structure. It was expected that subjects would show a bias for rhythm patterns that would be supported by either a 3/4 metre (e.g., LSSSS, SSLSS, and SSSSL) or by a 6/8 metre (e.g., LSSSS, and SSSLS), depending upon the accompaniment presented.

Finally, the effects of musical sophistication and training on the relative weights assigned to different phenomenal accents was of interest. For instance, one could ask if musicians and nonmusicians use harmonic and temporal accents differently in the perception of a rhythm pattern. As well, musicians and nonmusicians may or may not be similar in their inference and then employment of a metrical structure to guide their on-going attending in a analogous fashion to Jones' notion of the role of accent coincidences. Employing the same criteria used in Experiment 1, all participants in this series of experiments were divided into two groups on the basis of musical training. There is little doubt that musical experience, especially formal musical training, changes one's explicit knowledge-base of music, enhances one's vocabulary of musical terminology, and thus affects how one speaks, and possibly thinks of rhythmic structure. Furthermore, such a knowledge-base could alter the way a listener perceives the musical structure through conceptually-driven processing of incoming information. Similarities and differences in the perception of musical structure between musically experienced and inexperienced individuals can serve as an indication of those aspects of the perceptual process that are innate and generic, as well as those that have been mediated through experiential factors.

Experiment 2

This experiment explored the possibility that musicians and nonmusicians would employ harmonic and temporal cues differently in determining rhythm pattern perceptions. Furthermore, it was questioned whether or not musicians and nonmusicians would similarly infer and employ a 6/8 metrical structure with a bias for the rhythm patterns of LSSSS and SSSLS.

METHOD

Subjects. Eighteen experimentally naive volunteers participated in this study. There were 9 musicians with a range of 3.5 - 12 years and a mean of 6.8 years of musical experience. Their age ranged from 18 - 22 with a mean of 19.5 years and there were 3 males, 6 females. Nine subjects were nonmusicians with an age range of 19 - 21 and a median of 19.5 years. Two of these subjects were male and 7 were female. Subjects were recruited from introductory psychology classes, and were given course credit for participating in the experiment.

Apparatus. The diotic stimuli were generated by a Yamaha CX5M music computer and presented via a Yamaha CA-140 amplifier to Realistic Pro-2 earphones. Each subject sat at a table in a sound-attenuated, ventilated AEC chamber and responded to the stimuli by moving a computer mouse that communicated the subject's choices directly to the computer. A Hitachi 1481 Color Monitor was situated at eye-level directly in front of the chamber and was visible through a glass window. By moving the mouse on the table-top, the subject controlled the movements of the cursor visible on the screen of the monitor. The cursor had to be moved into one of five appropriate boxes, and the mouse button depressed to indicate the response choice for each trial. Data were automatically recorded and stored by the computer on disk.

Stimuli. Table 4 represents the stimuli used in this experiment. The stimuli consisted of the melody or temporal line in which a long duration note (where L = 444 msecs) was followed by 4 short duration notes (where S = 222 msecs). In musical notation, this is roughly equivalent to a quarter note followed by 4 eighth notes with a tempo of 135 beats per minute in common time. This pattern was repeated 10 times. The fifty notes presented were all Middle C. The reason for presenting all the tones at the same frequency was to eliminate melodic phenomenal and structural accents. On any given trial, one of fifteen different harmonic accompaniments were presented with the temporal line. Three progressions of five different locations each were used. The three progressions were I-V-I, I-IV-I, and IV-V-IV in the key of C Major. Within each cadence group each chord was presented six times (each time for a duration of 222 msecs) before changing. By systematically varying the location of the chord changes in relation to the long-duration tone, the melody was broken into five different rhythm patterns. The timbre was that of a FM-synthesized piano.

Procedure. A total of 60 trials were performed with 15 stimuli presented in random order four times. Subjects sat in a sound-attenuated chamber with a screen in front of them and were presented with the auditory stimuli over a set of earphones. The subjects were to listen to the stimuli and then choose which rhythm pattern they had heard. Subjects were under no time constraints and started each trial by placing the cursor into a start box. Data were also collected on each subject through a standard questionnaire on musical training, family musicians, general attitude towards music, age, gender, etc. Subjects were instructed to identify the rhythm pattern with each presentation, and told that they were free to use any strategy. The five possible choices were represented on the screen by five appropriately labelled boxes. Each

Table 4

<u>A Schematic Representation of Stimuli used in Experiments 2 and 3</u>

```
Temporal line:  L   S S S S L   S S S S L   S S S S L  ...

Experiment 2
 Harmony #1:    C C C C C C G G G G G C C C C C C G G...
 Harmony #2:    C C C C C G G G G G G C C C C C C G G G...
 Harmony #3:    C C C C G G G G G G C C C C C C G G G G...
 Harmony #4:    C C C G G G G G G C C C C C C G G G G G...
 Harmony #5:    C C G G G G G G C C C C C C G G G G G G...
Experiment 3
 Harmony #1:    C-----------G-----------C-----------G  ...
 Harmony #2:    C---------G-----------C-----------G--  ...
 Harmony #3:    C-------G-----------C-----------G----  ...
 Harmony #4:    C-----G-----------C-----------G------  ...
 Harmony #5:    C---G-----------C-----------G--------  ...
```

<u>Note</u>. This Table is read by following the location of a chord change upward to see where it intersects the temporal line. The accompaniment create five rhythm patterns based on a chord change: LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL. Although only the I-V-I progression is shown, I-IV-I and IV-V-IV progressions were also used.

possible rhythm pattern was tapped out prior to the presentation of the stimuli to ensure each subject understood what each schemata represented. Subjects were also told that there were no correct answers and that it was possible that all trials would sound as if they had identical rhythm patterns.

RESULTS AND DISCUSSION

An initial loglinear analysis was performed on the data, with three design and one response variable. The design variables were the musical Training of the subject (nonmusician or musician), the rhythm pattern as defined by the Location of the chord change (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL), and the Type of chord progression (I-V-I, I-IV-I, IV-V-IV). The response variable was the rhythm chosen by the subject (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL).

The loglinear analysis was implemented with the loglinear module in Statistica. An appropriate model was determined by first conducting tests of partial and marginal association. The results of these tests for association indicated the significant effects were: Choice (Partial $\chi^2(4) = 80.38$, $p < .001$; Marginal $\chi^2(4) = 80.38$, $p < .001$), Choice X Location (Partial $\chi^2(16) = 314.26$, $p < .001$; Marginal $\chi^2(16) = 313.22$, $p < .001$), and finally, Choice X Location X Training (Partial $\chi^2(16) = 87.01$, $p < .001$; Marginal $\chi^2(16) = 84.19$, $p < .001$). The interaction of Choice X Training approached significance (Partial $\chi^2(4) = 9.25$, $p = .055$; Marginal $\chi^2(4) = 8.53$, $p < .074$).

The second stage of the analysis involved the selection of a model that best fitted the observed data. This was done by automatic stepwise model selection. The best fitting model was Choice X Location X Training $\chi^2(100, N = 108) = 84.56$, p

= .87). Because this interaction contains all of the significant lower-order effects, the results can be completely specified by examining this interaction.

Table 5 shows the pattern of rhythmic choices as a function of the location of a chord change and musical training. The first thing to be noted in this data is that there was, as expected, a large percentage of choice for the rhythm patterns LSSSS and SSSLS. We interpret this as evidence of subjects inferring a 6/8 metrical structure, with the onset of the temporal accent serving as the initiation of a measure. Such an interpretation is supported by the results of Experiment 1 reported in Chapter 2.

Different patterns of responding were found for musicians and nonmusicians. Musicians primarily chose rhythm patterns on the basis of the location of a chord change. This is seen in the data by the large percentage of rhythmic choices along the negative diagonal: 71.3% - LSSSS, 43.5% - SLSSS, 47.2% - SSLSS, 54.6% - SSSLS, and 60.2% - SSSSL. This pattern of most responses along the negative diagonal is indicative of two things. First, the chord-change based rhythm pattern was usually chosen by the musicians; and second, when a musically experienced subject selected any particular rhythm pattern, it was usually the rhythm pattern determined by the location of a chord change.

In addition, musicians showed evidence of a preference for rhythm patterns based on accent coincidences. When there was a three-part coincidence consisting of a chord change, a long duration note, and an inferred metrical accent, resulting in the rhythm pattern of LSSSS, musicians selected this rhythm more often (71.3% of the time) than any other rhythm, under any other conditions. The rhythm

Table 5

Rhythmic Choice by Chord Change Location and Musical Training in Experiment 2

	Rhythm on the basis of the chord change				
Rhythmic choice	LSSSS	SLSSS	SSLSS	SSSLS	SSSSL
	Musicians				
LSSSS	71.3%	20.4%	17.6%	13.0%	17.6%
SLSSS	5.6%	43.5%	8.3%	5.6%	7.4%
SSLSS	2.8%	13.9%	47.2%	6.5%	0.9%
SSSLS	8.3%	9.3%	20.4%	54.6%	13.9%
SSSSL	12.0%	13.0%	6.5%	20.4%	60.2%
	Nonmusicians				
LSSSS	42.6%	24.1%	26.9%	21.3%	21.3%
SLSSS	9.3%	13.0%	11.1%	8.3%	13.0%
SSLSS	11.1%	26.9%	21.3%	11.1%	10.2%
SSSLS	14.8%	25.0%	29.6%	41.7%	25.9%
SSSSL	22.2%	11.1%	11.1%	17.6%	29.6%

pattern based on the accent coincidence of a chord change and the temporal accent (SSSSL) was selected more often than any other rhythm pattern (60.2% of the time) under any other conditions, except LSSSS. Finally, the rhythm pattern based on the coincidence of a chord change and an inferred metrical accent (SSSLS) was also selected frequently (54.6% of the time). Musicians did not show evidence of selecting rhythm patterns on the basis of either the temporal or metrical accent alone.

Nonmusicians, in contrast, did not show evidence of selecting rhythm patterns primarily on the basis of the chord change. Instead, they selected rhythm patterns mainly on the basis of an inferred 6/8 metrical structure (i.e., the rhythm patterns LSSSS and SSSLS). Selection of the rhythm pattern SSSLS on the basis of only the metrical accent was 14.8%, 25.0%, 29.6%, and 25.9% for each chord-change based rhythm pattern of LSSSS, SLSSS, SSLSS, and SSSSL, respectively. Selection of the rhythm pattern LSSSS on the basis of the metrical accent (and/ or possibly the temporal accent) was 24.1%, 26.9%, 21.3%, and 21.3% for each chord-change based rhythm pattern of SLSSS, SSLSS, SSSLS, and SSSSL, respectively. While it is true that nonmusicians may have selected this rhythm pattern of LSSSS under these conditions on the basis of the temporal accent alone, or in combination with the metrical accent, we believe it was based primarily on the metrical accent alone. Support for this view comes from the fact that the percent of selections for LSSSS under each chord-change condition was not higher than that which had been found for the rhythm pattern defined only on the basis of the metre (SSSLS). If selection of LSSSS was based on the coincidence of the metre and the temporal accent, we would expect higher percentage of responding for LSSSS than SSSLS. In addition, since subjects selected the rhythm pattern SSSLS on the basis of metre only, we believe it is more parsimonious to suggest that they selected the rhythm pattern LSSSS for the same reason. This is because the alternative, which is that they selected LSSSS on the basis of the temporal accent alone, would have to explain why metre was not used for this rhythm when it was used for SSSLS. As well, selection primarily on the basis of the metre can account for the conditions under which the lowest percentage of selections for the rhythm patterns of SSSLS, and LSSSS were found. The rhythm pattern SSSLS was selected least often (14.8% of the time) when the chord-change based rhythm pattern was LSSSS. This may have been due to the fact that when the chord change supported the rhythm LSSSS, there was a coincidence of a temporal accent, a chord change, and a metrical accent. The bias for the metre-supported rhythm of SSSLS is not seen under this condition because metre also supports LSSSS. Similarly, subjects chose the rhythm pattern of LSSSS least often (21.3% of the time) when the chord-change based rhythm pattern was SSSLS, possibly because the metrical bias for LSSSS was countered by the metrical bias for SSSLS.

Nonmusicians were qualitatively similar to musicians in that they selected rhythm patterns on the basis of accent coincidence. The rhythm pattern of LSSSS was selected more often (42.6% of the time) when it was supported by the coincidence of the three accents (metric, temporal, and harmonic), than was any other rhythm pattern under any other conditions. The rhythm pattern based on the accent coincidence of a chord change and an inferred metrical accent (SSSLS) was selected more often than any other rhythm pattern (41.7% of the time) under any other conditions, except LSSSS. Finally, the rhythm pattern based on the coincidence of a chord change and the temporal accent (SSSSL) was also selected frequently (29.6% of the time).

It appears that harmonic information was more salient for musicians than nonmusicians. In contrast, nonmusicians tended to emphasis the metrical aspects of a musical presentation when they perceived rhythm patterns. While the pattern of data supports this explanation, there is no direct evidence in the data that subjects did, in fact, infer a metrical structure of 6/8 time. Some indirect support comes from Chapter 2. In that study, when the harmonic accompaniment was the same as that which was used in this experiment, subjects reported an inferred 6/8 metrical structure 87.5% of the time. In order to further substantiate the claim that a metre was being inferred and that this inference affects the perception of a rhythm pattern, Experiment 3 was conducted in which the composite rhythm of the harmonic accompaniment was changed to support a different metrical inference (a 3/4 metre).

Experiment 3

In this experiment the harmonic accompaniment was changed to a chord being presented for a duration of 3 beats in common time before changing (i.e., 1332 msecs) rather than six times, each time being held for one half a beat (i.e., 222 msecs). Based on previous research (Experiment 1), this change should result in a change of the inferred metre from 6/8 time to 3/4 time. If subjects show a corresponding shift in preference for rhythm patterns consistent with the shift from 6/8 to 3/4 time, we would be more confident that the explanation given for the pattern of results in the previous experiment is accurate. Specifically, subjects should show a bias for the rhythm patterns of LSSSS, SSLSS, and SSSSL, if they infer a 3/4 metre with the long duration note coinciding with the onset of a measure.

The change in harmonic accompaniment involves presenting a temporal accent that coincides with each chord change. It was hypothesized that this may have two effects on the pattern of results apart from those associated with a change in metre. First, the coincidence of a temporal accent with each chord change should make rhythm patterns defined by the location of chord changes more salient for all subjects. Second, if nonmusicians are more affected by the temporal features of a musical presentation than are musicians, then we would expect nonmusicians to show a pattern of responding similar to that which was found for musicians in Experiment 2. That is, nonmusicians would select rhythm patterns consistent with rhythms defined by the location of a chord change, not, like musicians, because of the chord change per se, but rather because of the associated temporal accent that accompanies the chord presentation.

METHOD

Subjects. Forty-three experimentally naive volunteers participated in this study. There were 19 musicians with a range of 3 - 12 years and a median of 9 years of musical training. Their ages ranged from 18 - 25 with a median of 19 years and there were 9 males, 10 females. Twenty-four subjects were nonmusicians with an age range of 18 - 28 and a median of 19 years. Ten of these subjects were male and 14 were female. Subjects were recruited from introductory psychology classes, and were given course credit for participating in the experiment.

Apparatus. The equipment employed was the same as that used in the previous experiment.

Stimuli. The stimuli were similar to those used in Experiment 2, both in terms of the temporal line and the accompanying harmonic structure. Three progressions of five different locations each were used, as in Experiment 2. The three progressions were (I-V-I), (I-IV-I), and (IV-V-IV). Within each progression condition the chord changes occurred in different locations in relation to the temporal phenomenal accent such that the temporal line was broken into five different rhythm patterns (i.e., LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL). The only difference in the stimuli was the presentation of the harmonic component. Rather than presenting each chord of the harmonic accompaniment six times, each time being held for one half a beat in common time before changing, a chord was presented for a duration of 3 beats in common time and then changed. In this way, not only would each chord change have a corresponding temporal accent associated with it, but also the perceived metrical structure should be 3/4 time.

Procedure. The presentation of stimuli was the same as in the previous experiment.

Results and Discussion

An initial loglinear analysis was performed on the data, with three design and one response variable. The design variables were the musical Training of the subject (nonmusician or musician), the rhythm pattern as defined by the Location of the chord change (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL), and the Type of chord progression (I-V-I, I-IV-I, IV-V-IV). The response variable was the rhythm chosen by the subject (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL).

As in Experiment 2, the loglinear analysis was implemented with the loglinear module in Statistica. The results of the tests for partial and marginal association indicated the significant effects were: Choice (Partial $\chi^2(4) = 13.19$, $p < .01$; Marginal $\chi^2(4) = 13.19$, $p < .01$), Training (Partial $\chi^2(1) = 33.97$, $p < .001$; Marginal $\chi^2(1) = 33.97$, $p < .001$), Choice X Location (Partial $\chi^2(16) = 1212.82$, $p < .001$; Marginal $\chi^2(16) = 1212.08$, $p < .001$), and finally, Choice X Location X Training (Partial $\chi^2(16) = 538.55$, $p < .001$; Marginal $\chi^2(16) = 535.11$, $p < .001$).

The second stage of the analysis, the automatic stepwise model selection, indicated that the best fitting model was Choice X Location X Training $\chi^2(100, N = 2580) = 82.97$, $p = .89$). Because this interaction contains all of the significant lower-order effects, the results can be completely specified by examining this interaction.

The interaction of choice by location of chord change by musical experience is presented in Table 6. As expected, the harmonic phenomenal accent carried greater salience for musicians than for nonmusicians. Musicians showed the characteristic pattern of responding on the basis of the chord change as is evident in Table 6 by the large percentage of responses along the negative diagonal (85.5%, 63.2%, 78.1%, 70.6%, and 75% for each chord change based rhythm pattern of LSSSS, SLSSS, SSLSS, SSSLS, and SSSSL, respectively). Musicians showed evidence of both an inferred 3/4 metre, and a preference for rhythm patterns defined by accent coincidence (i.e., rhythm patterns LSSSS, SSLSS, and SSSSL). The musicians' highest percentage of responses was for the rhythm pattern LSSSS when it was defined by the accent coincidence of a temporal accent in the melody (the onset of the long duration note), a chord change, the temporal accent associated with the harmonic accompaniment, and

Table 6

Rhythmic Choice by Chord Change Location and Musical Training in Experiment 3

	Rhythm on the basis of the chord change				
Rhythmic choice	LSSSS	SLSSS	SSLSS	SSSLS	SSSSL
	Musicians				
LSSSS	85.5%	7.9%	4.4%	4.4%	5.3%
SLSSS	1.3%	63.2%	3.9%	7.0%	13.2%
SSLSS	3.5%	7.0%	78.1%	10.5%	2.6%
SSSLS	2.2%	12.3%	8.8%	70.6%	3.9%
SSSSL	7.5%	9.6%	4.8%	7.5%	75.0%
	Nonmusicians				
LSSSS	46.2%	18.8%	16.3%	16.3%	19.1%
SLSSS	10.8%	20.1%	14.6%	20.1%	26.4%
SSLSS	14.2%	19.1%	32.6%	21.2%	12.8%
SSSLS	14.9%	24.0%	21.2%	21.5%	12.8%
SSSSL	13.9%	18.1%	15.3%	20.8%	28.8%

an inferred metrical accent. The next highest percentages of responses for musicians were for coincidences of three accents (a chord change, the temporal accent associated with the harmonic accompaniment, and an inferred metrical accent resulting in the rhythm SSLSS -- 78.1%), and four accents (a chord change, the temporal accent associated with the harmonic accompaniment, an inferred metrical accent, and the offset of the temporal accent in the melody resulting in the rhythm SSSSL -- 75%).

That musicians did not respond more often for the rhythm pattern of SSSSL compared to SSLSS when they were supported by the accent coincidences, is indicative of the non-use of the temporal accent in the melody line by musicians in determination of phrase boundaries. It may be that musicians employed the temporal accent in the melody as a cue for the initiation of the major metrical unit (i.e., the measure). This explanation accounts for the highest percentage of responses for LSSSS when it was defined by the coincidence of accents, since the metrical accent associated with the onset of a measure is the strongest, and for the fact that there were not more responses for the rhythm pattern SSSSL when it was defined by accent coincidence.

Comparison of musicians' responses in this experiment with that of the previous experiment (Table 5) indicates a greater percentage of responses along the negative diagonal for this data. This was expected since the chord changes have an associated temporal accent. Compared to musicians, nonmusicians showed a different pattern of responses. Despite the temporal accent associated with the chord change, nonmusicians did not primarily select rhythm patterns defined by the harmonic event. Like musicians, nonmusicians selected mainly on the basis of accent coincidences with an inferred 3/4 time metrical accent (LSSSS -- 46.2%; SSLSS -- 32.6%; and SSSSL -- 28.8%). When a chord change alone defined the rhythm pattern of SSSLS, nonmusicians did not show evidence of significantly selecting that rhythm (21.5%) over the rhythm patterns of SLSSS (20.1%), SSLSS (21.2%), or SSSSL (20.8%). When a chord change alone defined the rhythm pattern of SLSSS, nonmusicians had a tendency to select the rhythm pattern of SSSLS (24%) rather than SLSSS (20.1%). This tendency reflects some ambiguity of metrical structure for nonmusicians, and in particular, for this rhythm pattern. The grouping of three short duration notes together (i.e., SL - SSS) by the chord change may reinforce the inference of a 6/8 metre rather than a 3/4 metre, since in the latter case there would be some form of syncopation (i.e., the onset of a note off the beat).

In summary, it appears that a change in the harmonic accompaniment can lead to the hypothesized change in inferred metre. Both musicians and nonmusicians showed evidence of principally inferring a 3/4 metrical structure, with the long duration note in the melody initiating a measure. Musicians had a tendency to principally employ the location of a chord change in the perception of a rhythm pattern, but also showed evidence of responding to accent coincidences. This supports models of musical perception based on accent coincidence (Berry, 1976; Benjamin, 1984; Jones, 1987, 1992). Nonmusicians also showed evidence of responding on the basis of accent coincidence, but unlike musicians did not principally employ harmonic information in the determination of rhythm patterns.

It is somewhat surprising that in both this experiment and Experiment 2, there were no significant differences found in the salience of chord changes as a function of the progression condition. The failure to find such differences is indicative of the problems involved in the incorporation of harmonic structural accenting in the investigation of rhythmic structure and may be due to several factors. First, it may be that the structural accenting of a chord change does not affect the perceived phenomenal event, however this seems unlikely. There is abundant evidence that supports the theoretical construct of stability (Bharucha & Krumhansl, 1983; Cook, 1978; Krumhansl, 1979; Krumhansl, Bharucha & Kessler, 1982; Krumhansl & Castellano, 1983; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979; Meyer, 1956; Stark, 1988). Second, it may be that the three progression conditions which were employed in the first two experiments were not sufficiently different. The three progression conditions were all common diatonic chord progressions. Finally, in all progression conditions, movement towards a tonal centre was matched by an equivalent size movement away from the tonal centre. It may be that any sense of tonal stability that resulted when a chord changed towards the tonal centre was countered by an equivalent amount of tension when the chord changed back.

To test this explanation, Experiment 4 was conducted in which the harmonic accompaniment was made up of common and uncommon diatonic chord progressions as well as a nondiatonic progression.

Experiment 4

In this experiment, an attempt was made to further differentiate the progression conditions. Rather than use progression which occur often in traditional Western music, (i.e., I-V-I, I-IV-I, and IV-V-IV), the overall tension of the progression was varied. This was done by changing the distance of the component chords from the tonal centre of C which is the note within the melodic or temporal line. In order to do this, a nondiatonic and unstable progression, a diatonic, but uncommon progression, and finally, a common diatonic progression were employed.

METHOD

Subjects. Sixteen experimentally naive volunteers participated in this study, 7 musicians and 9 nonmusicians. The musicians had a range of 5-8 years and a mean of approximately 6 years of musical training. Their age ranged from 19-41 with a median of 19 years and there were 3 males and 4 females. Nonmusicians had an age range of 19-24 with a median of 19 years. Three of these subjects were male and six were female. Subjects were recruited from introductory psychology classes, and were given course credit for participating in the experiment.

Apparatus. The equipment employed in this experiment was the same as that used in the previously reported experiments.

Stimuli. The stimuli in this experiment were similar to those used in Experiment 2 except that the harmonic presentations that accompanied the melody were different. Three progression groups of five different locations each were used as in Experiment 2. The three progression groups were, in the key of C Major: a stable and expected progression (I-V-I), a diatonic but unexpected progression (V-VII-V), and a nondiatonic and very unstable progression (C# major triad - B diminished triad - C# major triad). As in the previous experiments, the chord changes occurred in different locations within each progression condition in relation to the long duration tone such that the temporal line was broken into five different rhythm patterns. Each chord presentation was for approximately 222 msecs and was equivalent to the duration of a short duration note in the temporal line. Thus, as in Experiment 2, it was expected that subjects would infer a 6/8 metre.

Procedure. The presentation of stimuli was essentially the same as in the previous experiment. There were a total of 60 trials with the 15 stimuli presented in random order four times. Timing and task demands were the same as in previous experiments.

Results and Discussion

An initial loglinear analysis was performed on the data, with three design and one response variable. The design variables were the musical Training of the subject (inexperienced or experienced), the rhythm pattern as defined by the Location of the chord change (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL), and the Type of chord progression [common diatonic (C-G-C), uncommon diatonic (G-Bdim-G), and nondiatonic (C#-Bdim-C#)]. The response variable was the rhythm chosen by the subject (LSSSS, SLSSS, SSLSS, SSSLS, SSSSL).

As in the previous experiments, the loglinear analysis was implemented with the loglinear module in Statistica. The results of the tests for partial and marginal association indicated the significant effects were: Choice (Partial $\chi^2(4) = 14.22$, $p < .01$; Marginal $\chi^2(4) = 14.22$, $p < .01$), Training (Partial $\chi^2(1) = 13.94$, $p < .001$; Marginal $\chi^2(1) = 13.94$, $p < .001$), Choice X Training (Partial $\chi^2(4) = 2.42$, $p < .001$; Marginal $\chi^2(4) = 19.38$, $p < .001$), Choice X Location (Partial $\chi^2(16) = 233.83$, $p < .001$; Marginal $\chi^2(16) = 232.15$, $p < .001$), Choice X Chord type (Partial $\chi^2(8) = 17.20$, $p < .05$; Marginal $\chi^2(8) = 16.24$, $p < .05$), Choice X Location X Training (Partial $\chi^2(16) = 51.47$, $p < .001$; Marginal $\chi^2(16) = 52.33$, $p < .001$), and finally, Choice X Location X Chord type (Partial $\chi^2(32) = 45.67$, $p < .05$; Marginal $\chi^2(32) = 46.45$, $p < .05$).

The second stage of the analysis, the automatic stepwise model selection, indicated that the best fitting model was Choice X Location X Training, and Choice X Location X Chord type, $\chi^2(50, N = 960) = 41.53$, $p = .80$). Because these two interactions contain all of the significant lower-order effects, the results can be completely specified by examining them.

Table 7 shows the effect of rhythmic choice by location of the chord change by training. Comparing nonmusicians to musicians we find that musicians chose a rhythm on the basis of the chord change more often than did nonmusicians. Responses based on the location of the chord change ranged from 44.0% for LSSSS to 56.0% for SSSLS for musicians compared to a range of 28.7% for SSLSS to 40.7% for LSSSS for nonmusicians. Unlike the previous experiments, and in direct contrast to nonmusicians, musicians did not choose the rhythm pattern based on the three way coincidence of a temporal, a metrical, and a harmonic accent most frequently (only 44.0%). Instead, the rhythm pattern based on the coincidence of a metrical accent and the location of a chord change resulting in the rhythm pattern SSSLS was selected most often by musicians (56.0%). This percentage count was not much larger than that which was found for rhythm patterns based on either a chord change alone (52.4% for SLSSS, and 50.0% for SSLSS), or for the coincidence of the chord change with the offset of the temporal accent (51.2% for SSSSL).

Table 7

Rhythmic Choice by Chord Change Location and Musical Training in Experiment 4

	Rhythm on the basis of the chord change				
Rhythmic choice	LSSSS	SLSSS	SSLSS	SSSLS	SSSSL
		Musicians			
LSSSS	44.0%	6.0%	11.9%	10.7%	11.9%
SLSSS	31.0%	52.4%	4.8%	13.1%	6.0%
SSLSS	7.1%	8.3%	50.0%	10.7%	7.1%
SSSLS	3.6%	16.7%	21.4%	56.0%	23.8%
SSSSL	14.3%	16.7%	11.9%	9.5%	51.2%
		Nonmusicians			
LSSSS	40.7%	29.6%	25.0%	21.3%	27.8%
SLSSS	14.8%	33.3%	17.6%	13.9%	11.1%
SSLSS	13.9%	12.0%	28.7%	15.7%	13.0%
SSSLS	16.7%	11.1%	20.4%	34.3%	18.5%
SSSSL	13.9%	13.9%	8.3%	14.8%	29.6%

Nonmusicians showed the highest percentage of rhythmic choice when the temporal and harmonic accents coincided with an inferred metrical accent (LSSSS -- 40.7%) supporting the view that accent coincidence is of importance in the perception of rhythmic structure. They responded next highest for the rhythm pattern based on the coincidence of the chord change and an inferred metrical accent (SSSLS -- 34.3%). Rhythmic responses based on either a chord change alone (33.3% for SLSSS, and 28.7% for SSLSS), or the coincidence of a chord change and the offset of a temporal accent (29.6% for SSSSL) were the next highest percentage counts. As in the previous experiments, nonmusicians still, for the most part, showed evidence of responding on the basis of an inferred metrical accent. Responses on the basis of the metrical accent (or possibly the temporal accent as well) for the rhythm pattern LSSSS were 29.6%, 25.0%, 21.3%, and 27.8% when the chord changes supported the rhythm patterns SLSSS, SSLSS, SSSLS, and SSSSL, respectively. Unlike the findings for nonmusicians in the previous experiments however, the rhythmic responses based on the location of a chord change alone were greater than the rhythmic responses based on an inferred metrical accent alone in this experiment. This means that nonmusicians used a chord change as a phrase boundary cue more often when the chords formed an uncommon and unstable progression than when the progression was stable and common, and in the process behaved like musicians did when the harmonic progressions were common and diatonic.

Table 8 shows the significant interaction of rhythmic choice by location of a chord change by the type of chord progression. The pattern of responses for the common diatonic progression condition was very similar to the combined pattern of responses given by nonmusicians and musicians in the previous experiments. Like musicians in the previous experiments, subjects responded on the basis of the location of a chord change, and accent coincidences; and like nonmusicians, subjects responded on the basis of an inferred 6/8 metre with a metrical accent supporting the rhythm patterns of LSSSS and SSSLS. For the common diatonic chord progression, the characteristic largest percent of responses occurred when the relatively long duration note coincided with a chord change and an inferred metrical accent resulting in the pattern LSSSS (60.9%). As well, a large percent of responses occurred along the negative diagonal indicating that when a chord change supported a particular rhythm pattern, subjects selected that rhythm pattern, and when subjects selected a rhythm pattern it was primarily on the basis of the chord change. Responses on the basis of chord change alone for the common diatonic chord progression ranged from 35.9% for SSLSS to 43.8% for both SLSSS and SSSLS. Responses for rhythm patterns based on an inferred metrical accent (and possibly a temporal accent as well -- LSSSS) were 21.9%, 23.4%, 17.2% and 23.4% for each chord-change based rhythm pattern of SLSSS, SSLSS, SSSLS, and SSSSL, respectively. Responses for the metrically-supported rhythm pattern of SSSLS were 10.9%, 14.1%, 21.9%, and 21.9% for each chord-change based rhythm pattern of LSSSS, SLSSS, SSLSS, and SSSSL, respectively. As in Experiment 2, the lowest percentage of selection for the metrically-supported rhythm pattern SSSLS was when the chord change supported the metrically-supported rhythm pattern of LSSSS. Similarly, the lowest selection for the metrically-supported rhythm pattern LSSSS was when the chord change supported the metrically-supported rhythm pattern of SSSLS.

Table 8

Rhythmic Choice by Chord Change Location by Progressions in Experiment 4

	Rhythm on the basis of the chord change				
Rhythmic choice	LSSSS	SLSSS	SSLSS	SSSLS	SSSSL
C - G - C Progression Condition					
LSSSS	60.9%	21.9%	23.4%	17.2%	23.4%
SLSSS	12.5%	43.8%	12.5%	15.6%	3.1%
SSLSS	6.3%	9.4%	35.9%	10.9%	12.5%
SSSLS	10.9%	14.1%	21.9%	43.8%	21.9%
SSSSL	9.4%	10.9%	6.3%	12.5%	39.1%
G - Bdim - G Progression Condition					
LSSSS	46.9%	21.9%	12.5%	15.6%	17.2%
SLSSS	10.9%	35.9%	15.6%	14.1%	10.9%
SSLSS	7.8%	7.8%	35.9%	14.1%	12.5%
SSSLS	9.4%	18.8%	20.3%	45.3%	18.8%
SSSSL	25.0%	15.6%	15.6%	10.9%	40.6%
C# - Bdim - C# Progression Condition					
LSSSS	18.8%	14.1%	21.9%	17.2%	21.9%
SLSSS	42.2%	45.3%	7.8%	10.9%	12.5%
SSLSS	18.8%	14.1%	42.2%	15.6%	6.3%
SSSLS	12.5%	7.8%	20.3%	42.2%	21.9%
SSSSL	7.8%	18.8%	7.8%	14.1%	37.5%

The pattern of results for the uncommon diatonic chord progression was also, for the most part, similar to that of the common diatonic in previous experiments. There was the characteristic large percent of responses for the simultaneous occurrence of the temporal accent, metrical accent, and a chord change (46.9%) but it was not as strong as that of the common diatonic (60.9%). The trend of choosing on the basis of the chord change alone was evident in this condition as well; 35.9% for both SLSSS and SSLSS, 45.3% for SSSLS, and 40.6% for SSSSL. Responses for rhythm patterns based on coincidence of two or more accents was also found to be greater than the percentage of responses for rhythm patterns defined by only one phenomenal event.

What differed most markedly in the pattern of responses for the uncommon diatonic condition compared to the common diatonic condition was the tendency, or rather nontendency, of subjects to select a rhythm based on the location of an inferred 6/8 metrical accent alone.

With the nondiatonic chord condition there was no evidence of responding on the basis of the coincidence of the temporal, metrical, and harmonic accents. In fact, under such circumstances, subjects selected the rhythm pattern of SLSSS most often (42.2%) rather than the rhythm pattern of LSSSS (18.8%). With the exception of the three-way coincidence resulting in the rhythm pattern LSSSS, there was evidence that subjects under the nondiatonic condition still selected rhythm patterns based on the location of a chord change; 45.3% for SLSSS, 42.2% for both SSLSS and SSSLS, and 37.5% for SSSSL. Furthermore, the nondiatonic chord condition resembled the uncommon diatonic progression condition in that there was very little evidence of selecting rhythm patterns on the basis of either a temporal accent or the metrical accent alone. Unlike the findings for either diatonic condition, there was no evidence of responding on the basis of accent coincidence at all in the nondiatonic chord condition.

GENERAL DISCUSSION

An extrapolation across the results of the experiments in this chapter supports a number of general and specific conclusions, as well as raises several questions.

Unlike the position taken by Palmer and Krumhansl (1987a, 1987b) and Jones (1987), there was evidence that the two theoretical component categories manipulated in this series of experiments were not perceptually equivalent. The difference in the results of this investigation and those cited may be due to the fact that different component categories have been manipulated. The data generally support the position taken by Jones (1992) that phenomenal accents are not equally weighted; and specifically, are in agreement with the emphasis placed on harmonic events by traditional Western music theorists. Jones' (1992) notion of accent weights has been extended here to include the differences in perceptual salience between musicians and nonmusicians.

Musicians had a tendency across the experiments to primarily use the chord change as the basis for reporting a rhythm pattern, while nonmusicians primarily used an inferred metre as the basis for their perception of a rhythm pattern. The emphasis placed on metrically-supported rhythm patterns by nonmusicians compared to musicians was evident in their tendency to select rhythm patterns consistent with an inferred metre, regardless of where a chord change occurred in relation to a temporal accent. As well, the preference for metrically-defined, as opposed to harmonically-defined rhythm patterns, by nonmusicians was evident in the pattern of responses based on accent coincidence in Experiment 2. For nonmusicians, the percentage of responses for a rhythm defined by the two-way interaction of a chord change and a metrical accent (SSSLS) was greater than the two-way interaction of a chord change and a temporal accent (SSSSL), while for musicians, the opposite was the case.

The emphasis placed on harmony by musicians in this series of experiments may be indicative of the training they have received in traditional Western music theory with its heavy emphasis on harmonic functions. While many theorists would explain the differences in terms of musicians acquiring either implicit or explicit knowledge of a grammar (i.e., the rules that determine admissible and inadmissible constructions in some medium; e.g., Lerdahl & Jackendoff, 1983), there are alternative explanations. For example, musicians may have a greater knowledge of prototypical boundaries and characteristics (Krumhansl, Bharucha, & Kessler, 1982). In addition, as Halpern and Bower (1982) pointed out, not only do musicians learn some formal theory that explicitly sets out musical structure, but they also hear more music in social, academic, and professional settings; and gain greater understanding of music at the microscopic and global level when practicing and performing. The differences between the musicians and nonmusicians may, therefore, be due to the greater number of past episodes which musicians have at their disposal, and upon which they can draw when approaching the perceptual task employed in this study. This latter interpretation has in its support a number of musicological investigations showing that statistical properties of traditional Western music are remarkably similar to those which have been found in psychological investigations of music perception (Krumhansl, 1985; Palmer & Krumhansl, 1990).

Musicians and nonmusicians showed similarities in their patterns of responses as well. When diatonic chord progressions were used, both musicians and nonmusicians selected rhythm patterns that were defined by accent coincidences. The more accents that coincided, the greater the tendency was for subjects to select the rhythm pattern defined by the coincidence. For example, responses on the basis of a chord change either alone or in concert with other accents were greater in Experiment 3 than they were for the corresponding event in Experiment 2 or 4. This was undoubtedly due to the additional temporal accent that coincided with the chord change in Experiment 3. Within each experiment there was also support for the notion that the more accents coincide, the more perceptually salient the event will be (Benjamin, 1984; Berry, 1976; Jones, 1987). In Experiment 2, for example, the percentage of rhythmic responses was greater for both musicians and nonmusicians when there was a three-way coincidence of a temporal accent, a chord change, and an inferred metrical accent (LSSSS) than when there was either a two-way interaction of a chord change and an inferred metrical accent (SSSLS) or a two-way interaction of a chord change with the offset of a long duration note (SSSSL).

The only exception to the tendency to select on the basis of accent coincidence was in Experiment 4, when an uncommon and nondiatonic chord progression was used as the harmonic accompaniment. Under this condition, there was neither evidence of an inferred metrical structure, nor of responses based on coincidence of harmonic and temporal accents for either musicians or nonmusicians.

Finally, something must be stated regarding the inferential process of metre determination and how it interacted with other phenomenal and structural accents in the musical flow to result in a perceived rhythm pattern. In the three experiments presented in this chapter, metrically important locations were not marked by phenomenal accents such as the slight elongation of a note's duration or a subtle change in dynamics (Clarke, 1985; Gabrielsson, 1974, 1987). Instead, the composite rhythm was varied. In Experiments 2 and 4, the presentation of six chords within the time-span defined by the chord change resulted, it is believed, in the inference of a 6/8 metre. In Experiment 3, presentation of one chord held for the full duration before changing, resulting in the inference of a 3/4 metre. This conclusion is supported by the study reported in Chapter 1 in which subjects were requested to report the metre in which they were operating under the two conditions used in this study.

Based on the pattern of rhythmic responses, it is concluded that temporal sequences in which there was a coincidence of the temporal accent and an inferred metrical accent were preferred. The primary exception was in Experiment 3 for nonmusicians. While overall, nonmusicians selected a 3/4 metre as is evident in the high percentages for accent coincidences of a 3/4 metrical accent and a chord change location (resulting in the rhythm patterns LSSSS, SSLSS, and SSSSL), there was some ambiguity regarding the inferred metre and where in the musical flow metrical accents would occur when there were no accent coincidences. This ambiguity may have been due in part to the coincidence of a temporal accent with the chord change which resulted in a perceptually salient event which nonmusicians may have then used, instead of the temporal accent in the melody line, to mark a major metrical location.

In conclusion, the results of this study indicate that simple notions of perceptual accent weight which do not include recognition of prior musical experience and training will not be adequate. Reliable differences were found between musicians and nonmusicians, across the three experiments in terms of the perceptual salience of chord changes and temporal accents. In addition, there was some indication that structural accenting plays an important role in the salience of harmonic events, although more research in this area is necessary. Despite these sources of variability in studies of accent weights, several invariant features were found in this study across the three experiments. All subjects were found to report rhythm patterns more frequently when there were accent coincidences supporting the reported rhythm. In addition, all subjects similarly inferred a metrical structure and then used this inference as a guide for directing their attention to future locations in the musical flow. Both of these conclusions support Jones' notion that the rhythmicity of attention can be dynamically shaped on the basis of events defined by phenomenal accent coincidence in the music (Jones, 1976, 1984; Jones Boltz, & Kidd, 1982; Jones, Kidd, & Wetzel, 1981).

CHAPTER IV

Series 3: Inference of metrical structure from perception of iterative pulses within time-spans defined by chord changes.

In the previous experiments a harmonic and a temporal accent were pitted against each other in such a way as to form five possible rhythm patterns: LSSSS, SLSSS, SSLSS, SSSLS, SSSSL. In addition, two types of harmonic accompaniments that varied in the frequency of chord presentations (i.e., the composite-rhythm) but not the frequency of chord changes (i.e., the harmonic-rhythm) were presented. In Experiment 1, measures of both the perceived rhythm and the inferred metre were taken. It was found that subjects principally used the location of a chord change as the basis for their rhythm pattern perceptions, while the inferred metrical structure was based on the composite-rhythm (i.e., the frequency of chord presentations) within a time-span defined by: (a) the onset of the long duration note, (b) the offset of the long duration note, or (c) the chord changes. In all the previous studies, it was impossible to know within which time-span subjects perceived the iterative pulses that ultimately determined their inference of a metrical structure. The time-spans defined by the harmonic and temporal accents were all equivalent. However, in those few instances where selection of the metrical structure was not based on the composite-rhythm, it appeared to be based on the rhythm pattern defined by the location of the chord change, suggesting subjects may have been using the harmonic accent to group the iterative pulses that determined their inference of a metrical structure.

The following series of experiments were designed to test which phenomenal accents subjects use to group iterative pulses in the process of determining a metre. In order to do this, different time-spans defined by harmonic, melodic, and temporal phenomenal accents, and their coincidences were systematically pitted against each other. Subjects were requested to identify the metre of the stimuli as belonging to a category of either: (a) a triple metre (6/8 or 3/4 time), or (b) a duple metre (4/4 time or 2/4 time). In Experiment 5 the metres based on harmonic and temporal phenomenal accents were orthogonally varied. In Experiment 6, metres based on harmonic and melodic phenomenal accents were presented. Finally, in Experiment 7, temporal and melodic accent-based metres were presented. Subjects were asked to report the perceived metre. In this way the relative contribution of harmonic, temporal, and melodic phenomenal accents to the perception of metrical structure was delineated.

As in the previous experiments, all participants were divided into two groups on the basis of musical training. The results from the two groups were compared to determine the effects of musical experience.

Experiment 5

Method

Subjects. Twenty-six experimentally naive volunteers participated in this study. There were 13 musicians with a range of 5 - 10 years and a median of 8 years of formal music training. Their age ranged from 18 - 23 with a median age of 19 years. Of the 13 musicians, 7 were females and 6 were males. Thirteen subjects were nonmusicians with an age range of 18 - 25 and a median of 19 years. Eight of these subjects were female and 5 were males. Subjects were recruited from introductory psychology classes and were given course credit for participating in the experiment.

Apparatus. The diotic stimuli were generated by a Yamaha CX5M music computer and presented via a NAD 3020e integrated amplifier to AKG acoustic K340 earphones. Each subject sat at a table in a sound-attenuated, ventilated chamber and responded to the stimuli by moving a joystick that communicated the subjects' choices directly to the computer. A Hitachi 1481 Color Monitor was situated at eye-level directly in front of the chamber and was visible through a glass window. By moving the joystick, the subject controlled the movements of the cursor visible on the screen. The cursor had to be moved into one of two appropriate boxes to indicate the response choice for each trial. Data were automatically recorded and stored by the computer on disk.

Stimuli. The stimuli consisted of a temporal line and a corresponding harmonic accompaniment. There were two temporal conditions and two harmonic conditions. In the duple-metre temporal condition, one basic pattern of a long duration tone followed by two short tones was presented 18 times such that the pattern can be represented by the scheme LSSLSSLSS-etc. In the triple-metre temporal condition, a long duration tone followed by 4 short duration tones was presented 12 times such that the pattern presented can be represented with the scheme LSSSSLSSSSLSSSS-etc. The duple-metre harmonic condition consisted of four presentations of the same chord before it changed to a different chord. The cadence used in this experiment was C-G-C, or the I-V-I cadence in the key of C Major. Each chord presentation was set at a duration equal to the duration of a short-duration note in the temporal line. Thus the presentation of the harmonic accompaniment in this condition can be represented by the scheme C-C-C-C-G-G-G-G-C-C-C-C-etc. The triple-metre harmonic condition consisted of six presentations of the same chord before it changed to a different chord. The presentation in this condition can be represented by the scheme C-C-C-C-C-C-G-G-G-G-G-G-etc. The two harmonic conditions were orthogonally combined with the two temporal conditions resulting in four different conditions. In two of these four conditions, both the harmonic and temporal phenomenal accents support the same metre. In the other two conditions the harmonic and temporal accents support different metres. The tempo of each presentation was 135 beats per minute with the long-duration tone in the temporal line being equivalent to a quarter note and the short tones and chord presentations equal to eighth notes. The timbre was that of a synthesized piano.

Procedure. A total of 48 trials, 12 in each of the 4 conditions, were run. Subjects sat in a sound-attenuated chamber with a screen in front of them and were presented with the stimuli over a set of earphones. By placing the cursor in a start box at the top of the screen, subjects were first presented with four brief melodies; two of which were written in 4/4 time and two of which were written in a 6/8 time metre. The appropriate metrical category (triple, or duple metre respectively) was identified to the subjects on the video monitor as each melody played. After this initial training phase, subjects were instructed to listen to each trial and to identify the presented stimuli as being an example of either a triple or a duple metre. Subjects were to make their selection of metre by placing the cursor in one of two appropriately labelled boxes on the screen and then pushing a button on the joystick in order to register their choice. Subjects were under no time constraints and started each trial by placing the cursor in a start box.

Results and Discussion

An initial loglinear analysis was performed on the data, with three design and one response variable. The design variables were the musical Training of the subject (musician or nonmusician), the metrical category defined by the Temporal conditions (triple and duple metre), and the metrical category defined by the Harmonic conditions (triple and duple metre). The response variable was the metre chosen by the subject (triple and duple).

The loglinear analysis was implemented with the loglinear module in Statistica. The results of tests for partial and marginal association indicated the significant effects were: Choice X Harmonic condition (Partial $\chi^2(1) = 974.33$, $p < .001$; Marginal $\chi^2(1) = 970.14$, $p < .001$), Choice X Harmonic condition X Training (Partial $\chi^2(1) = 22.58$, $p < .001$; Marginal $\chi^2(1) = 22.57$, $p < .001$), and finally, Choice X Harmonic condition X Temporal condition (Partial $\chi^2(1) = 4.00$, $p < .05$; Marginal $\chi^2(1) = 3.72$, $p = .05$).

The second stage of the analysis fitted a model by automatic stepwise model selection. The best fitting model included all variables: Choice X Temporal condition X Harmonic condition X Training $\chi^2(0, N = 1152) = 0$, $p = 1.00$).

The significant interaction of choice by harmonic condition indicated subjects primarily selected metres consistent with the harmonic conditions, with a slight bias for duple metre over triple metre. When the harmonic condition supported a duple metre, subjects selected a duple metre 92.6% of the time, and when the harmonic condition supported a triple metre, subjects selected a triple metre 89.4% of the time. This bias may be reflective of a statistical property in traditional Western music favoring common time (e.g., 4/4 metre) over complex time (e.g., 6/8 metre).

The three-way interaction of choice by harmonic condition by training indicated that musicians chose a metre on the basis of the harmonic condition more reliably than did nonmusicians. When the harmonic condition supported a triple metre, musicians selected a triple metre 91.7% of the time compared to nonmusicians who selected a triple metre 87.2% of the time. When the harmonic condition supported a duple metre, musicians selected a duple metre 97.4% of the time while nonmusicians selected the duple metre 87.8% of the time. The slight preference for the duple metre was evident in the musicians responses (97.4% compared to 91.7% when the harmonic condition supported duple and triple metres, respectively), but was not seen in the responses made by the nonmusicians.

The three-way interaction of choice by temporal and harmonic conditions indicated that when the temporal condition supported a triple metre, subjects equally selected a triple and a duple metre when the harmonic conditions supported a triple and a duple metre respectively (89.7% of the time). Similarly, when the temporal condition supported a duple metre and the harmonic condition supported a triple metre, subjects selected a triple metre 89.1% of the time. Selection of metres in each of these three conditions was roughly equivalent and was based on which metrical structure was supported by the harmonic condition. When the temporal condition was duple metre and the harmonic condition supported duple metre, there was a relative inflation in the percentage of metrical selection in favor of the duple metre (95.5% of the time). This again is interpreted in terms of an overall bias in favor of common time metres in traditional Western music.

All of these significant effects are evident in Table 9 which shows the four way interaction of metrical choice by training by temporal and harmonic conditions. Both musicians and nonmusicians primarily selected a metre on the basis of the harmonic condition. When the harmonic and temporal conditions were both triple metres, musicians selected a triple metre 91% of the time, and nonmusicians selected this metre 88.5% of the time. When both the harmonic and temporal conditions supported a duple metre, musicians selected a duple metre 100% of the time while nonmusicians selected it 91.0% of the time. The higher percentage of responses for musicians and nonmusicians in favor of the duple metre reflects a bias in favor of common time metres in traditional Western music.

Apart from the quantitative differences between musicians and nonmusicians, musicians and nonmusicians may differ in an important way. When the temporal condition supported a duple metre and the harmonic condition supported a triple metre, nonmusicians selected a metre consistent with the harmonic condition most often (i.e., a triple metre -- 85.9% of the time). Similarly, when the temporal condition supported a triple metre and the harmonic condition supported a duple metre, nonmusicians selected a metre consistent with the harmonic condition (i.e., a duple metre -- 84.6% of the time). Musicians also selected on the basis of the harmonic conditions when the temporal and harmonic metres conflicted (92.3% of the time in favor of a triple harmonic based metre, and 94.9% in favor of a duple

Table 9

The Percentage of Metrical Choice under Different Conditions in Experiment 5

	Harmonic Conditions	
Metrical Choice	Triple Metre	Duple Metre
Temporal Condition : Triple Metre		
Musicians		
Triple Metre	91.0%	5.1%
Duple Metre	9.0%	94.9%
Nonmusicians		
Triple Metre	88.5%	15.4%
Duple Metre	11.5%	84.6%
Temporal Condition : Duple Metre		
Musicians		
Triple Metre	92.3%	---
Duple Metre	7.7%	100.0%
Nonmusicians		
Triple Metre	85.9%	9.0%
Duple Metre	14.1%	91.0%

harmonic-based metre). However, the slightly higher percentage of responses for the duple metre for musicians may be indicative of their use of knowledge of statistical properties in traditional Western music. Musicians, compared to nonmusicians, may have a greater tendency to select a common time over a complex time when they are presented with ambiguous stimuli, even though they more reliably select a metre on the basis of the harmonic condition.

On the basis of this data, it would appear that those features of a harmonic accompaniment which determine the perception of a metrical structure are very important in establishing the overall perceived metre of a musical presentation. Indeed, this study would seem to indicate that the harmonic accents of a musical presentation are more important than are the temporal accents defining the time-span within which subjects will perceive iterative pulses and from which they will then infer a metrical structure. Of course, unlike this experiment, most musical presentations would have various phenomenal cues working in cooperation to determine the metrical structure. Nevertheless, this experiment serves to show the salience of harmonic phenomenal accents, and thus adds support to the claim that phenomenal accents are not perceptually equivalent (Jones, 1992) and specifically, that harmonic phenomenal accents appear to be much more salient than temporal phenomenal accents in the determination of metrical structure.

Experiment 6

Method

Subjects. Twenty-four experimentally naive volunteers participated in this study. There were 12 musicians with a range of 6 - 10 years and a median of 7 years of formal music training. Their age ranged from 19 - 22 with a median age of 19 years. Of the twelve musicians, 6 were females and 6 were males. Twelve subjects were nonmusicians with an age range of 18 - 23 and a median of 20 years. Nine of these subjects were female and 3 were males. Subjects were recruited from introductory psychology classes and were given course credit for participating in the experiment.

Apparatus. The equipment employed in this experiment was the same as that used in Experiment 5.

Stimuli. The stimuli consisted of a melody line and corresponding harmonic accompaniment. There were two melodic conditions and two harmonic conditions. In the duple-metre melodic condition, a four-tone pattern was presented which consisted of Middle C followed by the E, G, and A immediately above it. This basic pattern was presented 18 times such that the pattern can be represented by the scheme CEGA-CEGA-CEGA-etc. In this way two melodic phenomenal accents combined to encourage adoption of a duple metre. The contour melodic phenomenal accent represented by a shift from an ascending pitch direction to a descending pitch direction at the transition point of A to C also represents a relatively large downward interval jump of 9 semitones. In addition to these two melodic phenomenal accents, there is also a melodic structural accent associated with this location in that there is a return to the tonic C. In the triple metre melodic condition, a three tone pattern was presented which consisted of Middle C followed by the E, and G immediately above it. This basic pattern was presented 24 times such that the pattern presented can be represented with the scheme CEG-CEG-CEG-etc. Just as in the other melodic condition, a triple metre was encouraged by a contour melodic phenomenal accent, a relatively large interval jump of 7 semitones, and a structural accent associated with the return to the tonic.

The duple metre harmonic condition consisted of four presentations of the same chord before it changed to a different chord. The cadence used in this experiment was C-G-C, or the I-V-I cadence in the key of C Major. Each chord presentation was set at a duration equal to the duration of a note in the melody. Thus the presentation of the harmonic accompaniment in this condition can be represented by the scheme C-C-C-C-G-G-G-G-C-C-C-C-etc. The triple metre harmonic condition consisted of six presentations of the same chord before it changed to a different chord. The presentation in this condition can be represented by the scheme C-C-C-C-C-C-G-G-G-G-G-G-etc.

The two harmonic conditions were orthogonally combined with the two melodic conditions resulting in four different conditions. In two of these four conditions, both the harmonic and melodic phenomenal accents supported the same metre. In the other two conditions the harmonic and temporal accents supported different metres. The tempo of each presentation was 135 quarter notes per minute with each tone and chord presentation equal to an eighth note, and the timbre was that of a synthesized piano.

Procedure. The procedure was the same as in Experiment 5.

Results and Discussion

An initial loglinear analysis was performed on the data, with three design and one response variable. The design variables were the musical Training of the subject (nonmusician or musician), the metre defined by the Melodic conditions (triple and duple metres), and the metre defined by the Harmonic conditions (triple and duple metre). The response variable was the metre chosen by the subject.

As in Experiment 5, the loglinear analysis was implemented with the loglinear module in Statistica. The results of the tests for partial and marginal association indicated the significant effects were: Choice (Partial $\chi^2(1) = 4.23$, $p < .05$; Marginal $\chi^2(1) = 4.23$, $p < .05$), Choice X Melodic condition (Partial $\chi^2(1) = 27.65$, $p < .001$; Marginal $\chi^2(1) = 18.49$, $p < .001$), Choice X Harmonic condition (Partial $\chi^2(1) = 408.74$, $p < .001$; Marginal $\chi^2(1) = 398.41$, $p < .001$), and finally, Choice X Harmonic condition X Training (Partial $\chi^2(1) = 206.31$, $p < .001$; Marginal $\chi^2(1) = 212.19$, $p < .001$).

The second stage of the analysis, the automatic stepwise model selection, indicated that the best fitting model included all variables: Choice X Melodic condition X Harmonic condition X Training $\chi^2(0, N = 1152) = 0$, $p = 1.00$).

Overall, there was a slight yet significant preference for a duple metre over a triple metre: 53.0% of the time, subjects chose a duple metre. The significant interaction of choice by melodic condition also showed evidence of this preference. When the melodic condition supported a duple metre, subjects selected a duple metre 59.4% of the time, and when the melodic condition supported a triple metre, subjects selected a triple metre 53.3% of the time. This pattern is indicative of a slight preference in favor of a metre consistent with the melodic condition, combined with an overall preference for a duple as opposed to a triple metre. The same pattern was evident in the significant interaction of choice by harmonic condition. When the harmonic condition supported a duple metre, subjects selected a duple metre 81.6% of the time, and when the harmonic condition supported a triple metre, subjects selected that metre 75.5% of the time. Again, this pattern is indicative of a strong preference in favor of a metre defined by the harmonic condition, combined with a slight preference for a duple over a triple metre. The three-way interaction of choice by harmonic condition by training indicated that musicians chose a metre on the basis of the harmonic condition much more than did nonmusicians. When the harmonic condition supported a triple metre, musicians selected a triple metre 94.8% of the time compared to nonmusicians who selected a triple metre 56.3% of the time. When the harmonic condition supported a duple metre, musicians selected a duple metre 96.2% of the time while nonmusicians selected a duple metre only 67.0% of the time. The slight preference for a duple metre was evident in both the musicians and nonmusicians responses (96.2% and 67.0% for duple metre compared to 94.8% and 56.3% for triple metre, for musicians and nonmusicians, respectively).

All of these significant effects are evident in Table 10 which shows the four way interaction of metrical choice by training by melodic and harmonic conditions. Musicians primarily selected a metre on the basis of the harmonic condition. When the harmonic condition was a triple metre, musicians selected a triple metre regardless of the melodic condition (95.8% and 93.8% for the melodic conditions of triple and duple metre, respectively). Similarly, when the harmonic condition supported a duple metre, musicians selected a duple metre regardless of the melodic condition (99.3% and 93.1% for the melodic conditions of a duple and a triple

Table 10

The Percentage of Metrical Choice under Different Conditions in Experiment 6

	Harmonic Conditions	
Metrical Choice	Triple Metre	Duple Metre
Melody Condition : Triple Metre		
Musicians		
Triple Metre	95.8%	6.9%
Duple Metre	4.2%	93.1%
Nonmusicians		
Triple Metre	71.5%	38.9%
Duple Metre	28.5%	61.1%
Melody Condition : Duple Metre		
Musicians		
Triple Metre	93.8%	0.7%
Duple Metre	6.3%	99.3%
Nonmusicians		
Triple Metre	41.0%	27.1%
Duple Metre	59.0%	72.9%

metre, respectively). The slight preference for duple metre over a triple metre is evident in the musicians' data when one notes the higher percentage of responses for a duple metre when it was supported by both the melodic and harmonic conditions (99.3%) compared to when both conditions supported a triple metre (95.8%).

This bias for a duple metre was more evident in the nonmusicians' pattern of responses. When both the harmonic and melodic conditions supported a duple metre, nonmusicians selected a duple metre 72.9% of the time. Similarly, when both conditions supported a triple metre, nonmusicians predominantly selected a triple metre (71.5% of the time). Although the slightly higher percentage of nonmusicians' responses for a duple metre (72.9%) compared to triple metre (71.5%) when both the melodic and harmonic conditions supported those metres can be interpreted in terms of the bias for a common time metre, this is even more evident when one examines the two conditions in which the melodic and harmonic supported metres are different. When the harmonic condition supported a duple metre and the melodic condition supported a triple metre, nonmusicians selected a duple metre 61.1% of the time. This can be interpreted in terms of a combination of a duple metre bias and a preference to select on the basis of the harmonic condition. It was only when the harmonic condition supported triple metre, and the melodic condition supported a duple metre that nonmusicians selected a metre inconsistent with the harmonic condition (59% of the time they selected a duple metre). This indicates that the bias for common time metres is greater for nonmusicians than for musicians, and is overcome only when both the melodic and harmonic conditions support a triple metre.

Experiment 7

Method

Subjects. Twenty-four experimentally naive volunteers participated in this study. There were 12 musicians with a range of 3 - 8 years and a mode of 7 years of formal musical training. Their age ranged from 19 - 23 with a mode age of 19 years. Of the twelve musicians, 7 were females and 5 were males. Twelve subjects were nonmusicians with an age range of 18 - 25 and a mode of 20 years. Six of these subjects were female and 6 were males. Subjects were recruited from introductory psychology classes and were given course credit for participating in the experiment.

Apparatus. The equipment employed in this experiment was the same as that used in the previous two experiments.

Stimuli. The stimuli in this experiment consisted of a melody with both melodic pitch and temporal phenomenal accents. There were two temporal conditions and two melodic pitch conditions. In the duple metre temporal condition, one basic pattern of a long duration tone followed by two short tones was presented 18 times such that the pattern can be represented by the scheme LSSLSSLSS-etc. In the triple metre temporal condition, a long duration tone followed by 4 short duration tones was presented 12 times such that the pattern presented can be represented with the scheme LSSSSLSSSSLSSSS-etc. In both the duple metre melodic pitch condition and the triple metre melodic pitch condition, the metrical structure was encouraged by the systematical occurrence of a melodic contour from ascending to descending, a relatively large interval jump, and a melodic structural accent associated with a return to the tonic note. Combining the two melodic conditions with the two temporal conditions resulted in four conditions, which can be seen represented in musical notation in Figure 13. In two of these four conditions, both the melodic and temporal phenomenal accents support the same metre (Example A -- duple metre, and Example D -- triple metre). In the other two conditions the melodic and temporal accents support different metres. The tempo of each presentation was 135 quarter notes per minute with the long-duration tone in the melody being equivalent to a quarter note and the short tones and chord presentations equal to eighth notes, and the timbre was that of a synthesized piano.

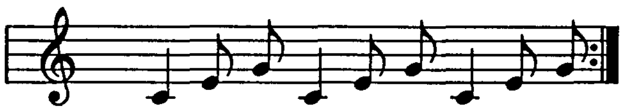
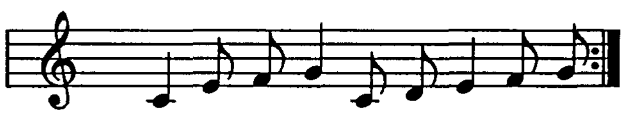
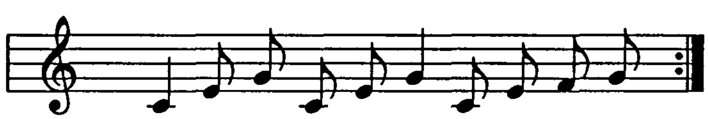
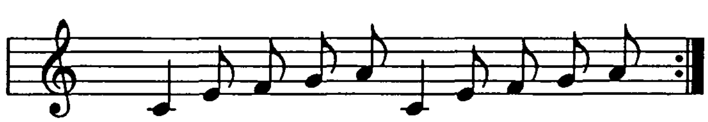
Procedure. The presentation of stimuli was the same as that of the previous two experiments.

Results and Discussion

An initial loglinear analysis was performed on the data, with three design and one response variable. The design variables were the musical Training of the subject (nonmusician or musician), the metre defined by the Melodic conditions (triple and duple metre), and the metre defined by the Temporal conditions (triple and duple metre). The response variable was the metrical condition chosen by the subject.

As in the previous experiments, the loglinear analysis was implemented with the loglinear module in Statistica. The results of the tests for partial and marginal association indicated the significant effects were: Choice (Partial $\chi^2(1) = 7.95$, $p = .005$; Marginal $\chi^2(1) = 7.95$, $p = .005$), Choice X Training (Partial $\chi^2(1) = 4.15$, $p < .05$; Marginal $\chi^2(1) = 3.56$, $p = .06$), Choice X Melodic condition (Partial $\chi^2(1) = 49.41$, $p < .001$; Marginal $\chi^2(1) = 43.84$, $p < .001$), Choice X Temporal condition

Figure 13. The Musical Notation of Stimuli used in Experiment 7.

A). Melodic and Temporal Duple Metres.

B). Melodic Triple and Temporal Duple Metres.

C). Melodic Duple and Temporal Triple Metres.

D). Melodic and Temporal Triple Metres.

(Partial $\chi^2(1) = 130.90$, $p < .001$; Marginal $\chi^2(1) = 125.05$, $p < .001$), and Choice XTemporal condition X Melodic condition (Partial $\chi^2(1) = 59.89$, $p < .001$; Marginal $\chi^2(1) = 53.74$, $p < .001$).

The second stage of the analysis, the automatic stepwise model selection, indicated that the best fitting model included all variables: Choice X Training X Melodic condition X Temporal condition, $\chi^2(0, N = 1152) = 0$, $p = 1.00$).

Overall, there was a slight yet significant preference for a duple metre over a triple metre: 54.2% of the time, subjects chose duple time. The significant interaction of choice by training indicated that the bias for a duple metre over a triple metre was stronger for musicians than it was for nonmusicians. Musicians selected a duple metre 56.9% compared to nonmusicians who chose a duple metre 51.4% of the time. The significant interaction of choice by melodic condition also showed evidence of this preference. When the melodic condition supported a duple metre, subjects selected a duple metre 63.9% of the time, and when the melodic condition supported a triple metre, subjects selected a triple metre 55.6% of the time. This pattern is indicative of a slight preference in favor of a metre consistent with the melodic condition combined with an overall preference for a duple metre as opposed to a triple metre. The same pattern was evident in the significant interaction of choice by temporal condition. When the temporal condition supported a duple metre, subjects selected a duple metre 70.5% of the time, and when the temporal condition supported a triple metre, subjects selected that metre 62.2% of the time.

The three-way interaction of choice by melodic condition by training indicated that musicians chose a metre on the basis of the melodic condition much more than did nonmusicians. When the melodic condition supported a triple metre, musicians selected a triple metre 63.2%. When the melodic condition supported a duple metre, musicians selected a duple metre 77.1% of the time. Again, the slight bias in favor of duple metre was evident in the musicians' pattern of responses. Nonmusicians, on the other hand, were qualitatively different. They did not show any indication of selecting on the basis of melodic information. If one considers the bias in favor of a duple metre, nonmusicians tended to select a triple and a duple metre independent of the melodic conditions. When the melodic condition supported a triple metre, nonmusicians selected a duple metre 52.1% of the time, and when the melodic condition supported a duple metre, nonmusicians selected a duple metre 50.7% of the time.

All of these significant effects are evident in Table 11, which shows the four-way interaction of metrical choice by training by melodic and temporal conditions. Nonmusicians primarily selected a metre on the basis of the temporal condition. When the temporal condition was a triple metre, nonmusicians selected a triple metre regardless of the melodic condition (61.1% and 70.8% for the melodic conditions of triple and duple metre, respectively). Similarly, when the temporal condition supported a duple metre, nonmusicians selected a duple metre regardless of the melodic condition (72.2% and 65.3% for the melodic conditions of duple and triple metre, respectively). The slight preference for a duple metre is evident in the nonmusicians' data when one notes the higher percentage of responses for a duple metre when it was supported by both the melodic and temporal conditions (72.2%) compared to triple metre responses when supported by both conditions (61.1%).

Table 11

The Percentage of Metrical Choice under Different Conditions in Experiment 7

	Melodic Conditions	
Metrical Choice	Triple Metre	Duple Metre
Temporal Condition : Triple Metre		
Musicians		
Triple Metre	86.1%	30.6%
Duple Metre	13.9%	69.4%
Nonmusicians		
Triple Metre	61.1%	70.8%
Duple Metre	38.9%	29.2%
Temporal Condition : Duple Metre		
Musicians		
Triple Metre	40.3%	15.3%
Duple Metre	59.7%	84.7%
Nonmusicians		
Triple Metre	34.7%	27.8%
Duple Metre	65.3%	72.2%

For the most part, and in direct contrast to nonmusicians, musicians showed evidence of selecting on the basis of both melodic and temporal information. When both the temporal and melodic conditions supported a duple metre, musicians selected a duple metre 84.7% of the time. Similarly, when both conditions supported a triple metre, musicians predominantly selected a triple metre (86.1% of the time). When the temporal condition supported a duple metre and the melodic condition supported a triple metre, musicians tended to select a metre on the basis of the temporal condition (i.e., a duple metre -- 59.7% of the time). However, when the temporal condition supported a triple metre and the melodic condition supported a duple metre, musicians tended to select on the basis of the melodic condition (i.e., they chose a duple metre 69.4% of the time). This pattern of responses can be interpreted in terms of a musicians' being able to use either melodic or temporal information when given an ambiguous musical passage. It appears that musicians tend to side with whichever component (melodic or temporal) supported a duple metre. This may be indicative of musicians relying on their knowledge of statistical relationships in traditional Western music when they are presented with ambiguous stimuli that can be interpreted in terms of two metres.

On the basis of this experiment, it would appear that in the absence of a harmonic accompaniment, both temporal and melodic features are used to determine the perception of a metrical structure. The relative strength of the melodic and temporal phenomenal accents was found to vary depending upon the musical training of the subjects and the conditions present. This experiment indirectly supports the salience of harmonic phenomenal accents. In comparison to Experiments 5 and 6, the temporal and melodic accents had a greater influence on the perception of a metrical structure in this experiment. This suggests that the harmonic features of the two previous experiments may have been so powerful a cue for the perception of metrical structure that subjects ignored temporal and melodic information when it conflicted with the harmonic accompaniment.

General Discussion and Summary

This series of experiments was developed to add further support to the hypothesis that phenomenal accents can be ordered according to a hierarchy of perceptual significance. Together, the three experiments show that of all the phenomenal accents tested, harmonic phenomenal accents were the most important in establishing a metre. In the two experiments in which a harmonic cue was pitted against a temporal or melodic accent, subjects reported hearing a metre that was based primarily on the harmony. In Experiment 6, the harmony was used more often than melodic accents to perceive a metrical structure, but the melodic variable was nevertheless found to be significant in determination of metrical choice. This suggested that the coincidence of melodic accents may be more salient in defining a time-span that is used to infer a metrical structure than is a temporal accent. This was confirmed in Experiment 7. The coincidence of melodic accents were used more often than temporal accents for metre determination. This experiment also offered indirect support for the saliency of harmonic features in metrical perception. This was the only experiment of the three in which a harmonic condition was not included, and under such conditions, both the temporal and melodic accents were used much more for perception of the metre than when a harmony was present.

In all three experiments a bias was found for a duple metre over a triple metre. In Experiments 5 and 6, when one of the experimental conditions included the manipulation of a harmonically supported metre, the bias was evident in a higher percentage of choice for a duple metre when both experimental conditions supported that metre compared to the percentage of choice for a triple metre when both experimental conditions supported a triple metre. In Experiment 7, while nonmusicians primarily selected a metre on the basis of the temporal condition, a slight preference for a duple metre was evident in the higher percentage of responses for a duple metre when it was supported by both the melodic and temporal conditions compared to responses for a triple metre when both conditions supported a triple metre. Musicians, on the other hand, tended to select a duple metre in all conditions except when both the temporal and melodic conditions supported a triple metre. This pattern of responses was interpreted in terms of the musicians being able to use either melodic or temporal information when given an ambiguous musical passage, and a tendency to select on the basis of whichever component (melodic or temporal) supported a duple metre. Musicians may have been relying of their knowledge of statistical relationships in traditional Western music when they were presented with the metrically ambiguous stimuli.

The data also allow us to draw some conclusions regarding accents weights and how they may combine to determine perceptually salient events in music. The fact that two melodic phenomenal accents were combined (both a contour change and a relatively large interval jump) with a melodic structural accent to serve as the melodic perceptual cue supports the argument that phenomenal accents are not equally weighted and that they do not combine in an additive fashion. If phenomenal accents are equally-weighted in terms of salience and combine in an additive fashion, then combining the two melodic phenomenal events with the melodic structural accent should have made the melodic cue more than twice as salient compared to either the harmonic or temporal accent, and it was not.

The differences between musicians and nonmusicians supports the conclusion that musicians use harmonic information much more often and reliably compared to nonmusicians. In contrast, nonmusicians are inclined to use temporal information more than musicians do to determine a metrical structure when harmonic cues are not available.

Finally, there is abundant evidence from other sources that temporal accents; specifically, long duration notes, initiate major metrical locations (Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1982, 1984; Longuet-Higgins & Steedman, 1971; Steedman, 1977). However, it appears that the time-spans which define in part the hierarchical pattern of iterative pulses will be harmonic accents if they are present. The data from this series of experiments in conjunction with that which was reported in Experiment 1 supports the contention that the cues which define the time-spans from which one perceives iterative pulses and infers a metrical structure need not be the same cues which are perceived to mark important metrical locations. In addition, the results support the role of harmonic-rhythm in traditional Western music theories of metre determination.

CHAPTER V

General Discussion and Conclusions

Implications for Empirical Investigations of Rhythmic Structure.

Generally speaking, the ideas presented in this thesis are not new. The ideas are in keeping with modern music-theoretical views. However, the conclusions drawn in this thesis differ from theoretical notions in at least one very important respect: The conclusions are based on empirical observation rather than theoretical constructs. Music theory is based on intuitions and insights into the structure of music which may or may not be verified empirically through the rigorous standards of scientific enquiry. Scientific testing such as that which has been done in this thesis allows for observations and the opportunity to test various aspects of rhythmic-theoretical views.

The research presented in this thesis addresses the perception of rhythmic structure. Although there has been to date a large body of literature on the perception of rhythmic structure, it has been fragmented with noticeable gaps for investigation of relevant variables. In particular, empirical investigations of the role of harmonic accents (e.g., harmonic-rhythm), have been lacking. It is possible to identify several reasons for the fragmentation.

One possible reason for the currently inadequate state of research on own right.

This thesis addresses some issues of debate and controversy in regards to research on phrase structure. There have been many different accounts of how structural and phenomenal accents combine to determine phrase structure. Some theorists (e.g., Jones, 1987; Benjamin, 1984; Berry, 1976) believe differences in phrase structure are determined by coincidences of equally salient temporal, melodic and harmonic events. Such views of accents could be referred to as being categorical; either an event is or is not an accent. These theories have a difficult time including harmonic structural accenting, since this construct is based upon the notion of differential accent strength or stability.

Other theorists believe that accents within a theoretical component category are not categorical but rather can differ in terms of accent strength or perceptual salience along a continuum. However, the theoretical component categories combine in an additive and equally weighted fashion to determine rhythmic structure (Palmer and Krumhansl, 1987a, 1987b). The experiments reported here indicated that accents based on changes in harmony are perceptually more salient than accents based on other musical events such as relatively long duration notes. Therefore, in contrast to Palmer and Krumhansl (1987a, 1987b), the two theoretical components of harmony and temporal aspects did not combine in an additive and equal fashion to determine the phrase structure of the music.

One can hypothesize as to the reasons for the discrepancy between the current findings and those of Palmer and Krumhansl (1987a, 1987b). Possible reasons may have to do with methodological differences between the two approaches or with a confound inherent in either approach. In the first instance, the current approach has been to ask subjects to directly select the rhythm pattern they have heard, while the methodological approach of Palmer and Krumhansl has been to employ a variation of Krumhansl's (1979) tone-profile technique which employs rating various probe-tones. Even given hypothetical evidence that subjects are inaccurate in directly selecting among rhythmic alternatives which pattern it is that they have heard, the methodology employed by Palmer and Krumhansl (1987a, 1987b) is much more problematic. Apart from the difficulties associated with the subjective nature of rating scales, this approach has the additional difficulty of having subjects abstractly quantify a phrase boundary.

In terms of a possible confound inherent in the two approaches, the possible role of familiarity effects on Palmer and Krumhansl's (1987a, 1987b) results was discussed at some length in the first chapter. Alternatively, it could be argued that the relative importance of the harmonic features in rhythm perception was inflated in these experiments. The argument is based on the premise that the melodic and temporal features were impoverished more than the harmonic features in the presented stimuli. Even though two chords being presented in an alternating fashion is impoverished by most standards, it is nevertheless a legitimate harmonic accompaniment. In contrast, a 'melody' without pitch information or with very little temporal variation does not sound very melodic. While this argument is valid, it fails to appreciate the reasons for the investigation, and begs the question that was being addressed: namely, that different phenomenal accents are not equally salient with any notion of accent strength being based upon mere coincidence. Rather, different phenomenal cues (e.g., a chord-change, and a long-held note) are differentially salient. To state that the harmonic accompaniment which consisted of a phenomenal event is more salient than a lone temporal accent is to support exactly what has been shown in this series of experiments.

The finding of significant differences in perceived structure as a function of musical experience is of importance. Both in terms of phrase structure and metrical structure, non-musicians were found to employ inferred metrical accents in determining rhythmic structure much more than were musicians. Musicians in contrast seem to employ the use of chord changes to determine the perceived rhythmic structure much more than did non-musicians. However, in all experiments, and for all subjects, regardless of musical training, harmonic phenomenal accents (in particular those associated with a change of chord) served as effective and functionally salient events in the music. They were used not only to determine rhythm patterns but also as a basis for metre inference.

It was somewhat surprising that the salience of a chord change was found not to be mediated by its structural significance in a tonal context. It was only when nondiatonic and unusual chord progressions were used that a qualitatively different pattern of rhythmic choices was found. Under these conditions, there was no evidence of selecting rhythms on the basis of accent coincidence or of an inferred metre affecting the perceived musical sequence. It was argued in Chapter 4 that any stability created by a movement towards the tonic chord was countered by the tension created when one moves away from the tonic; an argument which was in part supported by the results found in the series of experiments. Another possibility is that the nature of this task is such that evidence supporting structural accenting is not readily apparent. Reported rhythms were defined by the chord change locations and thus the actual progressions played a very little role in rhythmic determinations. If the choice of rhythms to be selected fell across chord change locations rather than within them, the structural accenting of progressions might be a more salient feature of the presentation.

Preliminary investigations have supported this hypothesis. When two chords are alternately presented with durations equal to one half the duration of the rhythm pattern (i.e., the two chords together equal the duration of the rhythm pattern), the reported rhythm is very much dependent on which of the two chords is most stable. The more stable of the two, within a particular tonal context, is perceived to terminate the rhythm pattern. This paradigm therefore offers an alternative measure of harmonic stability to the traditional rating scale methodology. In addition other preliminary research has varied the durations of the two chord presentations, pitting harmonic-rhythmic features against the stability of the chords progressions. Thus, although much research remains to be done, the procedures employed and reported in this thesis provide a powerful technique for investigating many harmonic-rhythm features.

Apart from the relevance of the research presented in this thesis for phrase and harmonic structure research, it is also important for current research and theoretical work being conducted on the perception of metrical structure. The debate about whether metre is an inferred mental construct or is communicated via phenomenal accents is clearly addressed. The research presented here supports both positions. Experiments 1 through 4 clearly support the view that in the absence of metrical cues such as subtle changes in the dynamics and temporal features of metrically significant locations, and/or the frequency of note onsets and offsets to similarly mark significant locations, individual listeners will perceive a metrical structure. Furthermore, the inferred metrical structure interacted with the location of chord changes to affect the perceived rhythmic structure.

The final series of experiments demonstrated that with an ambiguous musical presentation, subjects will infer a metrical structure based primarily upon the timing of harmonic events (e.g., chord changes) if this information is available. In its absence, subjects will use the temporal aspects of the presentation (i.e., the timing of note onsets and offsets) and melodic information such as the timing of regular melodic accents to infer a metrical structure.

Finally, the research serves as an initial attempt to draw the research tradition of harmonic structure into theoretical works on phrase structure and metrical structure. This is novel in two related respects. First in terms of phrase structure, although research has dealt with the notion of harmonic structure and the concept of stability and it has long been recognized that harmonic structural accenting is of importance in phrase structure, it opens up a new role for harmony in terms of it serving as a phenomenal event. The simple occurrence of a harmony change, regardless of its structural accent strength can serve as a cue for phrase structure. In the second respect, the use of chord changes as phenomenal events has lead to a means to investigate the music-theoretical notion of harmonic rhythm. Such an endeavor has only rarely been undertaken (e.g., Smith & Cuddy, 1989).

REFERENCES

Agmon, E. (1990). Music theory as cognitive science: Some conceptual and methodological issues. Music Perception, 7, 285-308.

Bartlett, J.C., & Dowling, W.J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. Journal of Experimental Psychology: Human Perception and Performance, 6, 501-516.

Benjamin, W. (1984). A theory of musical meter. Music Perception, 1, 355-413.

Berry, W. (1976). Structural functions in music. Englewood Cliffs, NJ: Prentice-Hall.

Bharucha, J., & Krumhansl, C.L. (1983). The representation of harmonic structure in music:Hierarchies of stability as a function of context. Cognition, 13, 63-102.

Boltz, M. (1989). Rhythm and "good endings": Effects of temporal structure on tonality judgments. Perception and Psychophysics, 46, 9-17.

Boltz, M., & Jones, M.R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? Cognitive Psychology, 18, 389-431.

Boretz, B. (1971). In quest of the rhythmic genius. Perspectives of New Music, 9-10, 149-155.

Bray, K.I., Snell, D.B., & Peters, R.M. (1967). For Young musicians (Vol.2). Waterloo: Waterloo Music Company Ltd.

Brown, H., & Butler, D. (1981). Diatonic trichords as minimal tonal cue-cells. In Theory Only, 5, 37-55.

Butler, D. (1983). The initial identification of tonal centers in music. Paper presented to the NATO Conference on the Acquistion of Symbolic Skills, University of Keele, England, July 6-9, 1982. In J. Sloboda & D. Rogers (Eds.), Acquisition of symbolic skills. New York: Plenum Press.

Butler, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. Music Perception, 6, 219-242.

Clarke, E.F. (1985). Structure and expression in rhythmic performance. In P.Howell, I. Cross, & R. West (Eds.), Musical structure and cognition. pp. 209-236.

Cook, N. (1978). A guide to musical analysis. London: Dent.

Cooper, G.W., & Meyer, L.B. (1960). The rhythmic structure of music. Chicago: Chicago University Press.

Cooper, W.E., Paccia, J.M. & LaPointe, S.G. (1978). Hierarchical coding in speech timing. Cognitive Psychology, 10, 154-177.

Cuddy, L.L., Cohen, A.J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. Canadian Journal of Psychology, 33, 148-157.

Dowling, W.J. (1968). Rhythmic fission and the perceptual organization of tone sequences. Unpublished Ph.D. dissertation, Harvard University.

Dowling, W.J. (1973). Rhythmic groups and subjective chunks in memory for melodies. Perception and Psychophysics, 14, 37-40.

Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. Psychological Review, 85, 341-354.

Dowling, W.J. (1982). Melodic information processing and its development. In D. Deutsch (Ed.), The psychology of music. New York: Academic Press.

Dowling, W.J., & Harwood D.L. (1986). Music Cognition. Toronto: Academic Press.

Dowling, W.J., & Fujitani, D.S. (1971). Contour, interval, and pitch recognition in memory for melodies. Journal of the Acoustical Society of America, 49, 524-531.

Essens, P.J., & Povel, D.J. (1985). Metrical and nonmetrical representations of temporal patterns. Perception and Psychophysics, 37, 1-7.

Gabrielsson, A. (1974). Performance of rhythm patterns. Scandinavian Journal of Psychology, 15, 63-72.

Gabrielsson, A. (1986). Rhythm in music. In J.R. Evans & M. Clynes (Eds.), Rhythm in psychological, linguistic, and musical processes. (pp. 131-167). Springfield, IL: Charles C. Thomas.

Gabrielsson, A. (1987). Once again: The theme from Mozart's piano Sonata in A Major (K.331). In A. Gabrielsson (Ed.), Action and perception in rhythm and music. Stockholm: Royal Swedish Academy of Music. pp.81-104.

Garner, W. (1974). The processing of information and structure. Hillsdale, NJ: Erlbaum.

Halpern, A.R., & Bower, G.H. (1982). Musical expertise and melodic structure in memory for musical notation. American Journal of Psychology, 95, 31-50.

Handel, S. (1973). Temporal segmentation of repeating auditory patterns. Journal of Experimental Psychology, 101, 46-54.

Harris, J.D. (1963). Loudness discrimination. Journal of Speech and Hearing Disorders Monograph Supplement, 11, 1-63.

Hughes, M. (1977). A quantitative analysis. In M. Yeston (Ed.) Readings in Schenker Analysis and Other Approaches. (pp. 144-164), New Haven, Conn.: Yale University Press.

Idson, W.L., & Massaro, D.W. (1978). A bidimensional model of pitch in the recognition of melodies. Perception and Psychophysics, 24, 551-556.

Jones, M.R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. Psychological Review, 83, 323-355.

Jones, M.R. (1981a). A tutorial on some issues and methods in serial pattern research. Perception and Psychophysics, 30, 492-504.

Jones, M.R. (1981b). Music as a stimulus for psychological motion: Pt. I. Some determinants of expectancies. Psychomusicology, 1, 34-51.

Jones, M.R. (1982). Music as a stimulus for psychological motion: Pt. II. An expectancy model. Psychomusicology, 2, 1-13.

Jones, M.R. (1984). The patterning of time and its effects on perceiving. Annals of the New York Academy of Sciences, 423, 158-167.

Jones, M.R. (1987). Dynamic pattern structure in music: Recent theory and research. Perception and Psychophysics, 41, 621-634.

Jones, M.R. (1990). Musical events and models of musical time. In R.A. Block (Ed.), Cognitive Models of Psychological Time. Hillsdale, NJ: Erlbaum.

Jones, M.R. (1992, February). The structure and function of accent patterns. Paper presented at the Second International Conference on Music Perception and Cognition, University of California, Los Angeles, CA.

Jones, M.R., & Boltz, M. (1989). Dynamic attending and responses to time. Psychological Review, 96, 459-491.

Jones, M.R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. Perception and Psychophysics, 32, 211-218.

Jones, M.R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. Journal of Experimental Psychology: Human Perception and Performance, 7, 1059-1073.

Jones, M.R., Summerell, L., & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. Quarterly Journal of Experimental Psychology, 39, 89-121.

Kallman, H.J., & Massaro, D.W. (1979). Tone chroma is functional in melody recognition. Perception and Psychophysics, 26, 32-36.

Knopoff, L. & Hutchinson, W. (1983). Entropy as a measure of style: The influence of sample length. Journal of Music Theory, 27, 75-97.

Krumhansl, C.L. (1979). The psychological representation of musical pitch in a tonal context. Cognitive Psychology, 11, 346-374.

Krumhansl, C.L. (1985). Perceiving tonal structure in music. American Scientist, 73, 371-377.

Krumhansl, C.L. (1990). Tonal hierarchies and rare intervals in music cognition. Music Perception, 7, 309-324.

Krumhansl, C.L., Bharucha, J.J., & Castellano, M.A. (1982). Key distance effects on perceived harmonic structure in music. Perception and Psychophysics, 32, 96-108.

Krumhansl, C.L., Bharucha, J.J., & Kessler, E.J. (1982). Perceived harmonic structure of chords in three related musical keys. Journal of Experimental Psychology:Human Perception and Performance, 8, 24-36.

Krumhansl, C.L., & Castellano, M.A. (1983). Dynamic processes in music perception. Memory and Cognition, 11, 325-334.

Krumhansl, C.L. & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. Psychological Review, 89, 334-368.

Krumhansl, C.L., & Shepard, R.N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. Journal of Experimental Psychology:Human Perception and Performance, 5, 579-594.

Lee, C.S. (1985). The rhythmic interprtation of simple musical sequences: Towards a perceptual model. In P. Howlell, I. Cross & R. West (Eds.) Musical Structure and Cognition. London: Academic Press. pp. 53-69.

Lehiste, I. (1975). The perception of duration within sequences of four intervals. Paper presented at the 8th International Congress of Phonetic Sciences, Leeds, England.

Lerdahl, F., & Jackendoff, R. (1983). A generative theory of tonal music. Cambridge, MA: MIT Press.

Longuet-Higgins, H.C., & Lee, C.S. (1982). The perception of musical rhythms. Perception, 11, 115-128.

Longuet-Higgins, H.C., & Lee, C.S. (1984). The rhythmic interpretation of monophonic music. Music Perception, 10, 424-441.

Longuet-Higgins, H.C., & Steedman, M.J. (1971). On interpreting Bach. In B. Meltzer & D. Michie (Eds.) Machine Intelligence, 6. Edinburgh: Edinburgh University Press.

Massaro, D.W., Kallman, H.J., & Kelley, J.L. (1980). The role of tone height, melodic contour and tone chroma in melody recognition. Journal of Experimental Psychology: Human Learning and Memory, 6, 77-90.

Merriam, A.P. (1964). The anthropology of music. evanston, IL: Northwestern University Press.

Meyer, L. (1956). Emotion and Meaning in Music. Chicago: University of Chicago Press.

Miller, G.A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. Journal of the Acoustical Society of America, 191, 609-619.

Moore, B.C.J. (1989). An Introduction to the Psychology of Hearing, (3rd ed.). Toronto: Academic Press.

Oller, D. (1973). The effect of position in utterance on speech segment duration in English. Journal of the Acoustical Society of America, 54, 1235-1247.

Palmer, C. (1989). Mapping musical thought to musical performance. Journal of Experimental Psychology: Human Perception and Performance, 15, 331-346.

Palmer, C., & Krumhansl, C.L. (1987a). Independent temporal and pitch structures in determination of musical phrases. Journal of Experimental Psychology: Human Perception and Performance, 13, 116-126.

Palmer, C. & Krumhansl, C.L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. Perception and Psychophysics, 41, 505-518.

Palmer, C. & Krumhansl, C.L. (1990). Mental representations for musical meter. Journal of Experimental Psychology: Human Perception and Performance, 16, 728-741.

Piston, W. (1948). Harmony (rev. ed). New York: Norton.

Povel, D.J. (1981). Internal representation of simple temporal patterns. Journal of Experimental Psychology: Human Perception and Performance, 7, 3-18.

Povel, D.J., & Essens, P. (1985). Perception of temporal patterns. Music Perception, 2, 411-440.

Povel, D.J., & Okkerman, H. (1981). Accents in equitone sequences. Perception and Psychophysics, 30, 565-572.

Restle, F. (1972). Serial patterns: The role of phrasing. Journal of Experimental Psychology, 93, 385-390.

Rosner, B.R., & Meyer, L.B. (1982). Melodic processes and the perception of music. In D. Deutsch (Ed.) The Psychology of Music. New York: Academic Press.

Royer, F.L. & Garner, W.R. (1966). Response uncertainty and perceptual difficulty of auditory temporal patterns. Perception and Psychophysics, 1, 41-47.

Royer, F.L. & Garner, W.R. (1970). Perceptual organization of nine-element auditory temporal patterns. Perception and Psychophysics, 7, 115-120.

Schachter, C. (1980). Rhythm and linear analysis: A preliminary study. In F. Salzer and C. Schachter (Eds.), The music forum, 4. New York: Columbia University Press.

Scott, D. (1982). Duration as a cue to the perception of a phrase boundary. Journal of the Acoustical Society of America, 71, 996-1007.

Seashore, H.G. (1937). An objective analysis of artistic singing. In C.E. Seashore (Ed.), Objective analysis of musical performance: University of Iowa studies in the psychology of music, 4, (pp. 12-157). Iowa City: University of Iowa Press.

Simon, H.A. (1968). Perception du pattern musical par auditeur. Sciences de l'Art, 5, 28-34.

Sloboda, J.A. (1983). The communication of musical metre in piano performance. Quarterly Journal of Experimental Psychology, 35, 377-396.

Smith K.C., & Cuddy, L.L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. Journal of Experimental Psychology: Human Perception and Performance, 15, 457-471.

Stark, M.E. (1988). The effect of musical experience on the perception of triads. (Unpublished Ph.D. dissertation), McMaster University.

Steedman, M.J. (1977). The perception of musical rhythm and metre. Perception, 6, 555-569.

Thomanssen, J.M. (1982). Melodic accents: Experiments and a tentative model. Journal of Acoustical Society of America, 71, 1596-1605.

Vos, P. (1977). Temporal duration factors in the perception of auditory rhythmic patterns. Scientific Aesthetics, 1, 183-199.

Vos, P., & Handel, S. (1987). Playing triplets: facts and preferences. In A. Gabrielsson (Ed.), Action and perception in rhythm and music (pp. 35-47). Stockholm: The Royal Swedish Academy of Music Publications.

Watkins, A.J., & Dyson, M.C. (1985). On the perceptual organization of tone sequences and melodies. In P.Howell, I.Cross, & R. West (Eds.), Musical Structure and Cognition. London: Academic Press. pp. 71-120.

Webster (1984). Webster's New World Dictionary: Second College Edition. New York: Simon & Schuster, Inc.

Welker, R.L. (1982). Abstraction of themes from melodic variations. Journal of Experimental Psychology: Human Perception and Performance, 8, 435-447.

Woodrow, H. (1909). A quantitative study of rhythm. Archives of Psychology, 14, 1-66.

Yeston, M. (1976). The stratification of musical rhythm. New Haven, CT: Yale University Press.

Youngblood, J.E. (1958). Style as information. Journal of Music Theory, 2, 24-35.