RATING CADENCE STABILITY

RATING CADENCE STABILITY:
THE EFFECTS OF CHORD STRUCTURE, TONAL CONTEXT, AND MUSICAL TRAINING

By

MARGARET ELIZABETH WEISER

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

September 1990

DOCTOR OF PHILOSOPHY (1990) McMASTER UNIVERSITY
(Psychology)

TITLE: Rating Cadence Stability: The Effects of Chord Structure, Tonal Context, and Musical Training.

AUTHOR: Margaret Elizabeth Weiser, B.A. (Guelph)

SUPERVISORS: Professor J. R. Platt and Professor R. Racine

NUMBER OF PAGES: vii, 139

ABSTRACT

Cadences are orderly progressions of chords which occur in classical and contemporary Western music. They can serve as anchoring points for the perception of musical key and of tonality. The rules governing the structure and usage of cadences have been set forth in music theory. In a series of experiments, listeners were able to rate the stability of simple two-chord cadences without explicit knowledge of music theoretical concepts.

The stability ratings obtained for the cadences presented in these experiments were affected by the listener's musical training (inexperienced or formally trained), by the cadence type (chord progression moving toward or away from the tonic), by the position of the root (in the lowest or uppermost note position of chord), by the direction of cadence resolution (upward or downward pitch change), and by the tonal context. The tonal context was an ascending or descending diatonic scale in the key of G-major or C-major played before each cadence trial.

Two hypothetical listening strategies were introduced as possible ways of perceiving musical cadences. Melody-tracking was defined as a simple analytic listening strategy

which focused on the soprano voice of each chord in a cadence. Voice-tracking was defined as a flexible strategy which allowed the listener to focus on either the soprano or the bass voice in the triad sequence.

Musically trained listeners' ratings generally conformed to the voice-tracking strategy. Their ratings followed conventional standards consistently and accurately. They rated G-C cadences most stable in a C-major context, and C-G cadences most stable in a G-major context. Untrained listeners' ratings did not consistently show this effect of key context. Their ratings generally conformed to the melody-tracking strategy. They considered downward resolution of a cadence more stable than upward resolution, and they gave higher stability ratings to cadences when the soprano voice instead of the bass voice led the cadence.

All listeners tended to rate plagal cadences as more stable than imperfect cadences, and perfect cadences as more stable than other cadences. They also generally gave higher stability ratings to soprano-led cadences than to bass-led cadences. The musically untrained listeners were able to assign meaningful stability ratings to cadences, despite their lack of musical terms. The trained listeners appeared to approach the task in a different way. Over the course of formal musical training, trained listeners may have learned to use a more flexible strategy while maintaining a high level of accuracy and consistency in their task performance.

ACKNOWLEDGEMENTS

The completion of this dissertation leaves me indebted to many individuals. Above all, I wish to express my thanks to my co-supervisors, Professors John Platt and Ron Racine, as well as to past and present members of my thesis committee at McMaster, Professors Lorraine Allan and Lee Brooks, for their encouragement and curiosity.

The experiments reported here could not have been performed without the programming talents of Gary Weatherill, who made it all possible.

I also wish to thank my fellow McMaster alumni, among them Marianna Stark, April Takeda, and Eleni Hapidou, for their friendship and support.

And finally, my thanks to Anne Weiser, for her musical expertise, and to Anthony Kerby, for his critical eye.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

PREFACE

Music presents an attractive focus for psychological research for several reasons. First, music in its various forms is apparently universal in human culture. The musical styles of different cultures offer an extensive field of study, and volunteers for empirical studies are usually plentiful. Secondly, both the performance and the understanding of music may be considered specialized skills (Sloboda, 1985). However, musicianship cannot easily be ascribed to a specific and isolated talent. It is probably more advantageous to consider musical behaviour as a hierarchy of integrated skills (Seashore, 1938; Frances, 1984) than as a single ability.

Musical skills do not appear to be strongly related to common cognitive or perceptual skills. While many people are attracted to music as a leisure activity, few excel in musical performance or appreciation. The display of musical excellence in some mentally retarded or autistic persons testifies to the inability of conventional intelligence measures to predict or account for musical talent (Sacks, 1987; Charness, Clifton, and MacDonald, 1988; Lucci, Fein, Holevas, and Kaplan, 1988). It is also easy to find normal individuals who are totally unskilled in musical performance skills, but are neither intellectually nor physically compromised according to any standard measure. This double dissociation of musical skills and intelligence underscores the uncommon qualities of musical abilities.

There are many cases in which specific musical skills can be dissociated from other cognitive and motor skills. This is evident in the preservation of particular musical abilities, such as the ability to recognize a melody or use musical notation, in trained musicians who have sustained brain damage resulting in concurrent aphasia, alexia and agraphia (Schweiger, 1985; Grossman, Shapiro and Gardner, 1981; Zatorre, 1985; Shapiro, Grossman and Gardner, 1986; and Sacks, 1987). The apparently independent quality of certain musical skills calls for an examination of unconventional strategies in cognitive processing (Sidtis and Bryden, 1978). Such strategies may conceivably prove useful in the assessment and remediation of gross cognitive deficits.

Thus, a third reason for the study of musical abilities is an examination of the way perceptual, cognitive and motor processes interact during the performance of musical skills. Each unique combination of these processes could be conceptualized as a strategy. Musical tasks may be used to assess the recruitment and cooperation of perceptual, cognitive, and motor processes according to the task demands in various contexts. Such tasks also permit us to observe the transfer of strategies across specific task contexts and sensory modalities.

Finally, the question whether musical meaning and structure are integrated with or are parallel to the structure of language and speech offers a challenging focus. Some theorists have argued for a strong linguistic analogy, stressing the supremacy of grammatical structure. In their view, musical works can be interpreted to convey a structured and meaningful message (Meyer, 1956), much like the meanings we obtain from verbal utterances. While spoken language may not be a necessary prerequisite for an individual's acculturation to elementary musical structures (Hargreaves, 1986), more advanced musical understanding may to some degree be verbally mediated.

Other theorists have insisted that a work of music must be considered on its own terms, and within a separate framework (Gardner, 1973; Langer, 1967). The latter view points out that there are few similarities between language and music, arguing that there are no elements of musical structure which correspond to linguistic concepts such as noun, verb, or preposition (Lerdahl and Jackendoff, 1985).

But if there are no "nouns" and "verbs" in music, how do we understand musical works as meaningful? Are there other basic elements in music? If such elements exist, they may be perceived by both musically skilled and unskilled listeners. A specialized vocabulary may not be necessary. In order to investigate this question systematically, we required a simple task that could easily be administered to a number of participants. Only simple musical fragments were used, to permit the collection of a large number of ratings from the participants.

The experiments introduced in this dissertation address the perceptions reported by listeners of widely differing musical background who are presented with simple musical patterns. The musical meaning of these stimuli may differ according to the listener's focus of attention. This paradigm allows the identification and description of different listening strategies according to the individual focus. Using an identified listening strategy may affect the musical meaning attributed to fragments of music, or even to an entire musical work.

To this end, a paradigm was developed which allowed us to systematically manipulate tonal relationships. Listeners heard simple chord progressions, and they were asked to rate their impression of harmonic resolution. The progressions were presented in different harmonic contexts, and the listeners' ratings were compared. This paradigm permitted us to quantify the listeners' perceptions of stability.

The first chapter of this dissertation will introduce the foundations of musical meaning underlying this paradigm. We will begin with a consideration of the basis of musical sound, first discussing general aspects of sound, and continuing with a consideration of the distinctive qualities of musical tones. Following this, the structural aspects of music will be examined, with a focus on the function and role of harmony. Harmonic progressions, or cadences, will be discussed in a separate section addressing their attributes and typical roles in music.

Turning to the listener, we continue with a brief review of relevant issues in research on the cerebral lateralization of musical skills. This discussion leads to an examination of possible listening strategies. The hypothetical strategies are introduced within the framework of a proposed three stage model of listening skills. Finally, an experimental paradigm is suggested in order to test the predictions generated by the listening strategy model.

The second chapter of the dissertation consists of a short statistical excursion to introduce the area of cluster analysis. The main assumptions and applications of cluster analysis are outlined, and various techniques are contrasted.

Chapters Three, Four and Five each present an experiment, including sections on methodology and results. This series of experiments aims to examine the ratings given by listeners to cadences in different contexts.

The dissertation concludes with Chapter Six, which discusses the outcome of the series of experiments and assesses the impact of the reported results on the listening strategy model.

CHAPTER ONE

FOUNDATIONS OF MUSICAL MEANING

Psychological Studies of Music

Given the task of adequately describing our experience of a work of music, one might first try to identify and define a series of acoustic events, and then discuss our psychological response to these events. The empirical discipline of psychoacoustics uses such an approach first to determine the characteristics of the audible stimuli presented to the ear, and then to explain our perception of these sounds as a result of the biological functioning of the auditory system (Moore, 1982).

Another approach to this imaginary task would be to consider music as a cultural product, and to describe the defining attributes of the particular style or composer from which this product originates. Such a description may permit us to explore the unique qualities of the music we hear. The nature of musical works within a tradition has been comprehensively addressed by numerous musicological studies (Rowell, 1983; Meyer, 1967; Kerman, 1985).

However, neither of these two approaches presents a viable route for the accumulation of experimental evidence regarding the process by which individual listeners are able to understand particular musical meanings. The acoustic analysis of sound sequences contributes to our knowledge of basic perceptual mechanisms, while musicology seeks to understand both universal and particular concepts of musical structure in various cultural traditions. Yet neither of these perspectives, psychoacoustic or musicological, focuses on the specific abilities or training of the individual listener. These behavioral issues can be most adequately addressed by experimental psychological studies of music, which allow us to observe the listener's response to controlled musical stimuli.

In contrast, a musicological analysis is the product of the theoretical judgements of a particular musicologist. These judgements cannot easily be separated from that individual's theoretical perspective. However, it is possible to obtain informative descriptions of salient regularities and incongruities within the musical excerpts presented to the listener. Documentation of these (ir)regularities permits a close examination of the musical structure in a given excerpt, but does not provide answers to questions regarding the process preceding the listener's response. The judgements obtained from musicological analyses are not in a form which allows detailed testing of the causal relations among the structural regularities described.

Psychological studies of music aim to investigate numerous empirically-based questions. How do we account for differences among listeners with varying degrees of expertise? Do trained listeners hear chords in a different way than do untrained listeners? Is the trained listener's experience of music obtained through an automatic set of responses to specific contexts, or through a flexible set of skills? Experimental psychology offers the techniques to study questions such as these. In many cases, musicological descriptions have already documented the nature of the regularities we observe, and we would be surprised by and perhaps even suspicious of results that diverge sharply from the music-theoretical framework.

Within the field of the experimental psychology of music, our primary areas of concern are as follows. We can use experimental tasks to examine the contribution of context-dependent factors such as listening habits, exposure to music, and amount of training. We can also assess the influence of specific parts of the musical stimulus by manipulating some of the components of the sound, such as the number of tones played, their pitch and their duration. Finally, we can design experiments to evaluate and refine our hypotheses about the causal relationships between musical structure and cognitive-perceptual elements and about the sources of individual variations.

We will not assume that musical abilities - or, indeed, any related cognitive or perceptual abilities - are equally distributed across the subject population. Musical ability is understood here as the level of proficiency that an individual displays on standardized musical tasks, such as reproduction of a melody.

An individual's abilities in the perception and performance of standardized musical tasks cannot be exclusively accounted for by the quantity of musical training that the individual has received. Nevertheless, our understanding of the individual listener's capabilities can be guided by our knowledge of the listener's musical training. In order for an individual to excel, any genetic predisposition to perform a given skill must, to be observable, be accompanied by some amount of diligent practice (Judd, 1988). One would expect, therefore, to find that an increase in practice is reflected in an increase of manifest skill.

The highly-practiced listener may be understood as an active interpreter who utilises both salient and less obvious cues in an attempt to gain information from sensory input (Neisser, 1967). We are predisposed to detect patterns in the flow of information with which we are confronted. When a monotonous, even sequence of tones is played, the listener tends to group the sounds together to form distinctive rhythmic patterns (Fraisse, 1982). Our perceptions of music, like our perceptions of other events in the world, can be shaped by our cognitions of these events and of the surrounding context. For example, a funeral dirge may appear comical or distressing when played at a wedding.

Our cognitions in turn are affected by our understanding of preceding events, and by our expectations of events to follow (Jones, 1981 and 1982). In twentieth-century Western culture, certain sequences of musical tones are usually associated with connotations of stability and closure. After the meaning of a certain sequence has become firmly established, it may be very difficult to learn to hear these sequences in a way different from that to which we have grown accustomed. Once one has learned to hear a melody in a polyphonic work of music, it is not easy to ignore the known melody and to listen for another, unfamiliar one. What music means to us depends largely on what we have learned to expect to hear, and on our ability to learn new expectations in adaptation to changing patterns and diverse contexts. Our listening experience is often biased by our tendency to hear music only in terms of what we have learned to listen for (Copland, 1939).

When someone learns to play or listen to music, we can, by observation, witness the gradual integration of cognitive strategies with the ability to recognize perceptual patterns (Sloboda, 1985). After becoming familiar with various styles of music, the listener learns to listen for the sounds and sound patterns typical of that tradition.

When someone learns to play an instrument as well, the acquired motor skills will have an effect on how the person listens to music. For example, when listening to a work of music, a percussionist will tend to focus on the rhythmic line of the drums rather than on the melody played by a solo instrument. By comparison, an untutored listener will probably be unable to identify or closely follow the contribution of the various percussion instruments.

The Basis of Musical Sound

What can we hear in music? Basically, any sound can vary according to four simple characteristics of a sound wave: frequency, amplitude, duration and form. The corresponding psychological attributes are pitch, loudness, time, and timbre. Rhythm, volume, and consonance are more complex attributes of sound that can also affect our perception of a tone. With respect to the distinction between music and other sounds, there is no generally accepted definition for music, usually considered a pleasant stimulus, or for noise, usually considered noxious. As modern composers (Copland, 1957) have pointed out, "noise" is a term reserved for sounds we do not wish to hear, while "music" is used for those we choose to listen to. The nature of the actual sounds that are assigned to these two categories varies according to task and tradition.

In the simplest case, the pitch or perceived highness of a musical sound is determined by the frequency of the audible sound waves. A greater frequency of sound waves gives rise to a sensation of higher pitch. Most musical sounds are complex tones whose waveform can be analyzed into a fundamental frequency and a number of higher-frequency components, or harmonics. The fundamental generally dominates our sensation of pitch, so that we have the impression of hearing a single musical tone (Wood, 1975).

The intensity of a tone, psychologically experienced as loudness of the sound, is measured by the amplitude of the sound wave. Greater amplitudes generally result in the perception of a louder sound.

The experienced extent of a tone in time depends on the physical duration of the sound. However, other factors, such as the rhythm and harmony of the musical pattern, can also affect our perception of a tone's duration (Fraisse, 1982).

Timbre, or tone quality, is partly produced by the interaction of the harmonic partials in a complex tone. The fundamental, or first partial, corresponds to the perceived pitch of a harmonic complex tone. It is the profile of the other partials in the amplitude envelope of the musical sound which gives rise to our perception of timbre (Moore, 1982). The first five higher-order partials of the harmonic series can be detected if one listens closely. The frequencies of the partials are in an ascending series in the ratios 1:2:3:4:5:6 etc. (Wood, 1975). The second partial is the octave above the fundamental and can be most easily heard. The upper partials are typically more difficult to discriminate due to their decreasing intensity, as well as their tendency to fall within a critical bandwidth. The critical band refers to the range of frequencies capable of masking a tone (Moore, 1982). The relative intensity and representation of the partials in a complex tone defines the characteristic timbre of various musical instruments.

Thus, the unique timbre of a particular instrument is affected by the strength of the separate frequency components, which are present in varying degrees of intensity and for varying durations over the course of a musical sound. But not all of the tones we hear are in fact physically present. When the musical tone is composed of a harmonic series based on a missing fundamental, the listener will report hearing that fundamental, even though it is not physically present. Cortical processing enriches our perception of tones, allowing us the impression that we hear the bass in a popular song even on a tinny pocket radio with an inadequate speaker which is actually physically incapable of producing the low-pitched tones we perceive.

The dynamic, constructive view of auditory perception has been explored by Terhardt (1978, 1982a, 1982b), who describes two different pitch percepts which can be elicited simultaneously by musical tones. He suggests that they are the result of two competing listening modes, which he referred to as the basic "analytic" mode and the more sophisticated "holistic" mode (Terhardt, Stoll, and Seewann, 1982b). Most musical tones permit multiple interpretations because their waveforms are complex and irregular, unlike the smooth, simple curves of a pure sine wave (Moore, 1982).

In Terhardt's model, analytic listening yields a tonal percept which is dominated by a single spectral pitch. Any of the low-numbered, higher-order harmonics may stand out as a spectral pitch; the number of candidates varies with the profile of the complex tone. In the case of a highly complex musical sound such as a triad, which consists of three complex tones, the listener's percept would be complex as well (Terhardt, Stoll, and Seewann, 1982b).

The tonal percept derived by holistic listening consists of a unitary virtual pitch which is inferred from all the separable spectral pitches. An example of a virtual pitch is the tone which may be generated from the spectral profile of the higher-order harmonics of a complex tone by means of an acoustical algorithm (Terhardt, Stoll and Seewann, 1982a). Such a virtual pitch does, in fact, correspond to the fundamental of the complex tone. Our perception of speech sounds, where the fundamental tends to perceptually dominate the individual harmonics, may be considered a common instance where holistic listening takes precedence over analytic listening.

According to Terhardt's model, perception of a musical chord in its entire complexity would typically involve the perception of several spectral and virtual pitches. The spectral pitches correspond to the salient harmonics in each of the three tones of the triad, while the virtual pitches correspond to each of the three fundamental tones in the triad. An additional pitch below the lowest fundamental tone in the triad might be heard as well (Terhardt, 1978), despite the fact that it is not physically present. Such additional virtual pitches may be obtained from the three fundamental tones of the chord, if the chord itself is perceived as a complex tone consisting of three pitches (Terhardt, Stoll and Seewann, 1982a). In this way, an additional virtual pitch corresponding to the triad's harmonic root, or "fundamental note", could be derived.

Terhardt suggests that in an actual musical listening situation, the two different ways of listening to musical chords compete simultaneously (Terhardt, Stoll, and Seewann, 1982b). An individual could be biased towards one of the two possible ways of listening. This could be due to individual variations in the listener's approach to the task, such as inherent preferences or momentary attentional state. Or the bias could be the result of contextual factors, such as the acoustic and functional context of the task.

Terhardt's model can be extrapolated to explain differences in the perception of chords. Predominantly holistic listening, which yields a virtual pitch percept, may result in the perception of a chordal fundamental predictable by rules of harmony. Analytic listening, on the other hand, may result in a highly variable triad pitch percept, which would depend on the relative salience of the various chord components, and which might be predicted on the basis of the listener's preferences and the sound qualities of the musical tone.

Before we consider a listening model which aims to identify the chord components which are salient for different groups of listeners, some fundamental musical terms must be discussed.

Structural Aspects of Music

How are we able to hear sequences of chords as harmonic progressions? Musical scales, which served as the background tonal context in the experiments reported here, can be played and heard as a sequence of tones. Scale structure can also be inferred from the relationships of the tones in a musical work. Diatonic scales such as C-major (C D E F G A B) or G-major (G A B C D E F# G) can be played in an ascending or descending series of seven steps. The characteristic pattern of a major diatonic scale is composed of five whole steps and two half steps. The half steps are between the scale positions III and IV, and between VII and

VIII. The roman numerals refer to the scale degrees, which define the pitch classes, or chroma, of the musical scale (Benjamin, Horvit, and Nelson, 1975).

A major triad can be defined as the combination of two consonant dyads. Consonance here refers to the sounding of musical notes with fundamental frequencies in simple integer ratios such as the octave (1:2), the fifth (3:2), the major third (5:4) and the minor third (6:5). When the fundamental frequencies of harmonic tones are in simple ratios, some of the upper partials will coincide (Moore, 1982). The most consonant combination of dyads consists of a minor third and a major third, which results in a consonant interval of a fifth from the lowest to the highest note of the triad (Wood, 1975).

The lowest tone of a major triad in root position and in its most close position is called the root, the middle is the third, and the top note is the fifth (Cannel and Marx, 1982). These numeric terms are based on the size of the tonal interval, which is the distance of the triad components from the root (Benjamin, Horvit, and Nelson, 1975). The resulting chord will be a major triad if the lower interval is a major third (M3) and the upper interval is a minor third (m3). The distance between the root and the fifth is a perfect fifth (P5). Using the notes of the diatonic scale, the tonic triad, built on the first note of the scale, will be major, as will triads built on the subdominant and dominant, or fourth and fifth notes of the scale. In C major these triads would be: on the tonic, CEG; on the subdominant, FAC; and on the dominant, GBD.

The position of the tones in the chord may vary. If the root of the triad is the lowest tone (bass) of the triad, this is called the root position. If the root tone is the highest note (soprano) of the triad and the third is in the bass, the chord is in first inversion (Hindemith, 1943; Priesing and Tecklin, 1959).

The conventions governing the interplay of pitch, duration, and intensity of musical sounds form an idiomatic musical language. While there is no single cross-culturally valid prescription for a definition of tonality (Serafine, 1983), the concept has usually been understood as a coherent system of musical meaning which depends on prescribed pitch relationships within one or more musical keys with a stable tonal centre (Rowell, 1983). While rhythm, melody, and harmony are heard by specialist and layman alike, listeners' perceptions of these concepts may differ dramatically. In the following sections, some essential points will be introduced.

Melody and harmony may operate together in music. Let us call the upper note in a triad the soprano voice, the middle note the alto voice, and the lowest note the bass voice. Now, if the listener attends to the movement of any single voice in the successive chords, a melody can be heard in a harmonic progression. In a similar fashion, a simple unaccompanied melody can be heard as implying a harmonic base which complements the line of the isolated melody voice. A melody could define the tonic just as well as a series of chords, since the presentation of the melody alone implies an underlying harmonic progression and a musical scale. As Palmer and Krumhansl (1987b) comment, musical excerpts with vertical (harmonic) organization also allow the perception of horizontal (melodic) organization. Consequently, a sequence of chordal harmonies may be comprehended by listening for a melody in any one of the voices in the harmonic progression.

Harmonic Progressions

A cadence, or harmonic progression, is a series of two chords presenting the same acoustic features as isolated chords, along with the additional information available from the sequence of chords. The term cadence is originally derived from the latin "cadere": to fall, which refers to the resolution of the harmonic progression by a downward pitch motion back to the tonic note. However, it is now accepted that cadences may either move away from or return to the tonic, and they may be played either as triads, or in classical four-part harmony (soprano, alto, tenor, and bass, with one note doubled), or in various other nontraditional combinations (Kitson, 1946; Priesing and Tecklin, 1959; Berry, 1987). Harmonic progressions can define a musical key and impart a sense of resolution, especially when they end on the major tonic triad (Krumhansl and Kessler, 1982).

Cadences usually close a musical phrase and allow a point of rest in the musical flow. In musical terminology, this rest is either medial, which indicates a need for continuation, or else final (Benjamin, Horvit and Nelson, 1975). The tonal context immediately preceding a cadence can be manipulated to examine the effect on listeners' perceptions, as measured in their judgments of stability.

Formally trained listeners are able to apply and recognize the use of strict voice-leading rules that prescribe the composition of ideal cadences. According to the model of voice-leading (Hindemith, 1943; Benjamin, Horvit, and Nelson, 1975; Berry, 1987), the motion of the harmonic progression can be defined in two aspects: linear and relative motion. The linear motion of a cadence describes the progression of a tonal voice between chords. This motion can be either conjunct, which refers to movement of a voice by a single tonal step at a time, or it may be disjunct, where the voice skips tonal steps.

Relative motion refers to the movement of the voices with respect to each other. In similar relative motion, two or more voices must move in the same direction, either upward or downward together. Contrary relative motion occurs when two voices move in opposite directions, for instance if one moves up while the other moves down. If two or more voices move together in the same direction and for the same diatonic interval, this is referred to as parallel motion. According to convention (Berry, 1987; Kitson, 1946), there are a few restrictions on the possible arrangements of a sequence of triads: jumps of an octave or more, as well as moves of a minor third or an augmented fourth (tritone), are not allowed according to formal rules.

Using parallel fifths in a chord sequence is also frowned upon in conventional Common Practice (Cook, 1987). If the two chords of a cadence are both in root position and in close voicing, both the bass and the soprano voice are separated by fifths. According to the rules of musical convention of the Western High Classical period (Baur, 1985; Rosen, 1972), this cadence is to be avoided. Such a cadence is considered by some to be a musical cliche without a satisfactory build-up of musical tension and final harmonious release (Berry, 1987).

Harmonic progressions can confirm or challenge the listener's expectancies of resolution in order to produce either a sense of tension or resolution. In music, every tone is usually preceded and followed by other tones. Stability refers to the perceived completion of a musical phrase. An unstable phrase is perceived to be unfinished, and music theorists conventionally explain this as an effect of the distance of the final chord's root tone from the tonic note in the prevailing tonality (Cook, 1987). This "distance" is a perceptual judgement based on the role of a given tone within a musical context. The root tone of a triad identifies the chord and serves as a reference point with respect to the tonic note, which is the first note of the prevailing scale. The tonic triad, which has the tonic note as its root tone, is the musical "centre of gravity" of that particular scale, and thus within a distinct tonal key.

Because both the tonic and the dominant note occur in the major triad, it has been argued that this structural redundancy due to the presence of two strong cues makes the major triad a powerful indicator of tonality (Roberts and Shaw, 1984). While other indices, such as the tritone, can be uniquely distinctive, the role of the tonic and the dominant is pervasive. Unstable tones are usually perceived in reference to more stable ones (Bharucha and Krumhansl, 1983), so that tones in close proximity are perceived as "leading to" the tonic. In similar fashion, less stable chords may be perceived in reference to more stable ones. But music theory does not explain the quality of a chord which determines its stability. Some theorists have chosen to discuss musical meaning in terms of an underlying harmonic structure (Lehrdahl and Jackendoff, 1983). The present discussion focuses on the surface sequence of events.

To determine which chord components make cadences musically meaningful, an analysis of the musical context in which the cadences are presented is important. The notes played before the cadence may set up musical expectations, which could determine the listener's attentional focus on salient aspects of the chords (Jones, 1987). These musical expectations are guided by the listener's understanding of the relationships of musical tones, and this understanding can be either implicit or else verbalized in particular musical terminology.

Adopting this representational approach, Jones (1981, 1982) theorized that an acoustic event is compared to an ideal prototype which is itself determined by an internalized set of rules. By either confirming or refuting prior expectations, a cadence would allow the listener to extrapolate the relations between chords in a musical work. A discrepancy between expectations of musical resolution based on the harmonic aspects of a musical piece, in comparison to expectations formed on the basis of the melody, could result in tonal ambiguity. Such tonal ambiguity often contributes to the attraction of the musical work (Sloboda, 1985).

Apart from the question of whether ideal prototypes of harmonic relations exist, the particular qualities of these hypothetical structures must be identified. If chords are always perceived with reference to the surrounding tones and to each other, it should be possible to determine hierarchies of stability and to identify both stable and unstable chords. This cognitive mapping task has been carried out in a series of studies (Krumhansl and Shepard, 1979; Krumhansl and Kessler, 1982; Bharucha and Krumhansl, 1983; Bharucha, 1984b) investigating the hierarchies of tones and of chords. A common observation was that the tonic chord (I) is the most stable element of the chord hierarchy. This is followed by the dominant chord (V) and then the subdominant chord (IV).

Harmonic progressions may themselves be characterized by another, simpler factor: melodic contour. Any pitch interval sequence yields a particular contour that can be recognized in several other scales, keys and melodies. For instance, a popular song can be recognized regardless of whether it is sung by a high female voice in A flat or by a low male voice in C sharp, as long as the listener pays attention to the melodic contour of the pitch interval sequence. A distinct pattern of upward or downward pitch motion can be recognized even when some of the individual pitches are altered, provided the overall contour is maintained. A pitch interval sequence, if it is at all distinctive, has a particular contour that can be recognized even when it is transposed to another key.

The distinctive contour of a melody can also be detected if it is embedded in a harmonic structure, such as a chord progression. Two or more successive triads allow the perception of a melody in a harmonic progression. According to formal theory (Berry, 1987), parallel movement of the three chordal voices across the chords emphasizes the pitch change. One way to judge the harmonic closure of a cadence would be to compare the pitches of the successive chords. A large difference, or pitch differential, should be easier to detect. Untrained listeners may generally prefer cadences with disjunct parallel motion and a large pitch change.

Within this background, there are two reasons why cadences appear acceptable experimental stimuli. The first reason is the consideration of ecological validity. Cadences commonly conclude musical phrases, and may therefore be considered true musical fragments. The second reason is related to stimulus control. Cadences, which consist of triads which can be defined and manipulated, seem ideal for this purpose. Any triad in a cadence can be described as a unique configuration of tones whose frequencies can be isolated and identified.

As discussed above, a cadence in its simplest form consists of two harmonious triads in close succession. The final "Amen" phrase of a traditional hymn is an example of the compelling and stable conclusion a cadence can provide when the final triad is a major chord built on the tonic note of the current tonal key.

The tonal context given by the sequence of notes played before the cadence can establish a tonality by setting up musical expectations in the listener. These cognitive-perceptual expectations are guided by the listener's understanding of the harmonic relationships of musical tones. A cadence allows the listener to confirm or modify his or her expectations concerning the relations of chords in a musical work. As Bharucha and Stoeckig (1986) have shown, when listeners are asked to rate the consonance of chords preceded by a series of chords, the speed and accuracy of ratings depend on the harmonic relations of the preceding context. The authors explained this as a priming effect, where the preceding chords serve as expectancy cues for related chords.

What other elements of the preceding context might affect the listener's ratings of a chord's consonance, or harmonic stability? Will these factors affect the listener's ratings of cadences as well? Can we predict listeners' judgments of cadence stability on the basis of their past musical experience? These are some of the questions that will be addressed in the experiments reported here. Before discussing alternative ways of listening which may affect listeners' ratings of cadences, a brief consideration of the neuropsychological basis of musical skill is appropriate.

Lateralization of Musical Functions

The question whether musicianship can be ascribed to a specific cerebral hemisphere is a controversial issue. The concept of lateralization of brain functions has given rise to the popular view that the artistic activity of the creative mind is caused by unconscious and intuitive action of the right cerebral hemisphere (Edwards, 1979). This widely held view is partially supported by the well-documented phenomenon of left cerebral hemisphere dominance for spoken language in right-handers (Schweiger, 1985). Researchers sought an equivalent basis for musical and spatial skills in the right cerebral hemisphere.

Early this century, case studies describing the preservation of some musical abilities after left hemisphere lesions were interpreted as evidence that musical functions are more easily compensated in the right hemisphere (Henschen, 1926). However, subsequent reports of preservation of musical function raised the possibility that for right-handed persons, musical functions are primarily localized in the right hemisphere, contralateral to language.

The belief that musical functions in dextrals are strictly localized in the right cerebral hemisphere has not been conclusively settled (Gordon, 1983). The reason for the current impasse may be found in the dichotic listening paradigm which underlies much of the theorizing. In this paradigm, developed by Kimura (1961), the subject is presented with competing speech or non-speech sounds to both ears, and their responses are compared under different conditions. Ear advantages on the dichotic listening task are commonly viewed as evidence of underlying hemispheric specialization. Efron (1985) described several problems which call this interpretation into doubt.

First, Efron reports that the most well-documented effect, a right-ear advantage (REA) for speech sounds, is a surprisingly weak effect found in less than 50% of the dextral population (Wexler, Halwes, and Heninger, 1981). This does not bode well for a confident display of hemispheric specialization of less lateralized skills, such as music.

Efron goes on to point out that alternate explanations of the results obtained from the dichotic listening paradigm have not been fully explored. The fact that significant ear advantages on a monaural paradigm can be observed for both human and non-human subjects calls into question the interpretation of ear advantages as a strong sign of hemispheric specialization into question (Divenyi, Efron, and Yund, 1977; Pohl, 1983). Efron (1985) hypothesizes that observed ear advantages may exist independently of any supposed hemispheric competition. He reviews evidence that there are considerable asymmetries in subcortical structures which may account for much of the disparity between the ears. Furthermore, Efron suggests that some right-left asymmetries could be ascribed to variations in the acoustic-to-neural transduction process at the cochlear level.

Apart from these anatomical considerations, Efron (1985) mentions two further issues which are directly relevant to the present discussion. On the one hand, we must examine the stimulus attributes: can we hypothesize that there is cerebral hemispheric specialization for particular stimulus types? On the other hand, the specific task demands must be carefully considered: do different methodologies permit the use of different strategies in the task? As Efron insists, it is important to separate the effect of the stimulus characteristics from the subject's use of cognitive strategies.

The use of specialized cognitive strategies need not be seen as synonymous with the development of hemispheric specialization, although cerebral asymmetries may be found. The question of whether there are two fundamentally different modes of processing need not be confined to a strictly localizationist view. Different cognitive strategies may exist as separate modes available to both hemispheres. It is beyond the scope of this investigation to provide an empirical answer to the question of hemispheric specialization.

More important here is the question of whether cognitive skills are primarily task-dependent or stimulus-dependent (Morais, 1982). It is also conceivable that both aspects may be involved. If, however, these hypothetical cognitive skills are primarily stimulus-dependent, a certain processing mode could be automatically activated when a particular type of stimulus is presented. Here, the guiding focus of the skill would be external. For instance, small black regular symbols on a white page would be "read" as letters and words.

Alternately, if cognitive skills are primarily task-dependent, a certain processing mode may be used for specific types of tasks, regardless of the nature of the stimulus to be manipulated. The guiding focus here would be internal, and would depend on the goal of the activity. For example, it is possible to detect a rhythmic pattern in music or in the complex designs of visual art.

Perceptual strategies may be considered habitual modes of cognitive processing. The assumption will be made here that these modes are not necessarily localized in a particular cerebral hemisphere. These perceptual modes can be quite conscious and deliberate, or they may be unconscious and automatic (Prinz, 1984). The question of whether the individual is conscious of this process is not a critical factor here. However, the transition from a consciously monitored, deliberate phase of skilled performance to one of automaticity is a necessary part of skill development.

Several classic studies have considered the question of automaticity in some detail (Posner and Cohen, 1980; Posner and Snyder, 1975). The key concept is one of attentional focussing, where the individual learns to allocate necessary perceptual resources to the task at hand. Two processes are hypothesized; the first stage is thought to be highly structured, conscious and deliberate, and the second becomes automatic and unconscious. This two-process view depends on the notion of levels of control.

With regard to the current discussion, it can be argued that during the early stages of acquiring a musical skill, the listener's attention must be highly focussed on the task at hand. We would expect to observe a deliberate and reflective performance in which the performer/listener is "caught up" in the task and is continually monitoring his or her performance.

Later, once the individual has acquired some depth of musical training, one might expect to observe signs of automaticity, as the listener/performer can allocate attention to interpretive and critical tasks even while performing the musical skill quite adequately.

Active Role of Listener

Listening to music, as a cognitive-perceptual process, can be more or less active (Efron, 1985; Judd, 1988). "What musicians call ear training...involves this sort of active, searching listening..." (Wilson, 1987, p.116). It is possible that representations or mental models of some kind can be stored in memory and compared to the new percept. If this view holds, an individual may first have to learn how to form such stable representations which would act as a model and guide for listening.

Before any training begins, most individuals have already had considerable exposure to music. The listener's untutored exposure to music must be distinguished from any explicit formal training he or she may have had. A person's reactions and exposure to music may vary, and every listener has developed different habits of listening which depend partly on their motivation and partly on their ability to hear out the separate voices in a musical work. For instance, many untrained listeners have only learned to focus on the vocal line; when asked to describe the role of harmony in music they often say that it serves to emphasize the message of the lyrics in the song.

This layman's description of the effect of music is compatible with the "associative", affect-driven theory of meaning in music (Meyer, 1956). The theory suggests that music conveys the emotional tenor of the lyrics by anticipating or echoing the message non-verbally. This could only be possible if there was a "basic vocabulary" for musical meaning. Meyer's theory does not account for the possibility that such a basic vocabulary does not exist in music. Without it, musical meanings would be highly tenuous, and might vary radically across different traditions.

Recent cross-cultural evidence (Castellano, Bharucha, and Krumhansl, 1984) seems to indicate that there are few universal meanings in music. The complexity of Indian music, with its 32 pitch classes, is highly challenging to the untutored Western ear. Bharucha (1984a) suggests that to understand the representation of musical stability, we must distinguish between event hierarchies and tonal hierarchies. Event hierarchies are classifications of specific musical events, which may include tones or chords. Any event hierarchy depends on an underlying tonal hierarchy, which is a classification of event classes in a particular culture. Thus, events that fit the tonal schemata for stable events stipulated by a particular tonal hierarchy will be considered stable in that culture.

Musically competent individuals who are familiar with the harmonic relationships of a particular tonal hierarchy are able to focus their attention on isolated sounds in a complex musical performance. Formal musical training requires the listener to learn and apply specific strategies of listening and performance according to an explicitly prescribed system of rules. A high level of training fosters a familiarity with musical terminology which permits the listener to accurately label details of harmonic structures.

An example of a high level of musical skill would be a composer's ability to mentally imagine or "audiate" an entire orchestral score. Another, more mundane example, would be an accompanist's ability to transcribe a piano score form the key of G to that of C, in order to accommodate the singer's vocal range. In both of these examples the skilled musician draws on an ability to use musical code to translate from the written representation to an audible performance.

Unfortunately, not all training is beneficial. Inadequate training methods may leave the listener as unskilled as before. This may become obvious on tasks that require close attention to tonal sub-structures, where the performer may operate with the added handicap of well-practiced and ingrained errors. For example, a singer without adequate formal training may strain the voice and tend to sing off key. On the other hand, an active, technically constructive listening mode is certainly not restricted to formally trained professionals (Rowell, 1983). By using self-taught strategies, talented individuals can develop great skill and flexibility on tasks which involve attention to detail. For example, self-taught musicians may learn to focus on the melody or rhythm line of any instrumental voice, instead of only on the vocals of a song.

Considering these possibilities, we can hypothesize two alternate ways of listening to chord progressions. One type of active listening could have an elemental focus. If the listener assigns a fixed and determined role to discrete elements in a harmonic cadence, then orders these components in an attempt to synthesize a meaningful whole, we might speak of an inflexible, element-focussed strategy. This way of listening should be fairly resistant to context bias, since the listener's attention is focused on the chord progression itself, rather than on the tones preceding and following the cadence. We would expect this strategy to be used by unsophisticated listeners.

An alternate listening perspective could have a more contextual focus. If the listener only provisionally isolates the tonal components of the chords in a cadence, while also judging the entire harmonic progression within a particular tonal context, we could consider this an example of a flexible, context-oriented strategy. This way of listening should be more sensitive to variations in context, since the listener can respond to changes in the sequence of tones in which the cadence is embedded by revising their judgement of the harmonic progression. We would expect trained listeners to prefer this strategy.

Listening Skills

If it is possible to use different ways of listening to music, it is likely that this activity involves more general cognitive skills. The process of learning a new skill is typically not smooth and continuous. Instead, it usually occurs by a rapid transition from one performance plateau in the task to another. Anderson (1982) has described three stages of skill acquisition: cognitive, associative, and autonomous, which correspond to three distinctive phases of activity. During the cognitive phase, a person must encode the skill in a rudimentary fashion in order to begin practice. In the associative phase, the individual detects and eliminates errors, and improves the concept of the task. Finally, in the autonomous stage, verbal self-monitoring becomes unnecessary, and the performance of the skill becomes smooth and masterful.

These three general stages of learning a skill can be discerned in the characteristic behaviour of the individual at each stage: first, the novice, who has had no training and shows little or no skill; second, the adept, who has experienced some training and exhibits some degree of skill; and finally, the expert, who has undergone extensive training and displays a high degree of skill.

A brief analogy may serve to illustrate the nature of these three stages. When an individual is asked a technical question, and requested to respond quickly and without recourse to any authoritative reference, both form and content of the answer will vary according to the person's familiarity with the material. If, for example, the man in the street is asked whether spiders are insects, he is likely to reply quickly and without hesitation that yes, of course spiders are insects. If a biology student is asked the same question, one might expect the individual to hesitate, and to eventually reply that spiders are, in fact, not actually insects. Finally, if the same question is posed to an entomologist, this authority will instantly reply that no, spiders are certainly not insects, but rather members of the family Arachnidae.

By analogy, familiarity and training in the conventional musical styles gradually foster the prompt and correct use of musical concepts over time. In the early stages of training, however, it is likely that the individual's performance initially becomes more erratic as naive concepts are replaced by more specialized and technical ones. Dowling's (1986) moderately trained subjects only performed at chance levels on a task requiring them to recognize melodies in different contexts. On the same task, both highly trained and untrained subjects performed well. Dowling suggested that the moderately trained listeners were still unable to utilize all the available cues effectively, while highly trained listeners were flexible enough to shift between context-invariant contour cues and context-sensitive scale-step cues. They had apparently learned to listen in a flexible manner.

Several questions arise with regard to the development of the cognitive skills necessary for different ways of listening to music. First, are well-defined, verbalized concepts necessary for musical perception? Or do basic, non-verbal musical units of meaning exist, which the listener can detect without any explicit knowledge of formal structure? What musical patterns could be construed as fundamental units?

Second, how do these putative musical units form our perceptions of music? Given a simple musical pattern, can some basic rules governing the relationship of these units be defined?

Third, is the function of such hypothetical musical units dependent on the tonal context in which they are embedded? The perception of a unique and coherent musical sound may be affected by changes in the relationships of its tonal components, and by the preceding sequence of musical tones.

Fourth, need we assume that theoretical structure is internalized during the course of learning a musical skill? Seashore (1938), for example, insisted that true musicianship depended on the use of auditory and intellectual representations of musical structure through imagery. While untrained and mediocre musicians may also have the capacity for imagery, Seashore argued that training and dedication was necessary to fully master these skills.

The fifth and final question asks whether different strategies of perception are employed during different stages of skill acquisition. Simply put, untrained listeners may not listen the same way as trained listeners. The typical ways of listening used to perceive basic musical features at different stages of listening skill should be defined. The use of these different ways of listening, also referred to as perceptual strategies, should be illustrated.

A Proposed Model of Listening Strategies

For the purpose of studying differences in skill level for listening to music, a three stage model of skill acquisition is set forth here. It is proposed that individuals in the first stage of skill acquisition, corresponding to Anderson's (1982) cognitive phase, be called novices. The novice will typically be found to use a rudimentary analytic strategy. Individuals using this strategy typically show a bias toward typical categories, and use concepts in an unconscious and unarticulated fashion.

In the second stage, Anderson's (1982) associative phase, we find the moderately trained individual, who employs an advanced analytic strategy. Here we can expect a more flexible approach to tasks; while errors of categorization still occur relatively often, we find rules used in a very conscious and deliberate way.

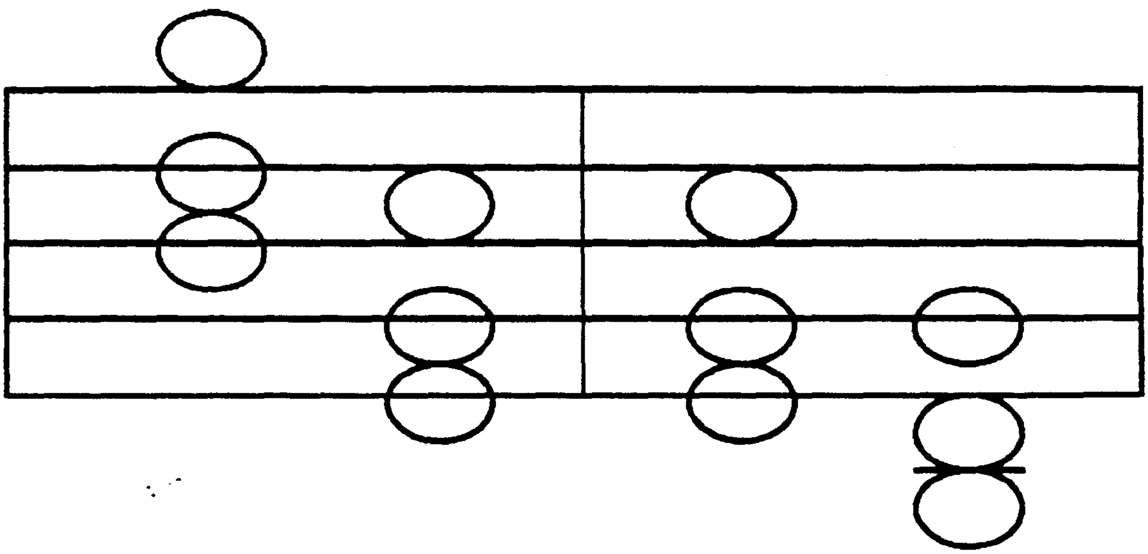
Finally, the third stage of skill acquisition will be occupied by the expert, who may use a holistic type of strategy. This stage would correspond to Anderson's autonomous phase (Anderson, 1982). Experts will exhibit a highly flexible approach to tasks; their performance will be virtually error-free, and they will appear to have attained a high degree of automaticity. They can also describe target concepts in words if required to do so. This complementary skill is beyond the capacity of the novice group, or even most members of the moderately trained group, who generally have little ability to describe the goals of their task.

In correspondence to these stages, it is proposed that two different modes of listening to musical cadences are possible. These modes, or attentional strategies, are (1) melody-tracking, and (2) voice-tracking. The first strategy, rudimentary melody-tracking, focuses on the characteristic pitch intervals in a melodic sequence in an inflexible, element-focussed manner. This strategy is compatible with Terhardt's (1982a, 1982b) model for the analytic perception of spectral pitch.

The perceived pitch of each triad in a cadence may depend on the most salient chord components. Usually the uppermost tones of each chord are most salient, due to the conventional placement of the melody in the soprano and the accompaniment in the lower voices. The soprano may be more salient due to the piercing quality of high-pitched tones. If the two triadic pitches of a cadence form a descending contour, the cadence may be considered stable, and the

Two examples of melody-tracking

focus only on soprano; prefer chords in first inversion position



G (Inv.) - C (Inv.)   C (Inv.) - G (Inv.)

degree of perceived stability should increase proportionally to the pitch differential (see Figure 1. N.B.: all examples imply treble clef only).

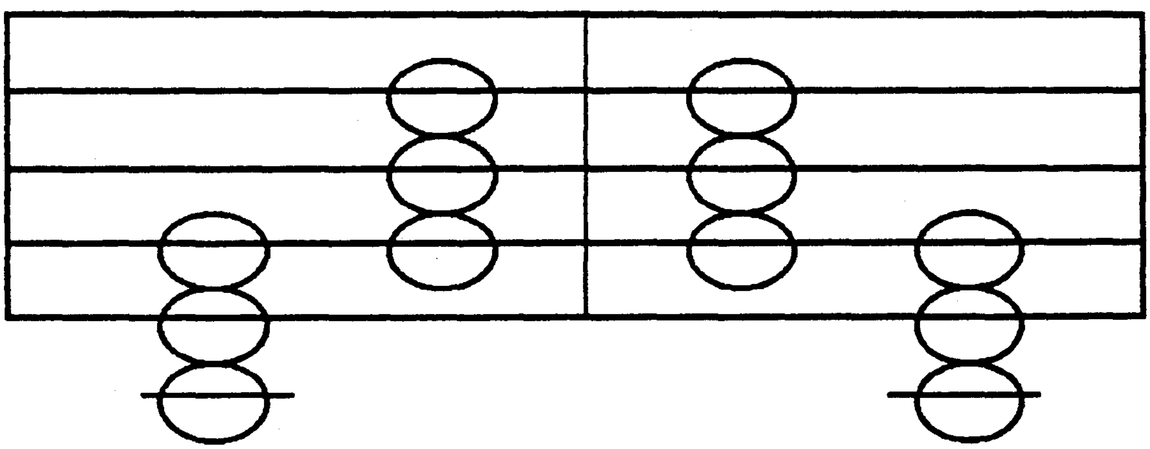
The second attentional strategy defined here, voice-tracking, is also compatible with Terhardt's (1982a, 1982b) model of analytic listening for spectral pitch. The immediate tonal context may have a strong effect on the perception of diatonic scale-step sequences. Listeners using a context-oriented voice-tracking strategy can follow the distinctive contour of any single voice. Conventionally, this would often be the uppermost note of a chord. Here, too, the uppermost voice is probably the focus of attention as a result of the conventional placement of the melody in the soprano voice and the accompaniment in the lower voice.

Yet voice-tracking listeners could also follow the lowest chord note. This flexibility may be evident in some other ways, as well. Although moderately trained listeners should still accept the traditional downward resolution of the cadence, due to their ability to focus on the melody line they should also rate upward resolution of a cadence as stable as long as the cadence melody of a single voice resolves to the tonic.

Successful use of an advanced voice-tracking strategy presupposes an understanding of pitch-interval relationships and an ability to correctly determine the tonic note and the tonal key on the basis of the relations

Two examples of voice-tracking

may focus on soprano or bass
root position chords acceptable



C (Root)- G (Root)   G (Root)- C (Root)

between notes in the scale preceding the cadence. This allows moderately trained listeners to discriminate melodies of equal or similar contour by recognizing subtle pitch-interval differences.

By isolating the three voices of a triad, listeners with some formal musical training should be able to follow a single distinctive melody line in a harmonic cadence. The principle of octave equivalence (Berry, 1987) dictates that the tonic note in a given key is both the beginning and the end of the scale, so that even upward resolution of a cadence into a triad in first inversion is formally acceptable, as well as resolution into root position. A focus on the melody line permits listeners to rate upwards resolving cadences as stable either if the melody is carried by the soprano voice and ends on the tonic in a first inversion triad, or if the melody is carried by the bass voice and ends on the tonic in a root triad (see Figure 2).

Only highly trained listeners who have the ability to verbalize scale-step relationships are expected to be able to utilize flexible retrieval strategies. They may be able to attend to either virtual or spectral pitch. In contrast to the reliance of moderately experienced listeners on relatively crude contour cues, professional musicians can link melodic pitches to their tonal functions. Musicians are comfortable dealing with musical tones in music-theoretical terms, using scale degree positions such as tonic and dominant to describe complex musical structure in an effective description of tonal relationships.

Highly trained musicians may be able to select the root as the most definitive element of a triad and follow it in a melodic sequence. This analytic ability can be called "root-tracking", and it would enable them to place the tones of the melody in relation to the tonic in a particular harmonic context. By holistically abstracting the unique configuration of a particular musical fragment, these listeners could recognize a melody or a cadence in any harmonic context. Root-tracking may be considered a specialized form of voice-tracking.

In brief, in order to determine whether a cadence is stable, the listener may use one of two listening strategies: first, by attending to the soprano voice, called melody-tracking; or second, by following any single voice in a chord progression, called voice-tracking.

Therefore, the crucial factors potentially affecting performance on this task are: the listener's background (trained or untrained), the type of cadence (final or medial), the position of the tones in each triad in the cadence (root or inversion), the motion of cadence resolution (upward or downward), and both the immediate and the extended musical context.

The experimental paradigm introduced here identifies separate groups of listeners on the basis of their stability judgments. These judgments are presumed to reflect ways of listening acquired through musical experience. It was expected that a harmonic progression would be rated stable if it offered a musically acceptable resolution of the cadence, preferably into the tonic triad of the prevalent key. Individuals who have achieved an equivalent level of theoretical or practical knowledge of music may all assign the same meaning to a musical excerpt. They may share a strategy of listening that has made particular musical elements especially salient.

The three experiments reported here are part of a series of investigations into listening strategies. Instead of listening to an entire musical piece, or even to a fragment of it, the subjects heard only two chords in succession: a simple cadence. Instead of requesting the listener to generate the appropriate scale afterwards, a scale was played beforehand, and the listener was asked to rate the stability of the cadence in the context of the preceding scale. This paradigm allowed systematic manipulation of the tonal relationships within and between the chords, and permitted quantification of the listener's perception of stability.

In the first experiment, listeners with some formal musical training were expected to use voice-tracking, an advanced, context-oriented analytic strategy. In the second experiment, untrained subjects also participated, and they were expected to employ melody-tracking, a rudimentary, element-focussed analytic strategy. In the third and final experiment, both trained and untrained listeners were tested on a new stimulus set to examine their attentional focus more closely.

CHAPTER TWO

STATISTICAL METHODOLOGY

The experiments discussed here were planned within the framework of a mixed factorial design, which lent itself well to multivariate analysis of variance (MANOVA). After identifying the main effects and interactions associated with the variability of the mean cadence ratings, cluster analysis procedures complementary to the MANOVA were used to illustrate the rating patterns across the cadence types that were characteristic of different groups of subjects.

Like any classification procedure, cluster analysis techniques attempt to order a set of objects, in this particular case mean cadence stability ratings, into several separate and distinctive groups. Unlike many other methods of classification, cluster analyses do not impose a classifying scheme upon the data, but attempt to reveal natural categories. The objects can be considered as points in multi-dimensional space, where the number of dimensions is equivalent to the number of variables (Everitt, 1980). Clusters can then be defined as specific delimited areas in this space which contain a higher density of object points relative to the surrounding space. The experimenter's choice of variables defines the problem space and affects the results of the analysis.

Let us consider a concrete example. Given a class of students, we may hypothesize that each student's grade on the final exam may be associated with at least two factors: time spent studying for the exam, and their previously obtained midterm grade. The membership of any group, such as this hypothetical class, can be investigated by one of several different similarity measures. Similarity measures permit the formation of a proximity matrix which compares the objects, or cases, after they have been evaluated on some measure of similarity (Aldenderfer and Blashfield, 1984). Correlation coefficients and distance measures are similarity measures currently in widespread use. Either approach allows us to present the data in a form which reveals how "similar" or "close" each object is to every other object in the set. Considering our class of students, we may wish to classify students on the basis of their grades, in order to investigate if those who study less are more likely to fail.

Correlation coefficients are frequently used measures of similarity which were originally developed independent of cluster analysis. Pearson's product-moment correlation coefficient (Pearson r), which is usually used to correlate variables, can also be used to correlate cases. Because of its inherent insensitivity to differences in the magnitude or elevation of variable or case values, the Pearson r has been described as a shape measure (Aldenderfer and Blashfield, 1984). It is also insensitive to scatter, the dispersion of the case values. In our example, finding that Sam's grades correlate positively with Anne's does not necessarily mean that these students are similar in any meaningful way. He may have obtained 19 on the midterm and 49 on the final, while she obtained 60 on the midterm and 90 on the final. While both students clearly improved their grade by 30 points, only one is failing the class.

When a distance measure is used instead of a correlation coefficient, the dissimilarity of objects is measured by an index of distance between data points (Gordon, 1981). This means that similar objects will obtain low values on this similarity coefficient, while very dissimilar ones will obtain high values. Two completely identical cases would obtain a similarity coefficient value of 0 on a distance measure. Considering our class of students, on the basis of their final marks Sam (49) and George (43) could be sorted into a "failing" group, while Anne (90) and Marie (88) could be assigned to the "passing" group. This example illustrates one shortcoming of distance measures: they are affected by differences in magnitude, or elevation, which can result in a loss of information from smaller magnitude cases. While Anne's and Marie's grades diverge by only 2 points, in contrast to the 6 point difference between George's and Sam's final grade, this difference between the two groups may not be reflected in the dissimilarity matrix.

Euclidean (geometric) distance is a popular dissimilarity measure which allows us to depict the relations of objects in two-dimensional space. The values of Euclidean distance coefficients have no inherent meaning; only the relative change between the coefficients obtained for different pairings of objects is important (Aldenderfer and Blashfield, 1984). The distances between all possible pairings of cases are compiled on the dissimilarity matrix, and the smallest value indicates the most similar pair.

Once a similarity measure has been selected, the actual process of clustering can proceed by several different methods. The first type of cluster analysis, the partitioning method, sorts a set of objects into a number of separate, non-overlapping groups. Each of these groups, or clusters, is different from every other according to one or more criteria, and no higher level of organization is assumed.

A partitioning method usually proceeds through three steps: cluster initiation, allocation to clusters, and relocation (Everitt, 1980). The first step, cluster initiation, requires that the maximum number of clusters be stipulated. This stipulation delimits the amount of differentiation possible. The choice of these starting points, which will serve as the initial estimates for the cluster centres. can be randomlv determined. or it can be set accordinα to the experimenter's workinα hvpothesis. However. unlike other methods of classification. the cateαories defined bv the startinα points in a cluster analvsis are potentiallv flexible, not fixed. In our example, we mav choose to initiallv divide the class into two αroups of students.

The second step in a partitioninα method, the allocation of objects to the clusters. generally occurs on the basis of each object's proximity to the nearest cluster centre. Proximity is used here in the sense of the distance between isolated objects and well-formed clusters (Everitt, 1980). For instance, our hypothetical student will be assigned to a cluster if her mark is similar to the average grade of the other students in that cluster. The degree to which clusters are permitted to overlap or be assimilated is variable, and depends on the particular procedure employed. Finally. relocation of objects into clusters takes place in order to optimize the cluster configuration (Everitt, 1980). Objects are relocated if their values do not match the criterion value. For instance, once all students have been assigned to clusters, new sub-clusters may be formed, and the membership of each grouping will be re-assigned.

The second type of cluster analysis, hierarchical clustering, sorts individual objects into well-associated groups. Each of the newly formed clusters, while being considered unique, may also share an attribute with several other sub-clusters within a common grouping. In this way, the entire set of objects is composed of a number of nested clusters. The hierarchical nesting method establishes relatively homogeneous groups of objects, while preserving heterogeneity across different levels. The final hierarchy of nested clusters can be illustrated on a dendrogram. The dendrogram, a graphic representation of clustering, uses the distance (or height) on the vertical axis to reflect the dissimilarity of individual items and branches of the tree-structure.

For our class of students, we might find that two main groups develop: students who are passing, and those who are failing. Students who are failing might also generally report less study time.

There are two ways in which an hierarchical structure can be obtained for a set of items: by agglomerative or divisive methods (Everitt, 1980). The divisive methods of hierarchical clustering are fundamentally analytic. They proceed by dividing the entire set of objects into successively smaller partitions. Ultimately, the divisive process will result in a large number of clusters equal to the number of separate objects in the set. Due to the probability of maximizing differences between individuals, this classification procedure is fraught with certain difficulties.

The alternative techniques, using agglomerative methods of hierarchical clustering, are primarily synthetic. They proceed by successively gathering, or fusing, objects into groups. New object clusters are formed by the affiliation of similar object pairs. Dissimilarities between objects are recalculated after each new linkage. Unchecked, agglomeration will lead to the formation of an all-encompassing cluster which will include all the objects in the set. The danger here lies in the possibility of minimizing differences between groups, as well as between individuals, while maximizing differences within groups. This eventually leads to the conglomeration of all objects, as the resemblances between individuals and groups are enhanced. The fusion of individuals and of clusters can proceed on the basis of either similarity or distance measures. Four agglomerative methods will be briefly considered here.

The single-link (nearest neighbor) agglomerative method fuses two or more single objects into clusters. adds single objects to existing clusters, and fuses clusters together. In this case, the distance between clusters is defined as the distance between their closest objects (Everitt, 1980). Using a matrix. the distance between two objects, d_{ij}, can be computed from the values in the rows (i) and the columns (j) for each object. Thus, if there are two groups, each with two objects, 1 and 2 in one group, and

3 and 4 in the other group, the allocation of another object (5) can be decided by comparing the distance between the new object and each of the members of the existing groups:

$d(1,2)_5 = \min \{d_{15}, d_{25}\}$

$d(3,4)_5 = \min \{d_{35}, d_{45}\}$ (Everitt, 1980, p.9)

In the single-link method, the new object will be allocated to the group to which it is closest. With each successive fusion the total number of clusters is reduced by one. All the clusters are gradually fused, one by one, so that the last linkage links all of the clusters, and thus all of the individual objects.

The class of students we have been considering would be classified one by one, comparing each student to another until similar pairs are identified, gradually adding less similar individuals, and then grouping pairs together.

The complete linkage (furthest neighbor) agglomerative method can be considered the opposite approach to the single-link method (Everitt, 1980). This method fuses objects and clusters on the basis of a cluster distance measure defined as the distance between the two furthest objects. Large distances between a new object (5) and the members of a group will lead to placement of the new object in another group:

$d(1,2)_5 = \max \{d_{15}, d_{25}\}$

$d(3,4)_5 = \max \{d_{35}, d_{45}\}$ (Everitt, 1980, p.11)

This complete linkage method will not fuse clusters as quickly as the single-link method, however the overall pattern of similarities finally obtained may often resemble the dendrogram generated by the previous method.

Applying this method to our class example, we would find that students are assigned to existing clusters only if there is more similarity between the candidate and the "furthest" member than there is between the candidate and any other clusters. For instance, a student with a grade of 55 would be assigned to the "failing" group if her grade was closer to the lowest grade of the failing cluster (19 - 49) than to the highest grade of the "passing" cluster.

Ward's method of hierarchical clustering is another agglomerative method. A dissimilarity matrix, based on the sums of squares of the criterion values, is used to calculate the Euclidean distance between objects and to classify them. Ward's method is based on the minimum variance criterion, which adds each new object to the cluster to which it contributes the least within-group variability (Everitt, 1980). Within-group coherence is maintained at the cost of maximizing between-group disparity. Ward's method does not account very well for outliers, which must be tagged on to an established cluster.

Considering our example, Ward's method might yield a classification of the students into an excellent group (80+) and a failing group (50-). But there might also be a few students with grades in the 60-65 range, who could be added on to one or the other group.

The group average method of agglomerative hierarchical clustering, the last to be discussed here, represents a compromise solution between single-link and complete-link methods. In this case, the distance between two clusters (k) and (ij) is defined as the average of all the distances between pairs of objects in the two groups.

$$dk(ij) = \frac{n_i}{n_i + n_j} dk_i + \frac{n_j}{n_i + n_j} dk_j$$

(Everitt, 1980, p.17)

The group-average method is more conservative than Ward's method; it does not maximize cluster separation during early stages of the clustering process. This robust method is thus least likely to generate spurious clusters (Aldenderfer and Blashfield, 1984). Outliers are more easily assimilated to the main body of a cluster, and this tends to preserve cluster membership at the cost of minimizing differences between individuals.

Looking once more at our student example, we would find that individuals are added on to existing groups by comparing their grade with the average grade in the existing group. An individual with a grade of 55 would be assimilated into the passing group (average 70) instead of the failing group (35).

By virtue of its robust qualities, the group-average method of agglomerative hierarchical clustering was selected to examine and illustrate the structure of the data obtained here. The purpose of this investigation was to explore the possibility that each individual's responses on the cadence rating task would naturally fall into one of two patterns. It was hypothesized that each listener, whether musician or non-musician, could be assigned to a group according to his or her pattern of responses. By examining the attributes of the stimuli, it should then be possible to identify those cadences whose presentation most clearly differentiates the two groups. Furthermore, by extrapolating from the hypothesized underlying listening strategies it should be possible to predict which cadences will act as group discriminators.

CHAPTER THREE

EXPERIMENT 1

The main purpose of the first experiment was to examine musically trained listeners' ratings of cadence stability, and to determine if their responses were distributed in a predictable pattern. A low stability rating indicated that the listener considered the harmonic resolution of a particular cadence to be incomplete. Conversely, a high stability rating indicated that the listener considered that particular cadence to be well resolved. Listeners' responses to different cadences were compared to a model of musical listening. The proposed listening model attempts to account for the listeners' performance by considering the role of five factors. These factors were identified and controlled in the experiment.

The first factor in the listening model was the amount of musical training the listener had received prior to participating in the experiment. Only those individuals with at least four years of formal individual training, who were currently practicing, were considered moderately trained listeners. In this first experiment, only moderately trained listeners participated.

Skilled performance in the cadence rating task was understood as the ability to consistently distinguish medial from final cadences, regardless of formal training or prior experience. For the purposes of this investigation, such a reliable performance was considered a sign of expertise. When distinguishing skilled from unskilled subjects, it should be clear that either group may contain both trained and untrained listeners. Trained listeners possess a repertoire of musical concepts and are fluent in the musical language of our culture (Swain, 1986), but this facility may not guarantee expertise in the experimental task. Untrained listeners, on the other hand, may have grown accustomed to the musical shape of acceptable harmonic resolution in cadences, without any formal training. If so, the process of musical acculturation must include some familiarity with elementary patterns of tension and resolution in harmonic cadences.

The second factor examined in this series of experiments was the type of cadence. For the sake of consistency, a conventional musical terminology will be employed when discussing cadences. Three cadences commonly employed in contemporary music (Priesing and Tecklin, 1959) were used in the first experiment. The <u>authentic</u>, or perfect cadence (V-I) moves from the fifth scale position (dominant) to the first (tonic), and listeners usually consider it very stable and resolved. Eight examples of G-C, a perfect cadence in C-major, are shown in the top row of stimuli in

Figure 3.

The less common plagal cadence (IV-I) moves from the fourth scale position (subdominant) to the first (tonic), and it is also usually found to be stable and resolved. Eight examples of F-C, a plagal cadence in C-major, are shown in the third row of Figure 3. The half-cadence, or imperfect cadence, (I-V) moves from the first scale position (tonic) to the fifth scale position (dominant), and listeners usually consider this cadence unstable and unresolved. Eight examples of C-G, an imperfect cadence in C-major, are shown in the second row of stimuli in Figure 3.

For comparison, an uncommon deceptive cadence, (V-IV), was also used (Priesing and Tecklin, 1959). Eight examples of this cadence are shown in the bottom row of Figure 3. This cadence moves from the fifth scale position (dominant) to the fourth (subdominant). Such a cadence is rarely used in popular music, and trained and untrained listeners alike were expected to consider it both unstable and unresolved.

The third factor to be investigated by the cadence rating task was the position of the root note of each triad in the cadence, referred to as triad position. The root note could be in the lowest position of the chord, in which case the chord is said to be in "root position". Examples of cadences in root position can be found in the first two columns of stimuli shown in Figure 3. Alternately, the root note could be in the uppermost position of the chord, while the third is in the bass, and this configuration is called the "first inversion" position (Hindemith, 1943). Examples of cadences in first inversion position can be found in the third and fourth columns of stimuli shown in Figure 3.

The fourth factor examined was the direction of resolution of the cadence. Across the two successive chords of the cadence, there could be either an upward or a downward pitch change as the cadence resolved into a higher or lower octave, respectively. Examples of upward resolution can be found in the even-numbered columns of stimuli shown in Figure 3, while examples of downward resolution can be found in odd-numbered columns.

The fifth and final factor investigated was the tonal context. This context consisted of an ascending melodic scale played before each cadence on every trial. The key of the scale alternated randomly between G-major and C-major. The 32 cadence stimuli shown in Figure 3 were each presented once in a C-major context and once in a G-major context, resulting in a 64 separate cadence combinations.

These five factors may affect the way an individual listens to a musical fragment. According to the proposed model of musical listening, two different ways of listening to cadences were hypothesized. The first, melody-tracking, was thought to be a rudimentary analytic strategy. When using this strategy, the listener is able to attend to the soprano voice. The second possible way of listening is called voice-tracking. Here the listener is able to follow any voice in a progression of triads.

To summarize, the first experiment examined the ability of trained listeners to rate the musical stability of simple two-chord cadences. The musical task was developed as a test for the different listening strategies described above. It was hypothesized that while untrained listeners may use melody-tracking, a rudimentary analytic strategy, the moderately trained listeners in this first experiment should use voice-tracking, a more advanced analytic strategy. They should be able to follow any tonal voice of a sequence of triads. Later experiments were formulated to investigate the performance of untrained listeners in comparison to trained listeners.

METHOD

Subjects. Sixteen experimentally naive volunteers participated: one was male, fifteen were female. All reported normal hearing, and 15 were right-handed. Their ages ranged from 18 to 33 yrs, with a median of 19 yrs. Median years of formal musical training was 7, ranging from 5 to 11 yrs. All of the participants were considered trained subjects. The subjects were recruited from a first-year psychology course; participation in an experiment was a requirement for course credit.

Apparatus. A Yamaha DX5M music computer with a built-in FM synthesizer presented the binaural stimuli, via a Yamaha CA-140 amplifier, to Realistic Pro-2 earphones. The subject, seated at a table in a sound-attenuating chamber, responded by moving a computer "mouse" which communicated the subject's choices directly to the computer. A Zenith colour monitor was situated at eye level directly in front of the subject. It was visible through the glass window of the experimental chamber. By moving the "mouse" on the table-top, the subject controlled the movements of the cursor on the screen of the monitor. The cursor had to be moved into the appropriate box indicating the response choice for each trial, and the response parameters were automatically recorded and stored on disk.

Stimuli. Each stimulus item consisted of one cadence, which was presented as a successive pair of chords. Four different two-chord cadence types were used, and each cadence type was preceded by a scale context of either C-major or G-major. The cadence types used were G-C (V-I or authentic in the key of C-major; I-IV in the key of G-major), C-G (I-V or semi in C-major; IV-I or plagal in G-major), F-C (IV-I or plagal in C-major; VII-natural-IV in G-major) and G-F (V-IV or deceptive in C-major; I-VII-natural in G-major).

The order of the preceding scale contexts (G-major or C-major) was randomized, and the scale was played in the ascending order of notes (from C3 or "middle C" to C4, and from G3 to G4). On each trial, each of the two chords in the cadence was presented in one of two positions: root or first inversion. For a chord in root position, the tonic triad in C-major was presented as C3-E3-G3, and the first inversion of the same triad was E3-G3-C4.

In total, a primary set of 16 cadences was obtained, shown in the odd-numbered columns on Figure 3. To control for the perception of subjective melodic motion across the pair of chords, this primary set was expanded to a set of 32 cadences. To allow an examination of the importance of motion of resolution, half of these moved from a lower into an upper octave (for example, from G2-B2-D3 to C3-E3-G3), while the other half moved from the upper into the lower octave (from G3-B3-D4 to C3-E3-G3). When the test cadences were presented to the subjects, the order of upward- vs downward-resolving cadences was counterbalanced along with the different scale contexts.

All tonal stimuli were presented in equal-tempered tuning. Each chord consisted of 3 simultaneous synthesized pure sine tones, which formed a triad. The tones were played at a comfortable listening level of approximately 60 dB SPL.

Procedure. A total of 128 trials were performed, with 64 stimuli presented binaurally in random order in each of 2 sessions. At the beginning of the session, the subject was instructed on the use of the "mouse", and given 2 "practice" trials. Each trial was preceded by an ascending diatonic scale in either C-major or G-major. The two chords of the cadence were immediately followed by the rating of stability, or harmonic resolution, on a 5-point scale. The boxes indicating possible choices were arranged in a semi-circle between the two poles labelled 1 ("tense, unstable") and 5 ("resolved, stable"). The higher the rating, the greater the reported stability. The response boxes were identified to the subject as 1) unstable, tense; 2) fairly unstable; 3) indifferent; 4) fairly stable; 5) stable, resolved.

The timing of events on each trial was as follows. First, the 7 tones of the diatonic scale were played in ascending sequence at a rate of 2 tones per sec. After a 50 msec pause, the first triad sounded for 500 msec, followed by another 50 msec pause, and finally by the second triad for 500 msec. There was no time limit for the subject's response; all trials were self-paced. At the onset of each trial, the cursor had to be returned to the centre square before the next stimuli would be played.

Additional information was requested on a questionnaire given to the subject after completion of the experiment. The subject was asked how many tones were present in each chord, and was requested to explain what he or she was listening for when trying to decide if the cadence was resolved. Subjects were also asked if they could identify by name any of the scales, chords or cadence types heard.

Analyses. The mean ratings for each trial were compiled in a multifactorial matrix across factors and subjects. Ratings obtained for each cadence in a particular context were compared to its ratings in other tonal contexts and by other subjects, as well as to ratings for other cadences.

All three experiments were designed within the framework of a mixed factorial design which lent itself well to multivariate analysis of variance (MANOVA).The use of this parametric test for the analysis of these data can be justified by Fienberg's (1980) argument that integer-valued variables, although usually labelled discrete, may also be considered continuous if they can be ordered along one conceptual continuum. In this case, parametric assumptions are justified and analysis of variance techniques are appropriate.

In the present experiment, the participants' cadence stability ratings were on an integer scale from one to five. This scale represented the continuum of perceived stability, ranging from totally unstable (1) to totally stable or resolved (5). The integer values can therefore be considered arbitrarily fixed verbal labels on an underlying and presumably continuous range of perception. The experimental constraints required the participants to adopt the artificial scale as a guide for their perceptual judgments, and it was seen that the subjects generally had no difficulty adhering to this guide. All of the subjects were able to utilize the full range of scale values. Parametric assumptions seem therefore legitimate in this case.

The quantitative data obtained from all participants on every trial in each of these experiments were collapsed across categories. Mean ratings were calculated for every individual participant across each of the stimulus types, and also across individuals for the different stimulus types.

The data was subjected to analysis by MANOVA conducted on an IBM-4381 mainframe computer using the SPSS-X package, and the mean ratings were also examined by cluster analysis (MICRO-CLUSTER, B.Edmonston, Cornell, 1985).

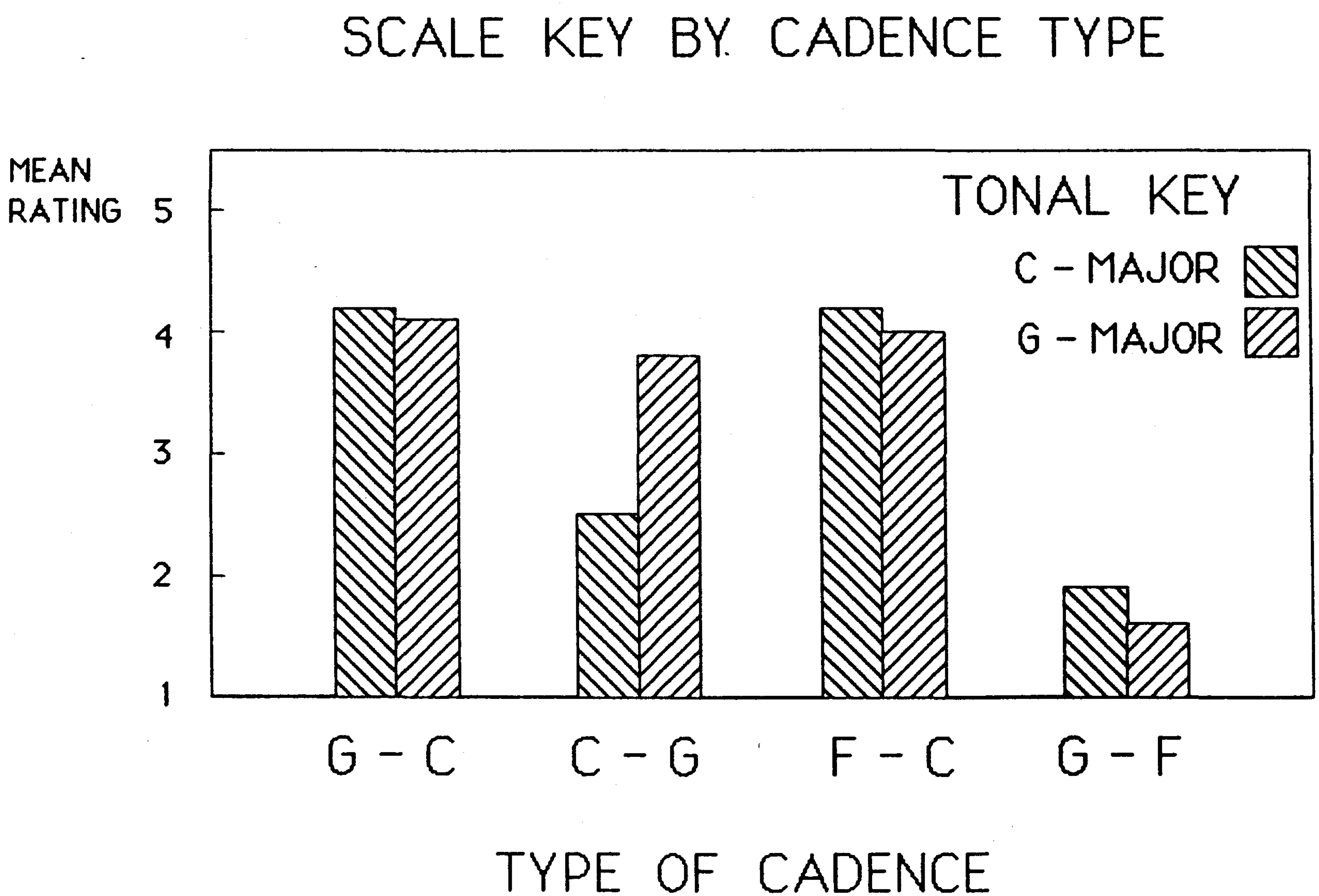
RESULTS AND DISCUSSION

The mean stability ratings given to the various cadences allow the classification of the cadence stimuli in various contexts, and the patterns of ratings reveal information about the listeners themselves. The musically trained listeners participating in this experiment showed no apparent difficulties performing the task. When questioned afterwards, some reported that they found the task challenging. Some of the participants were aware of the harmonic structure of the chords used in the experiment, and were in fact able to identify the chords correctly.

Based on their responses on the rating task, all of the listeners who took part in this first experiment could easily distinguish cadences ending on the tonic note from those which ended on other notes of the scale. These moderately trained listeners generally rated all of the cadences ending on middle C more stable than those which did not.

This same group of trained listeners also rated downward-resolving cadences more stable than cadences which resolved upward. While this finding at first glance resembles the profile expected from untrained listeners, it may simply illustrate trained listeners' familiarity with the conventional "cadential fall" in Western classical and popular music. Cadences with a falling contour were considered more stable (mean rating 3.4) than those with a rising contour (3.1). The analysis of variance revealed a main effect of this motion factor; $F\ (1,14) = 11.56$, $p < .05$.

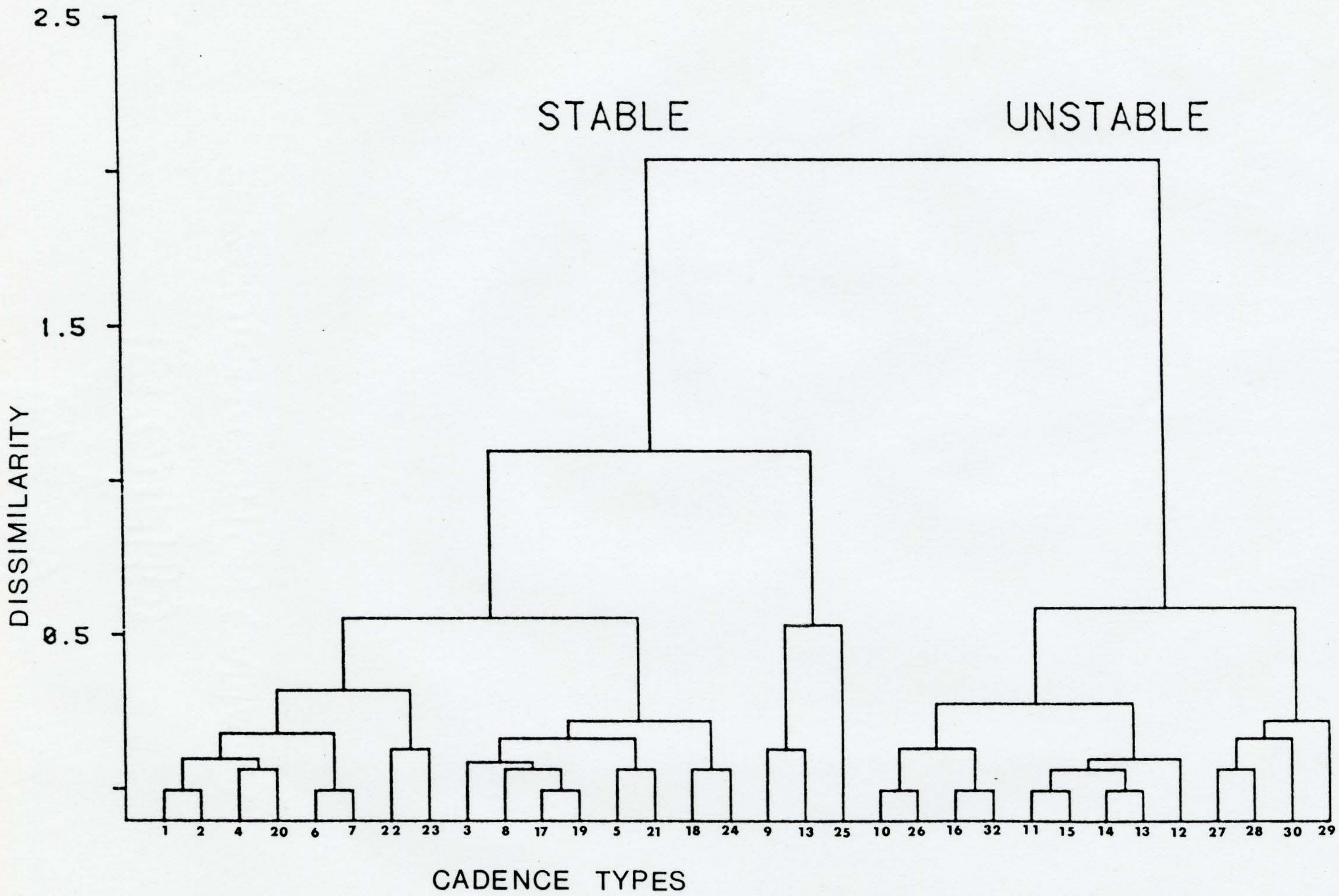
As expected, the four types of cadences were not all rated equally stable. This can be seen in the low ratings given by listeners for the highly unstable G-F cadence as opposed to the other three types (G-F: 1.8; G-C: 4.1, F-C: 4.1; C-G: 3.1). Cadences ending on C-major were generally rated more stable than other cadence types. The main effect of cadence type: $F\ (3,42) = 58.53$, $p < .001$, confirmed this difference.


SCALE KEY BY CADENCE TYPE
MEAN RATING
5
4
3
2
1
TONAL KEY
C - MAJOR
G - MAJOR
G - C
C - G
F - C
G - F
TYPE OF CADENCE

When a C-major scale was played before the cadence stimulus to create a C-major context, listeners generally considered the C-G cadence to be fairly unstable in comparison to the same cadence preceded by a G-major scale (C-G in C, 2.5; C-G in G, 3.8). In Figure 4, a histogram depicts the mean stability ratings for each cadence type (chord progression) preceded by either tonal key as vertical columns with variable shadings. It is clear that the C-G chord progression was the only cadence type that was affected by the key of the preceding tonal context. This is in accordance with conventional music theory, which describes C-G as an unstable semi-cadence in C-major. The same chord progression, however, played in the alternate G-major context, is a plagal cadence, where it was in fact generally rated higher. The trained listeners appear to have interpreted the contextual cues offered by the preceding scale as indicating a G-major tonality. In this tonal context, the cadence ending on G is a stable resolution to the tonic chord.

The G-F cadence was rated more unstable than the other cadences, regardless of the preceding tonal context (G-F in C, 1.9; G-F in G, 1.6). Ratings for the two cadences ending on C did not differ when they were preceded by different scales (G-C in C, 4.2; G-C in G, 4.1; F-C in C, 4.2; F-C in G, 4.0). The differences between the four cadence types in the two contexts are supported by the key

in C major key



by cadence type interaction: $\underline{F}$ $(3,42) = 11.33$, $\underline{p} < .001$.

The difference between cadences played in the two key contexts can also be illustrated by means of cluster analysis. The dendrogram in Figure 5 shows that by using the mean ratings assigned by listeners, the 32 cadence stimuli presented in C-major can be categorized into two clusters: stable (G-C and F-C) and unstable (C-G and G-F) cadences. The dendrogram illustrates the subjects' ability to discriminate the four cadence types in the C-major context in accordance with formal norms.

Three "unstable" cadences are included in the stable cluster in this figure (Cadences 9, 13, and 25). Considering the musical notation for these three examples (see Table 1), we find that in each of these cadences, the first chord is in root position. In two cadences (9, 25) the second chord is also in root position, and there is a slight downward motion for all three voices. In the other case (13) the second chord is in first inversion, and there is downward motion in two voices and unison in the upper voice. In all three of these cadences, the soprano of the second triad is not lower in pitch than the bass of the first triad. All the notes played in the two successive chords in each of these three cadences (9, 13, 25) can be found in a range limited by the soprano in the first chord and the bass in the second, in a closely overlapping harmonious sequence.

In the G-major context, however, the clustering of the cadence stimuli was markedly different. Figure 6 offers a graphical representation of the relative similarity of the stability ratings obtained for each cadence stimulus. The stimulus ratings fall into two large groupings, a "stable" and an "unstable" cluster. The unstable cluster, which contains cadences with the lowest mean ratings, consisted only of the G-F type cadences, regardless of chord position or tonal context. All other cadences were included in the stable cluster. The low ratings for the G-F cadences, which contain the F-natural chord, may indicate that this triad was considered alien to the G-major context, as expected according to standard Western musical convention.

Depending on the preceding scale, listeners gave different stability ratings to cadences composed of various combinations of triads in root position or first inversion. Cadences ending on a chord in the first inversion position were generally rated more stable when preceded by a G-major scale (I/I 3.5, R/I 3.4) than when preceded by a C-major scale (I/I 3.1, R/I 3.2). This difference was statistically confirmed by the interaction of key by triad position; $F(3,42) = 3.63$, $p < .05$. This finding could be due to a bias toward chords containing G4, which is the last note of the ascending G-major scale. The G4 is in the soprano (uppermost voice) in the first inversion of the G-major chord, and may therefore have been easier to detect than a G in the bass. This bias effect is unique to the G-major context; cadences ending on C were generally rated more stable in C-major, regardless of the position of the closing chord.

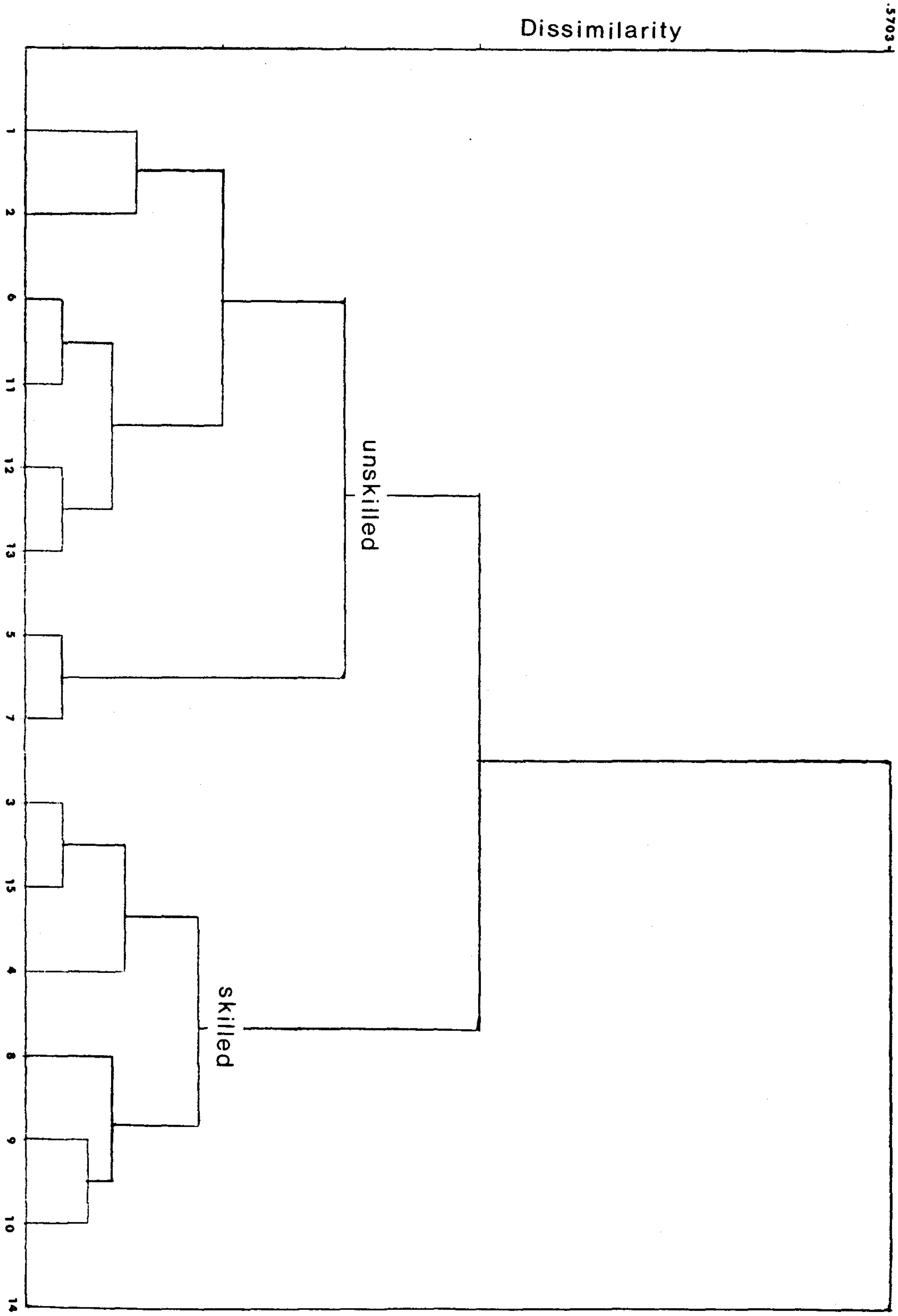
According to the listeners' responses, some cadences were noticeably more stable if the chords resolved downward instead of upward. Two cadence types showed a particularly strong effect of motion of the chord progression: the C-G type cadences (down 3.4, up 2.8) and the F-C cadences (down 4.3, up 3.9). While listeners generally gave higher stability ratings to cadences which ended on the tonic, as opposed to some other note in the scale context, this stability differential was particularly large for these two cadence types. The motion by cadence type interaction: $F(3,42) = 4.99$, $p < .05$ indicates that the difference between upward and downward motion was significant for these two types. For the other two types, there were no significant differences between cadences resolving upward or downward: the G-C type was generally considered highly stable, while the G-F type was rated highly unstable.

Chord position also influenced the impact of the motion of resolution, as seen in the interaction of cadence motion by chord position: $F(3,42) = 8.33$, $p < .001$. For cadences ending on the first inversion, downward resolving cadences are rated somewhat higher, but this difference is not significant (I/I down 3.4, I/I up 3.2; R/I down 3.5, R/I up 3.1). When both chords were in the root position, there was a clear preference for the downward resolution (R/R down

3.7; R/R up 3.1). But if the first chord of the cadence was in the first inversion position and the second chord was in the root position, downward resolution of the cadence was not rated more stable than upward resolution (I/R down 3.0; I/R up 3.2). According to voice-leading rules, resolution into the root position is more strongly supported by an upward movement in seconds or thirds, than by a downward motion in fourths or sixths.

The chord position was important for the effect of the cadence type as well. When the first chord was in root position and the second chord was in the first inversion position, a C-G cadence was considered particularly stable compared to other types. This can be seen in the cadence type by chord position interaction; $\underline{F}$ (9,126)= 4.94, $\underline{p}$<.001. Two cadence stimuli (13 and 14 in Table 1) are exemplary: the high ratings generally given to these cadences may reflect the effect of G4 by means of a comparison between the most salient voice, which is the soprano, and the last note of the scale (G4) immediately preceding the cadence.

Many of the highest-rated cadences could be predicted by voice-leading rules. Cadence 8 contains a unison on G4, cadence 24 has only linear 3rds and 2nds, cadence 13 has a unison on G4, and cadences 17 and 19 move down by perfect fourths (see Table 1). It is interesting to note that cadences with perfect parallel fifths were also rated highly (1, 3, 12, and 14). Analytically, the difference between the


Dissimilarity
.5703
1
2
6
11
12
13
5
7
3
15
4
8
9
10
14
unskilled
skilled

various cadence stimuli can be seen in the interaction of cadence motion by cadence type by chord position; $\underline{F}$ (9,126)= 3.03, $\underline{p}<.05$.

Looking at both tonal contexts, all the cadence stimuli were clustered according to the trained listeners' mean ratings, and the hierarchy of relations obtained were similar to the basic tenets of conventional music theory (Baur, 1985; Berry, 1987). Clustering of all listeners' mean ratings of cadences across both contexts according to a hierarchical agglomerative method separated a "stable" cluster of cadences (G-C, C-G types of cadences) from an "unstable" cluster (F-C, G-F types). The two separate clusters can be found on two different branches of a dendrogram. Regardless of the key context, two types of cadences, the C-G and the G-C cadence received higher mean ratings than the other two cadence types. According to musical convention, these two cadences are the most stable of all the cadences presented here in the G-major and C-major tonal contexts, respectively.

All of the participating listeners were clustered on the basis of their mean ratings to the various cadence forms. When subjects were clustered across stimuli in this way, a dendrogram (see Figure 7) showed that although all subjects had reported several years of training, their ratings were variable. Apart from subject 14 (who reported 8 years of piano lessons), the subjects fell into two clusters, with six in one group (mean 3.48) and eight in the other group (3.18).

The results of this experiment were promising but inconclusive. First, only self-described trained listeners participated, and their pattern of responses cannot be assumed to also characterize an untrained population. Secondly, the ascending scales used to set up a tonal context may have biased the listeners toward an attentional focus on the first inversion position, because all the scales ended on a high note. In order to test the hypothesis that cadence stability ratings may be guided by the final tone of the preceding scale, a second experiment was formulated. In the next experiment, the direction of the scales played before the cadences randomly alternated between ascending and descending. This was expected to affect both trained and untrained listeners' use of listening strategies.

CHAPTER FOUR

EXPERIMENT 2

In this second experiment, untrained listeners were recruited along with trained listeners, in order to extend the relevance of findings into the general population. Along with the same stimulus and context factors introduced in the first experiment, two additional factors were addressed in the second experiment. These new factors were the effects of training, and the direction of the preceding scale. The scale direction was expected to affect the listeners' stability ratings for cadences depending on the final note of the preceding scale context. The effect of this second manipulation was expected to differ for the two groups of listeners.

It was hypothesized that the performance of untrained listeners, whose judgments may be based on a rudimentary melody-tracking strategy, should not be significantly affected if the cadences were preceded by different tonal scales. The untrained listeners' ratings were expected to be primarily dependent on the contour of the cadence itself, independent of the preceding context. They were expected to consider cadences with a downward motion of resolution more stable than those which resolved upward. One way listeners could assess the downward motion of a chord was to compare the most salient pitch of each triad to the last note of the preceding scale, which was still in recent memory. By matching the pitch of the last scale note to the soprano voice in each of the chords of the cadence, the listener could determine if the cadence resolved upward or downward.

It was further hypothesized that the manipulation of the tonal context could affect the performance of listeners with a moderate amount of musical training. Their sensitivity to context may involve a direct comparison of each cadence with the scale presented prior to it. Pitch-matching could be a contributing element for moderately-trained subjects as well. In the ascending scale context of the first experiment, the final note was a high note, which may have biased listeners' attention towards cadences ending on this note. In the descending scale context, introduced in the present experiment, the final note was a low note.

Listeners who use a melody-tracking strategy could thereby be distinguished from listeners using voice-tracking. Melody-tracking listeners would tend to follow the top note of the chord. They were expected to give low ratings for cadences whose soprano voice does not move downward. Voice-tracking listeners, on the other hand, should also be able to focus their attention on the lowest note of the chord. If the final chord is in root position, the trained listeners should be able to give the cadence a stable rating even if it resolves upward.

METHOD

Subjects. Twenty-one experimentally naive volunteers participated; all reported normal hearing. On the basis of their years of formal training, eleven participants were assigned to the untrained group, and 10 to the trained group. In the untrained group, 5 participants were male, and 6 were female. Their age ranged from 19 to 33, with a median of 20. There were 9 right-handers and 2 left-handers. Median years of formal musical training in the untrained group was 1, ranging from 0 to 3.

In the trained group, 3 participants were male, and 7 were female. The age of the trained participants ranged from 19 to 23 years, with a median of 19. There were 7 right-handers and 3 left-handers. Median years of formal musical training was 5, ranging from 4 to 8. Six of the subjects in this experiment were actively recruited from the McMaster community and received a small remuneration; the others were all first-year psychology students for whom participation in an experiment was a requirement for course credit.

Apparatus. The equipment employed in this experiment was the same as for Experiment 1.

Stimuli. The stimuli presented in this experiment were the same as in Experiment 1, with the addition of a second complete set of 32 stimulus cadences played in a descending scale context, resulting in a total of 64 stimuli. For simplicity, only root position and first inversion chords were used.

Procedure. The procedure was identical to Experiment 1, although there were twice as many trials. The presentation of ascending and descending scales before the cadence was randomized.

RESULTS AND DISCUSSION

As in the prior experiment, downward-resolving cadences generally received higher ratings than upward-resolving ones (down: 3.47, up: 3.05). Cadence contour evidently is a cue for ratings of stability for both trained and untrained listeners. The difference between upward and downward motion of resolution was statistically significant; F (1,19)= 37.33; $p < .001$). The difference was particularly large for C-G cadences (C-G down 3.4, up 2.6) in comparison to other cadence types, and this interaction was statistically significant; F (3,57)= 8.01; $p < .001$.

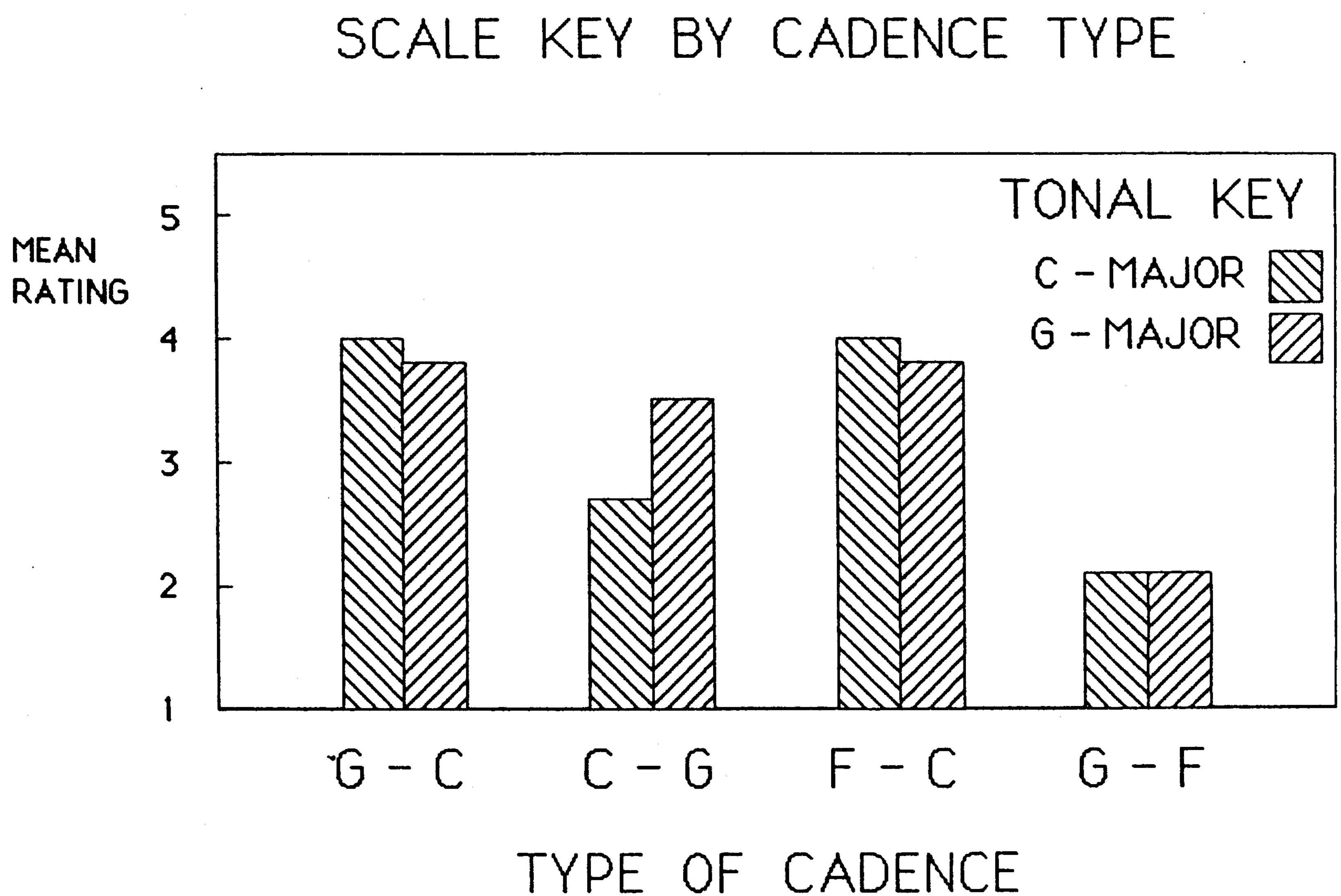
In contrast to the listeners in the first experiment, listeners here generally considered cadences presented in the G-major key context more stable (3.34) than those same cadences presented in the C-major context (3.17). This finding was significant; F (1,19)= 12.39; $p < .05$. The result had been anticipated by the clustering of stimuli across subjects in the previous experiment, where the number of cadences in the "stable" cluster was greater in the G-major context than in C-major. The results of the second experiment may also have been affected by the added participation of untrained subjects, most of whom may have matched the soprano voice of the first triad in the cadence to the pitch of the last note of the preceding scale. Such a pitch-matching tactic would be highly feasible, because the ascending G-major scale context offers a salient cue whenever the last note of the scale is echoed in the cadence.

Cadences ending on C-major were generally given higher mean stability ratings (G-C 3.93, F-C 3.96; compared to C-G 2.99, G-F 2.06), and this effect was statistically significant; $\underline{F}$ (3,57)= 55.24, $\underline{p}$<.001). The stability ascribed to cadences ending on C-major cannot be a simple effect of familiarity in this experimental context, because the cadences were presented as many times in the key of C-major as they were in G-major. Nor can it be due to the repeated presentation of C-major chords, because an equal number of G-major chords was presented within each experimental session. The apparent preference for cadences ending on C-major can be interpreted as evidence for the longstanding influence of Common Practice views, which stress the importance of the perfect cadence for tonality (Cook, 1987; Berry, 1987). If a perfect cadence is generally considered more stable than a plagal cadence, the preference for C-major becomes understandable; no perfect cadences were presented in G-major.

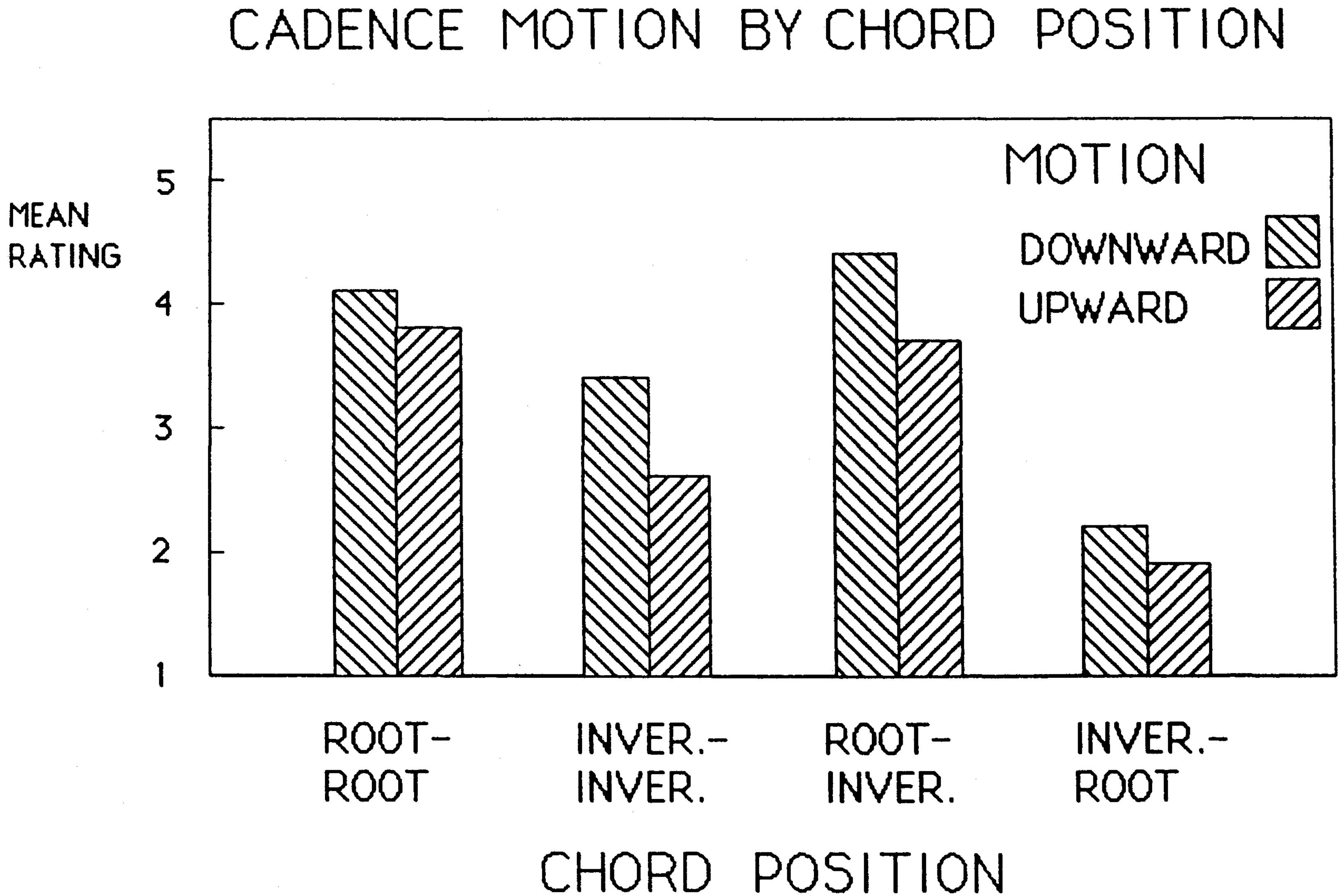
As in the first experiment, the stability ratings assigned by listeners differed according to the chord positions in the cadences: F (3,57)= 2.85; $p < .05$. Of all possible combinations of chord positions in the cadences, the I/I position generally received the highest stability ratings (R/R 3.23, I/I 3.36, R/I 3.23, I/R 3.13).

When the preceding scale descended, listeners' judgements of the stability of the following cadence did not depend on key (G-major: 3.26; C-major: 3.22). Yet when the preceding scale ascended, the key of the scale context was important (G-major: 3.35; C-major: 3.15); F (1,19)= 10.02; $p < .05$. In G-major, the following cadence may have been rated more stable if the soprano voice echoed the last note of the scale. This would support the interpretation that the last note of an ascending scale can be compared to the perceived melodic line of the successive chords.

The cadence types were also rated differently in the two different tonal keys, regardless of the direction of the preceding scale: F (3,57)= 13.60; $p < .001$. The C-G cadence type was rated more stable when preceded by a G-major scale than when it was preceded by a C-major scale (C-G in C-major 2.7, in G-major 3.5). This effect can be explained as the result of conventional harmonic standards. The chord



SCALE KEY BY CADENCE TYPE
MEAN RATING
5
4
3
2
1
TONAL KEY
C - MAJOR
G - MAJOR
G - C
C - G
F - C
G - F
TYPE OF CADENCE

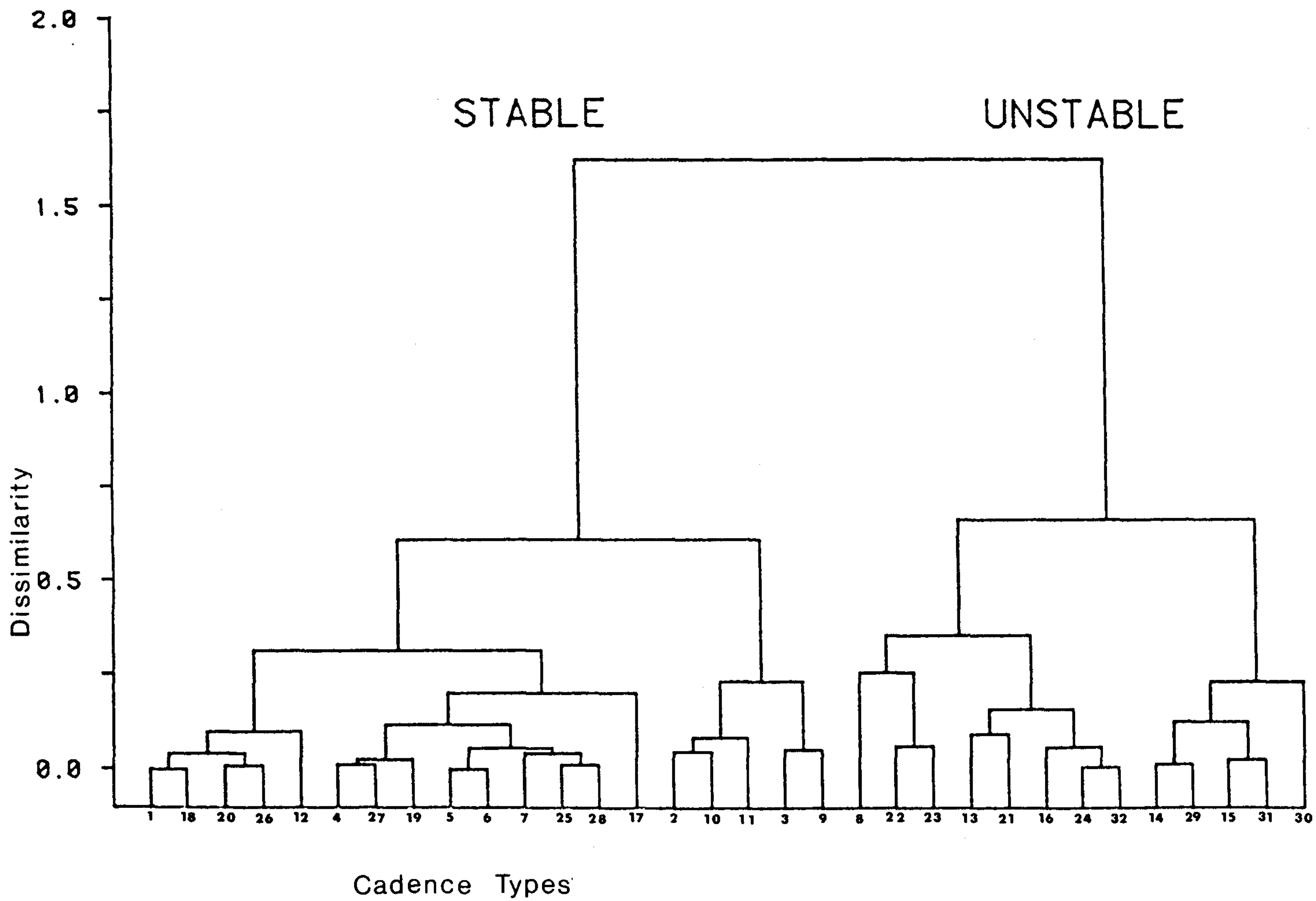

CADENCE MOTION BY CHORD POSITION
MEAN RATING
5
4
3
2
1
MOTION
DOWNWARD
UPWARD
ROOT-ROOT
INVER.-INVER.
ROOT-INVER.
INVER.-ROOT
CHORD POSITION

sequence C-G was assigned a higher stability rating in the G-major context, where it was a stable plagal cadence, than in the C-major context, where it was an unstable semi-cadence. This is illustrated in Figure 8, where the discrepancy between the mean rating for the C-G cadence type in the C-major tonal context, shown by the third bar of the histogram, in comparison to the rating for the same cadence in the G-major context, shown by the fourth bar, is readily apparent. This finding replicates the results of the prior experiment, and confirms the role of the preceding scale in establishing tonality. It should also be noted that even the untrained participants assigned their stability ratings in accordance with rules of traditional Common Practice.

Downward-resolving cadences were generally rated more stable than upward-resolving ones if the final chord was in the first inversion position; F (3,57)= 10.33; $p<.001$. The effect is illustrated in Figure 9, which shows that although a root/root cadence (up 4.1, down 3.8) is rated as just as stable as a root/inversion cadence (up 4.3, down 3.7), the inversion/root cadence is rated very unstable, especially for upward-resolving cadences (I/I down 3.4, up 2.6). The lower ratings for cadences beginning with a first inversion triad may be explained by the fact that triads in first inversion have their root notes in the uppermost position, which increases the probability that listeners who are following the soprano voice can detect the root. This was expected to be more important for the untrained listeners than for the trained listeners, who were predicted to show an ability to focus on other voices. The fact that no group differences were found does not permit the conclusion that training effects do not exist. It may simply emphasize the necessity of determining the listener's prior training through some empirical measure, rather than through self-report.

When both chords were in root position, only untrained subjects gave higher mean ratings to downward-resolving cadences in comparison to upward-resolving cadences (R/R down 3.9, up 3.1; I/I down 3.1, up 2.5; R/I down 4.1, up 3.7; I/R down 2.1, up 2.0). While this confirms untrained listeners' general preference for a distinct descending cadential contour, it also indicates that this preference does not apply equally to all chord configurations: $F(3.57) = 5.08$; $p < .05$.

Trained subjects clearly also rated downward-resolving cadences as more stable when both chords were in the first inversion position (I/I: down 3.8, up 2.7), or if the two triads were in different positions (R/I: down 4.4, up 3.8; I/R: down 2.2, up 1.9); $F(3,57) = 5.08$; $p < .05$. However, unlike untrained listeners, trained subjects gave equivalent stability ratings to both upward- and to downward-resolving cadences when both chords were in root position (R/R: down 4.3, up 4.0). The trained listeners may


STABLE
UNSTABLE
Dissimilarity
2.0
1.5
1.0
0.5
0.0
1 18 20 26 12 4 27 19 5 6 7 25 28 17 2 10 11 3 9 8 22 23 13 21 16 24 32 14 29 15 31 30
Cadence Types

be able to attend to an alternate voice by focussing on the root note in the lowest voice of the triad, rather than on the conventional soprano voice. Unlike the untrained subjects, they found either direction of resolution acceptable in this case.

The effect of triad position on the listeners' ratings of the cadences differed significantly for the different cadence types; $\underline{F}$ (9,171)= 8.03; $\underline{p}<.001$. For two of the cadence types, the I/I position received the highest ratings (G-C 4.23, F-C 4.23) in comparison to the other combinations of chord position.

When a G-C type cadence was preceded by a G-major scale, the I/I position received the highest ratings (4.19). Both the I/I (3.8) and R/I (3.79) positions received high ratings in the C-G type, which is a plagal cadence in G-major. Here the listeners' judgements agree with conventional practice, and the interaction is significant; $\underline{F}$ (9,171)= 5.27; $\underline{p}<.001$.

Clustering of the stimuli across subjects divided the cadences into two clusters (see Figure 10). The "stable cadence" cluster on the left arm of the dendrogram contained most of the G-C cadences and several cadences from the other three types. This picture was strikingly different from the straightforward clustering of cadences according to type obtained in the prior experiment (compare Figure 5). The reason may lie in the disparate background of the


SKILLED
UNSKILLED.
.65
.50
.35
.20
.05
Dissimilarity
T T T T T T T T T T
All Subjects
T:trained

participants. Some individuals who reported musical training may actually not have had the skills necessary to distinguish the different cadence types in a consistent and accurate fashion.

In fact, a clustering of the participants across the cadence stimuli on the basis of their mean ratings divided the participants into two separate groups: "skilled" listeners (mean 3.55) and "unskilled" listeners (2.91). The skilled group shown on the left (see Figure 11) consisted of those able to consistently and reliably distinguish medial from final cadences. They compared favorably with the trained participants of the prior experiment, even though the reported average training of the present group was much lower. Several individuals falling into the skilled category on this experiment actually reported no training at all. On the other hand, several individuals in the unskilled group, shown on the right side of the dendrogram, had previously reported several years training on an instrument. It is possible that some untrained individuals are using skillful and flexible context-sensitive strategy without having received formal training.

The prediction that the direction of the preceding scale would affect the stability ratings was confirmed. The ascending scale may indeed bias the listeners' attention toward the uppermost notes of the cadence. In this regard, trained listeners seemed to be affected as much as untrained. This was unexpected. The untrained listeners participating in the second experiment may have been matching the cadence chords to the last note and the direction of the preceding scale. If the notes in the scale were ascending and the cadence began on the final scale note, then resolved downward, the cadence was more likely to be perceived as stable, due to the "fall" in the contour. If the scale descended and the cadence began on the final scale note and resolved upward, the cadence was more likely to be rated unstable by untrained listeners.

Voice-tracking by moderately trained listeners may also be biased by the ascending scale simply because it emphasizes a simpler focus on the upper (melody) voice. The stimuli employed in this experiment may not have challenged their abilities sufficiently to stimulate changes in strategy. To provide this challenge, a new stimulus set was devised.

In the next experiment, traditional four-part harmony was used to form the cadences, in order to obtain a clearer picture of the listeners' attentional focus. The standard triad configurations used in the first two experiments gave the soprano voice undisputed prominence. By using a doubled root to counterbalance the prominence of the soprano, the new cadences offered the listener a choice between following either the soprano or the bass notes in the harmonic progression.

CHAPTER FIVE

EXPERIMENT 3

For the third and final experiment in this series, a new stimulus set was devised to permit more detailed examination of listeners' reactions to stable and unstable cadences. The cadences selected were chosen to challenge two specific hypothetical perceptual strategies, melody-tracking and voice-tracking. These cadences differed according to the relative perceptual salience of their four voice components. The four voices, ranging from high to low pitch, were soprano, alto, tenor, and bass. The pitches of these voices were determined by the three tones of a major triad, along with a doubled root tone.

As in the first two experiments, the listeners' stability ratings for these cadences were expected to be influenced by the conventional norms of Common Practice. All of the cadences were presented within the same fragmentary musical context used in the prior experiments. Each cadence was preceded by either a C-major or a G-major scale, and the tones of these scales were played in an ascending order for half of the presentations, and in descending order for the other half. The order in which the alternation between the two possible scale keys and scale directions occurred was randomly determined.

The listeners' ability to distinguish the two different cadence types, G–C and C–G, was expected to vary according to their formal musical training. The two groups of listeners, trained and untrained, were expected to be able to distinguish the C-major from the G-major key context with variable success; more experienced listeners were expected to have a slight advantage. The effect of three other factors, namely the direction of preceding musical scale, the particular voice which led in the cadence (either the soprano or the bass), and the motion of cadence resolution (either a downward or an upward contour), was also expected to vary for the two different groups. It was expected that evidence could be found for the use of the two distinctive listening strategies.

Although a general preference for downward resolution was predicted, because this exemplifies the commonly used cadential "fall", any such effect was expected to be strongest for untrained listeners, regardless of the tonal context preceding the cadence. Only trained listeners were thought likely to give stability ratings which would reflect a flexible focus of attention. They should be able to focus on either the soprano or bass voice.

It was also expected, however, that even moderately trained listeners would still tend to focus on the soprano voice, because the soprano is usually the most salient vocal line in Common Practice (Rosen, 1972; Cook, 1987). The ascending scale context was intended to emphasize this hypothesized effect of the soprano voice, while the descending context was expected to counteract the listeners' focus on the soprano voice. The melodic contour of the preceding scale could either contrast with or be mirrored by the contour of the cadential progression. This comparative process allows a focus on the soprano to be emphasized when the cadence resolved downward, because of the contrasting direction of the ascending scale in comparison to the descending contour of the melodic line in the cadence.

METHOD

Subjects. Forty-nine experimentally naive volunteers participated: 22 were male, 27 were female. All reported normal hearing; 36 were right-handed, 2 left-handed, and 11 ambidextrous (score derived from the Oldfield handedness questionnaire). The age of the participants ranged from 17 to 25 yrs., with a median of 19. Median years formal training was 4, ranging from 0 to 14. Based on their years of training, 28 participants with 3 years experience or less were assigned to the untrained group, and 21 others to the trained group. All of the subjects were first-year undergraduate psychology students for whom participation in an experiment was a requirement for course credit.

Apparatus. The equipment employed in this experiment

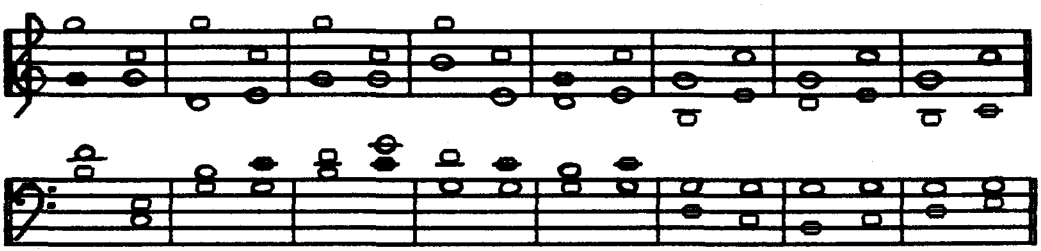
Chord sequences used in Experiment 3.

G – C 1. 2. 3. 4. 5. 6. 7. 8.

bass voice leading

9. 10. 11. 12. 13. 14. 15. 16.

soprano voice leading



C – G 17. 18. 19. 20. 21. 22. 23. 24.

bass voice leading

25. 26. 27. 28. 29. 30. 31. 32.

soprano voice leading

was the same used in Experiments 1 and 2.

Stimuli. The stimuli presented in this experiment were very similar to those presented in the prior experiments, and are shown in Figure 12. Only two types of chord progressions were used: C-G and G-C. This resulted in three cadence types as well as one non-cadential progression. In the C-major context, the stable perfect cadence (G-C) and the less stable imperfect (C-G) cadence were presented. In the G-major context, a stable plagal cadence (C-G) and an unstable non-cadential progression (G-C) were presented. The positions of the four tones within the doubled-root chords in each cadence were determined according to rules of voice-leading. Either the bass voice (the lowest tone in the chord) or the soprano (the highest) was the leading voice in the cadence. The motion of this leading voice, which carried the root progression, was either upward or downward.

Procedure. The presentation of the stimuli was essentially the same as in the prior experiments. The 32 basic cadences were presented in random sequence in both the C-major and the G-major key context, with either the ascending or the descending versions of these scales, for a total of 128 trials. Timing and task demands were the same as in the previous experiments.

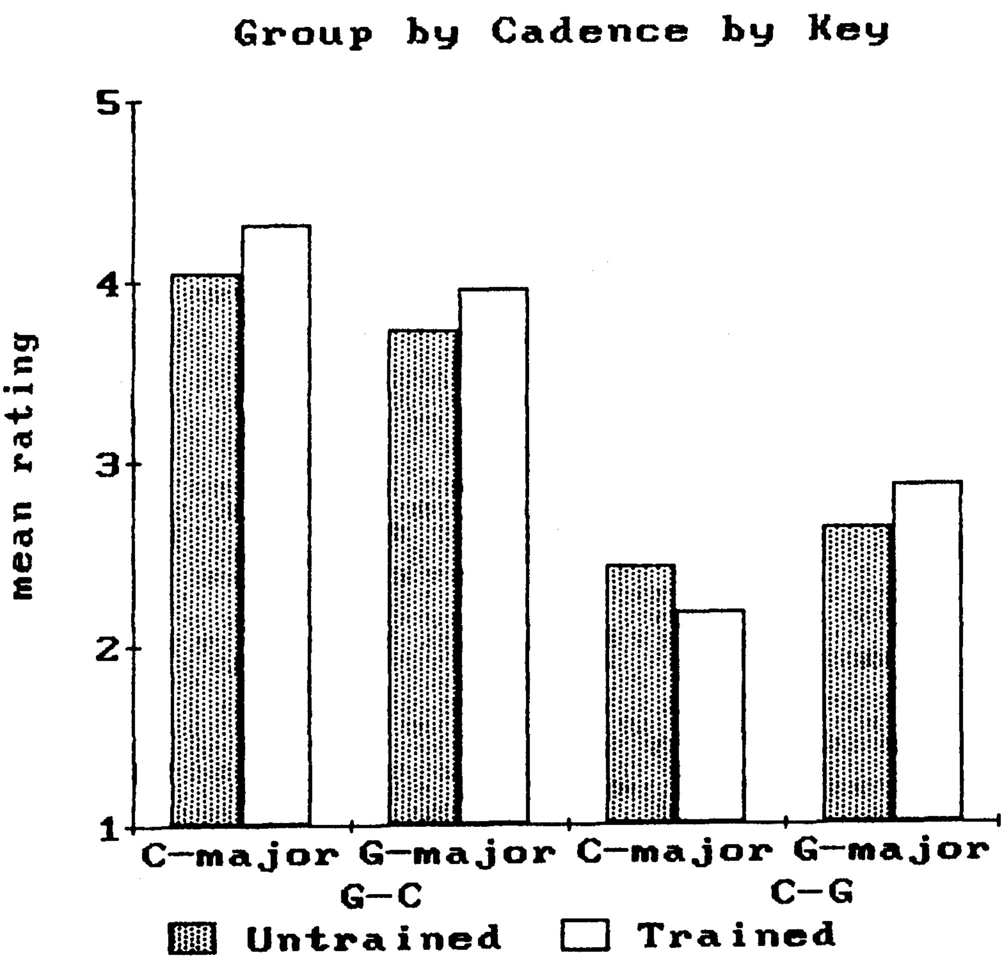
Statistics. As before, an analysis of variance was performed on the mean stability ratings for all of the stimulus cadences. The effects of the between groups factor, training, and of the five stimulus variables, cadence (G-C or C-G), key (C-major or G-major), scale (ascending or descending), voice (bass or soprano) and motion (upward or downward motion of voice) were examined.

RESULTS AND DISCUSSION

The formally trained listeners were not exempt from the common preference for a "falling" cadence. Cadences with downward motion of the leading voice were generally rated more stable (3.49) than cadences which resolved upward (3.05). This illustrates the preference of listeners in both groups for a clear cadential fall: $\underline{F}$ (1,47)= 92.05; $\underline{p}<0.001$.

Both trained and untrained listeners appeared to have a preference for harmonic resolution into C-major. Cadences ending on the C-major triad (G-C: 4.01) were generally rated more stable than those ending on the G-major triad (C-G: 2.53). This continues the trend toward the C-major tonality which was observed in the earlier experiments; $\underline{F}$ (1,47)= 175.03; $\underline{p}<0.001$. It appears that the C-major key was clearly established in the experimental context, perhaps at the cost of the G-major key.

The tonal key of the preceding scale clearly affected the listeners' perception of the cadence types. Perfect cadences were considered more stable than unusual progressions: the C-major context elicited higher average



Group by Cadence by Key
mean rating
5
4
3
2
1
C—major G—major C—major G—major
G—C
C—G
Untrained
Trained

bass 3.14), but the latter difference was not as large. This seems to confirm the prediction that the listener is more likely to attend to the upper voice of the cadence; $\underline{F}$ (1,47)= 4.28; $\underline{p}$<0.05.

For G-C progressions, cadences led by the soprano voice were given higher stability ratings (4.23) than those led by the bass voice (3.79), and this difference was clearly significant: $\underline{F}$ (1.47)= 21.78; $\underline{p}$<0.001. However, there was also a difference between the two groups of listeners in this regard: $\underline{F}$ (1,47)= 7.62; $\underline{p}$<0.05. Graphically, the ratings given by trained subjects are seen to be higher for the soprano voice only in G-C. This can be seen in Figure 14, where the trained subjects' ratings are depicted by light bars. Untrained listeners' ratings, depicted by shaded bars, show that they apparently considered cadences led by the soprano voice to be more stable, regardless of cadence type.

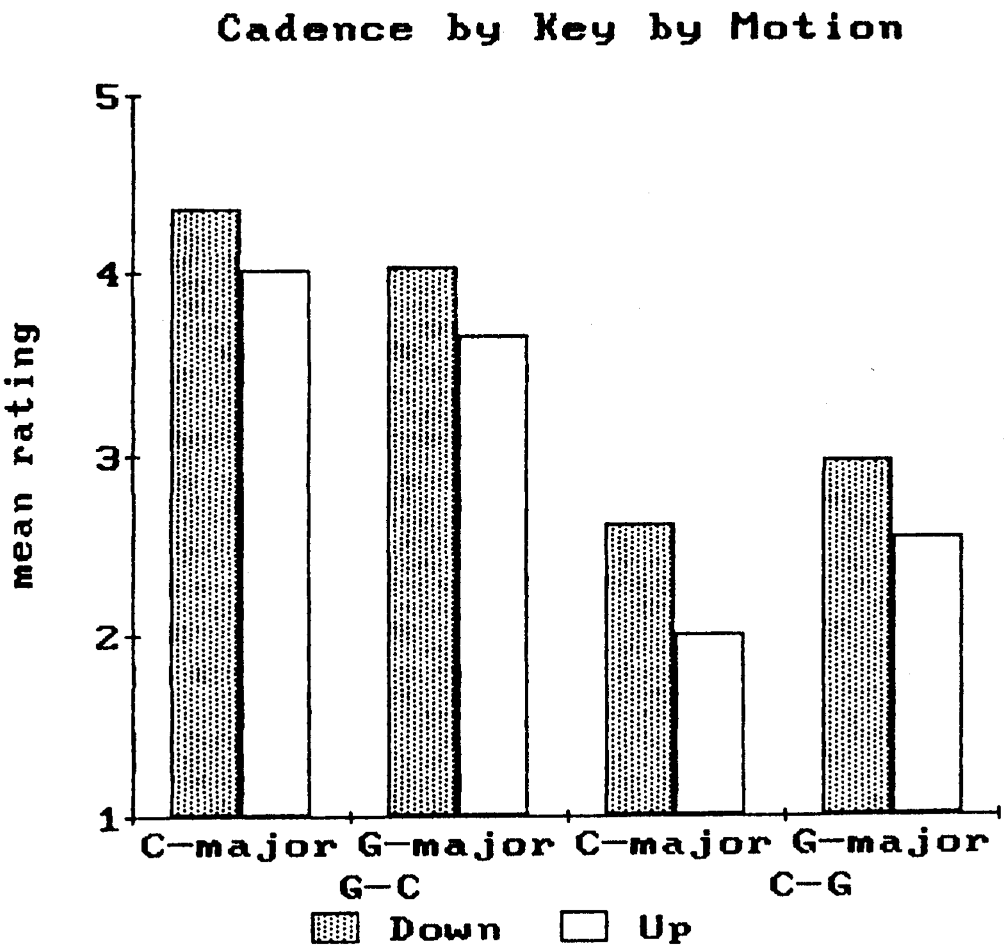
While listeners generally ascribed greater stability to the G-C progressions than to the C-G progressions, as well as to downward-resolving cadences rather than upward-resolving cadences, there was a significant interaction of these factors as well; $\underline{F}$ (1,47)= 5.94; $\underline{p}$<0.05. According to stability ratings, the difference between the two types of progressions was greater for upward-resolving cadences as opposed to downward-resolving cadences (G-C: down 4.19, up 3.83; C-G: down 2.79, up 2.26).

This interaction illustrates the biasing influence of the downward resolution of a cadence, which partially obscures the difference between progressions ending on C versus those ending on G.

While downward-resolving cadences generally obtained higher stability ratings, both in the ascending scale context (down 3.47, up 3.09) and in the descending scale context (down 3.51, up 3.00), there was a significant interaction of scale direction and motion of cadence resolution: F (1,47)= 6.99; $p<0.05$. A descending context may have emphasized the downward resolution of a cadence. This cannot be attributed to training, as the ratings given by the two groups of listeners did not differ significantly in this regard.

When there was upward cadence motion, both the bass-led (3.32) and the soprano-led (3.67) cadences received high stability ratings, and the soprano-led cadences were rated more stable. When the cadence motion was downward, the stability ratings for both the bass-led (2.95) and the soprano-led (3.14) cadences were lower, and while the soprano-led cadences were still rated more stable, the difference between the two was not as large. This interaction of cadence voice and motion of resolution was statistically significant; F (1,47)= 4.36; $p<0.05$.

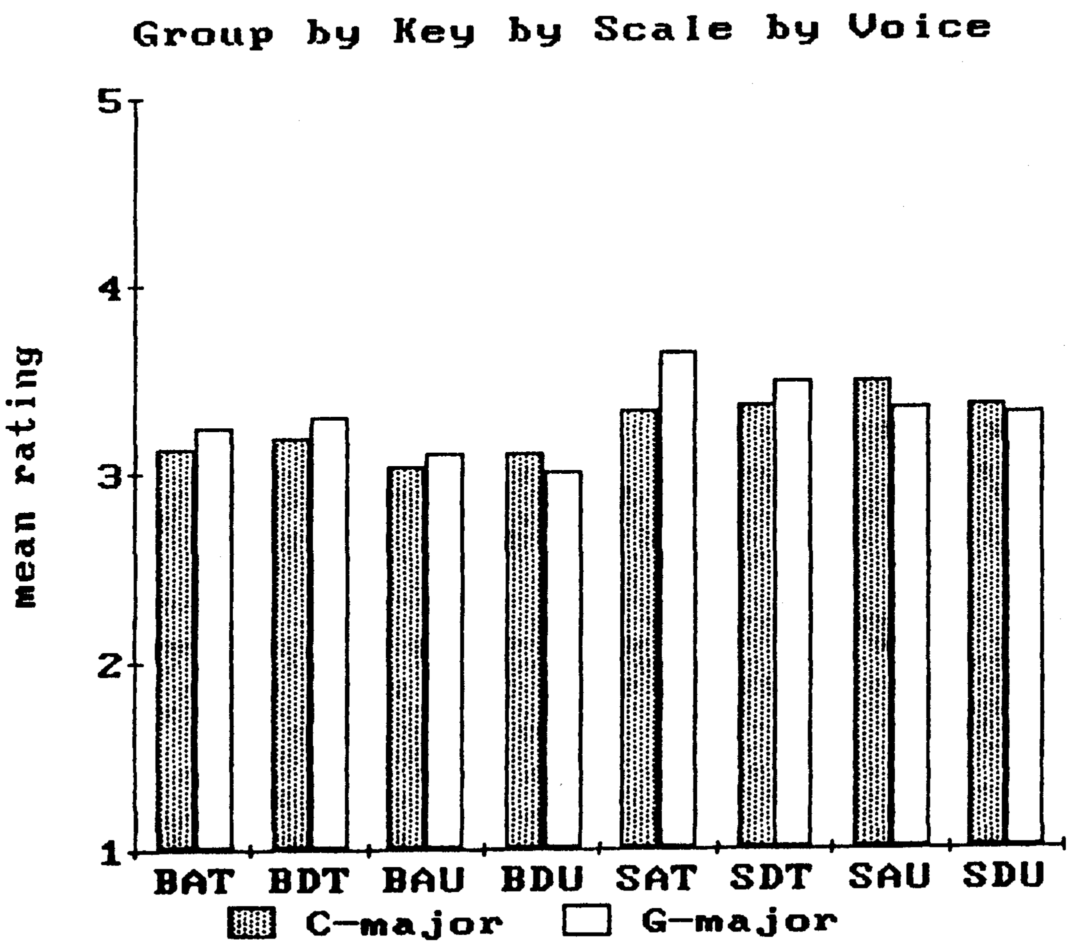
Furthermore, there was a unique interaction of cadence type, scale context, and leading voice; F (1,47)=



Cadence by Key by Motion
mean rating
5
4
3
2
1
C—major
G—major
C—major
G—major
G—C
C—G
Down
Up

11.62; $p<0.05$. Based on their stability ratings, listeners preferred plagal to imperfect cadences, and perfect cadences to non-cadential progressions. They also preferred perfect to plagal cadences, as well as soprano-led cadences to bass-led cadences. The C-G type was rated more stable in the G-major context (soprano 2.85, bass 2.65) than when preceded by a C-major scale (soprano 2.32, bass 2.28). Conversely, the G-C type was rated more stable in the C-major context (soprano 4.43, bass 3.94) than when preceded by G-major (soprano 4.03, bass 3.65). Again, the two groups of listeners did not differ in their judgements, which are in accordance with Common Practice. Both the trained and the untrained group seem to be primarily following the soprano voice.

The interaction of cadence type, key of scale context, and motion of resolution: $F(1,47)= 4.18$; $p<0.05$; is depicted in Figure 15. While G-C progressions were generally found to receive higher ratings than C-G progressions, as shown on the first four bars on the left, the key of the tonal context and the motion of the leading voice in each cadence interacted differently for the two types of progressions. For C-G progressions, higher stability ratings were obtained for upward-resolving plagal cadences (2.53) than for upward-resolving imperfect cadences (1.99), which were rated least stable overall. Downward-resolving plagal cadences (2.97) were rated most stable,

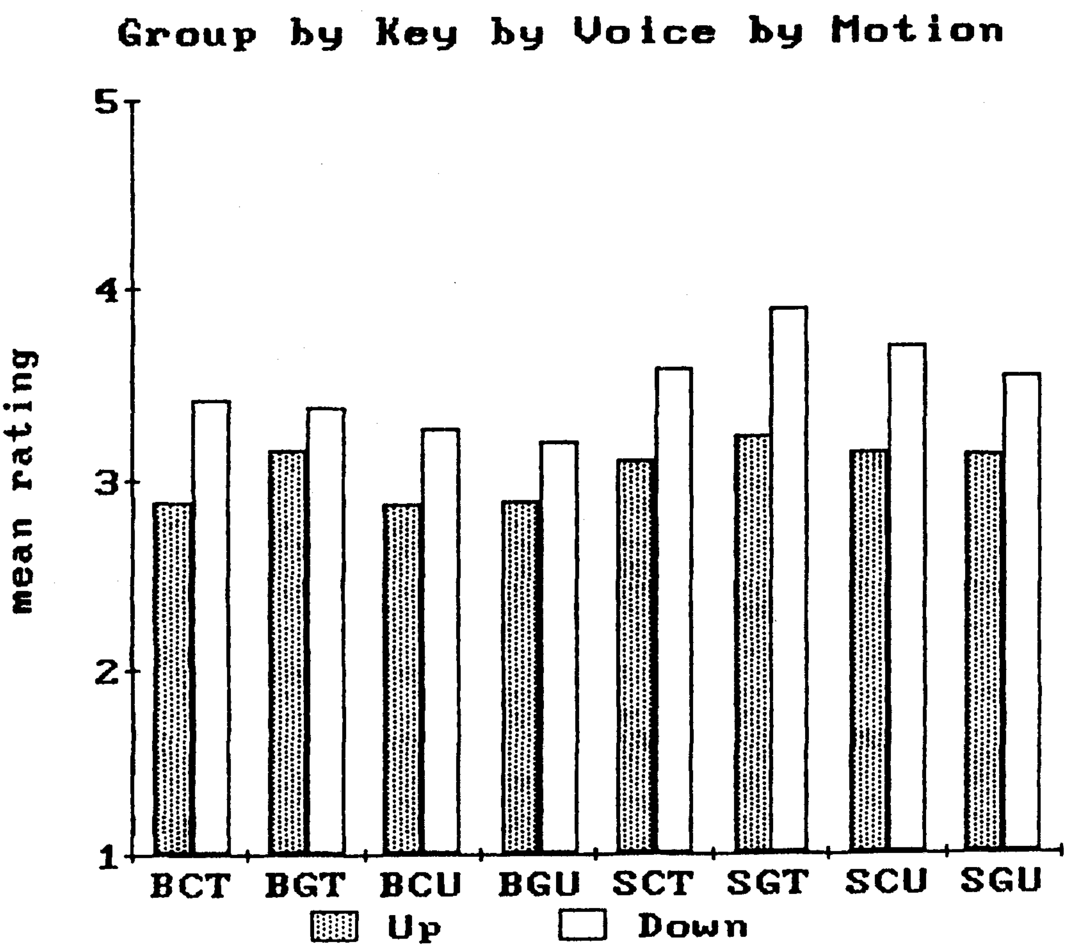


Group by Key by Scale by Voice
mean rating
5
4
3
2
1
BAT
BDT
BAU
BDU
SAT
SDT
SAU
SDU
C-major
G-major

even in comparison to downward-resolving imperfect cadences (2.61). Considering only the G-C progressions, the stability ratings were similarly distributed between upward- and downward-resolving plagal and imperfect cadences, but the differences were not statistically significant. Thus, the highest stability ratings were given to downward-resolving perfect cadences (4.36), as opposed to non-cadential progressions (4.03). Similarly, upward-resolving perfect cadences (4.01) were rated more stable than non-cadential progressions (3.65).

G-C type cadences with downward motion were rated more stable than upward resolving progressions, especially for soprano-led cadences (soprano down 4.30, up 4.16; bass down 4.09, up 3.50). In contrast, for the C-G type, only soprano-led cadences with downward motion were rated highly (3.04), while all other combinations were rated less stable (soprano up 2.13; bass down 2.54, up 2.39). This illustrates the interaction of cadence type, motion of resolution, and leading voice: $\underline{F}$ (1,47)= 35.36; $\underline{p}<0.001$.

When considering the specific interaction of training, scale key, scale direction, and leading voice, we found that only one combination separates clearly from the overall pattern, and this can be easily seen in Figure 16. A letter code is used to identify the factors depicted in the figure. The letter "B" identifies bass-led cadences, while "S" identifies soprano-led cadences. The letter "A"

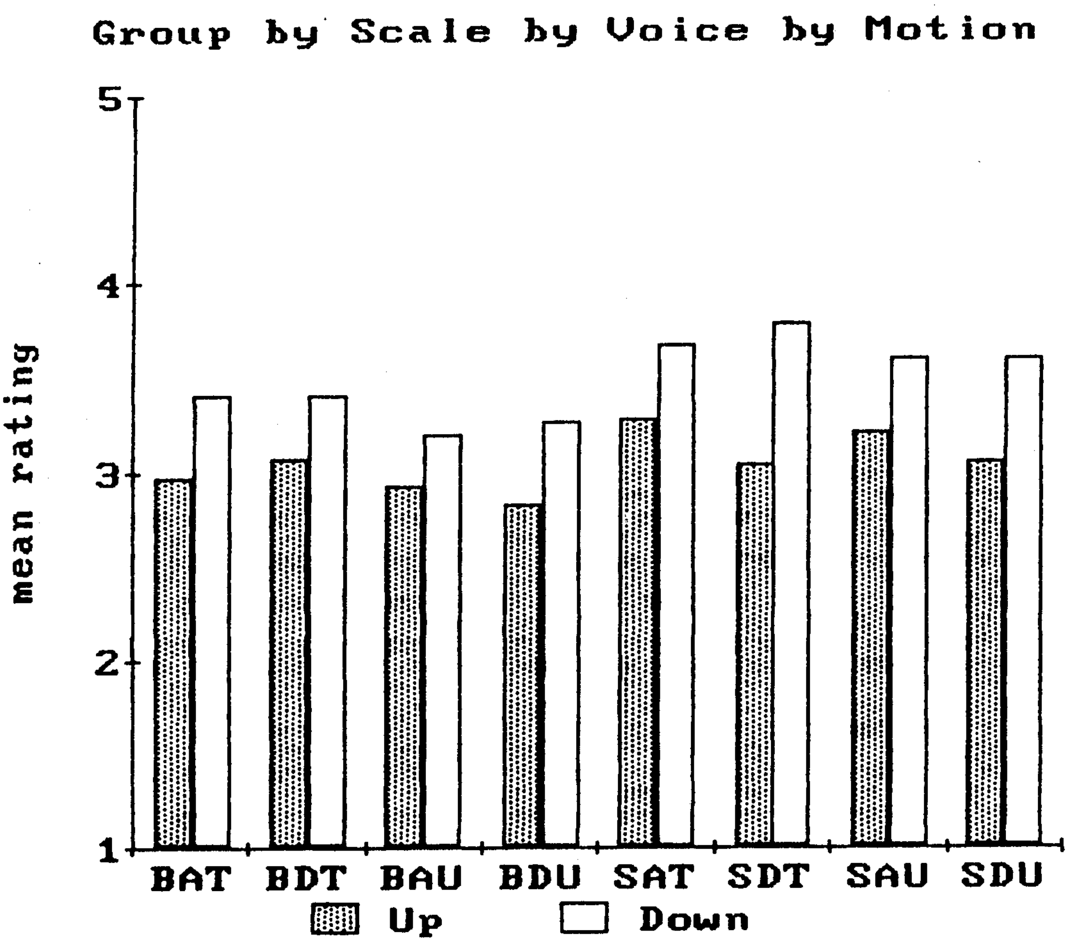


Group by Key by Voice by Motion
5
4
3
2
1
mean rating
BCT
BGT
BCU
BGU
SCT
SGT
SCU
SGU
Up
Down

identifies ascending scales preceding the cadence, while "D" identifies descending scales. Finally, the letter "T" represents the trained group, while "U" represents the untrained group. The shading of the bars indicates the key of the preceding scale: dark bars represent cadences preceded by a C-major scale, while light bars represent cadences preceded by a G-major scale.

Observing the graph (Fig. 16) we find that only trained listeners (bars coded "T") rated cadences with the root tones in the soprano voice (coded "S") clearly higher in the ascending G-major context (white bars) than when the same cadences were preceded by scales in the other key (C-major 3.32, G-major 3.63). The highest average ratings assigned by the untrained group were also given to soprano-led cadences, but their ratings for the C-major context were slightly higher than for G-major (F (1,47)= 6.36; d.f., $p<0.05$).

The trained group showed the strongest preference for cadences with the root note in the soprano voice which resolved downward and were preceded by a G-major scale. This is shown on Figure 17, which uses a code similar to the one described above to represent the factors. Thus, the third unshaded bar from the right, labelled SGT, represents the mean rating 3.88, which was obtained from trained listeners for downward-resolving soprano-led cadences preceded by a G-major cadence. The same chord pair preceded by a C-major

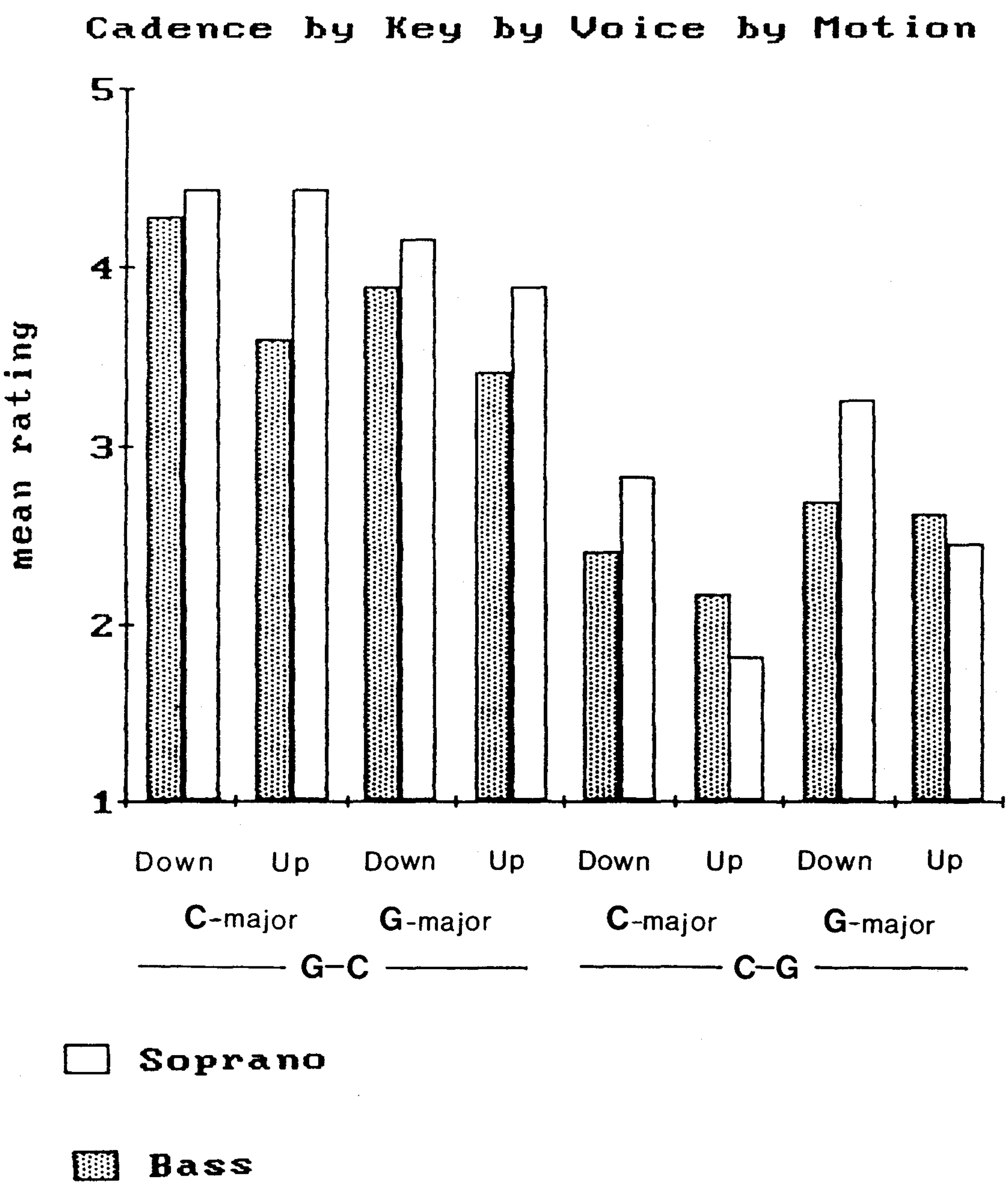

Group by Scale by Voice by Motion
mean rating
5
4
3
2
1
BAT
BDT
BAU
BDU
SAT
SDT
SAU
SDU
Up
Down

scale was rated less stable (3.57). Considering only upward-resolving cadences (indicated by the unshaded bars), the trained group also considered the G-major context more stable than the C-major for cadences led by the bass voice (C-major 3.69: G-major 3.53). These differences were verified as a complex interaction; $\underline{F}$ $(1,47) = 5.79$; $\underline{p} < 0.05$.

The two groups of listeners were influenced differently by the interaction of the scale direction, the leading voice of the cadence, and the direction of resolution: $\underline{F}$ $(1,47) = 4.95$; $\underline{p} < 0.05$. The untrained group gave higher ratings to downward-resolving as opposed to upward-resolving cadences, regardless of whether the cadence was preceded by an ascending or a descending scale. This can be seen on Figure 18, where the shaded bars correspond to upward-resolving cadences.

The trained group gave much higher ratings to downward-resolving soprano cadences preceded by a descending scale than to upward-resolving soprano-led cadences in the same context. This can be observed by comparing the shaded and unshaded bars labelled SDT on Fig. 18. However, when an ascending scale preceded the cadence, the trained group also gave high ratings to upward-resolving, as opposed to downward-resolving soprano-led cadences; a comparison of the shaded and unshaded bars labelled SAT illustrates this effect.

While the average stability ratings given by

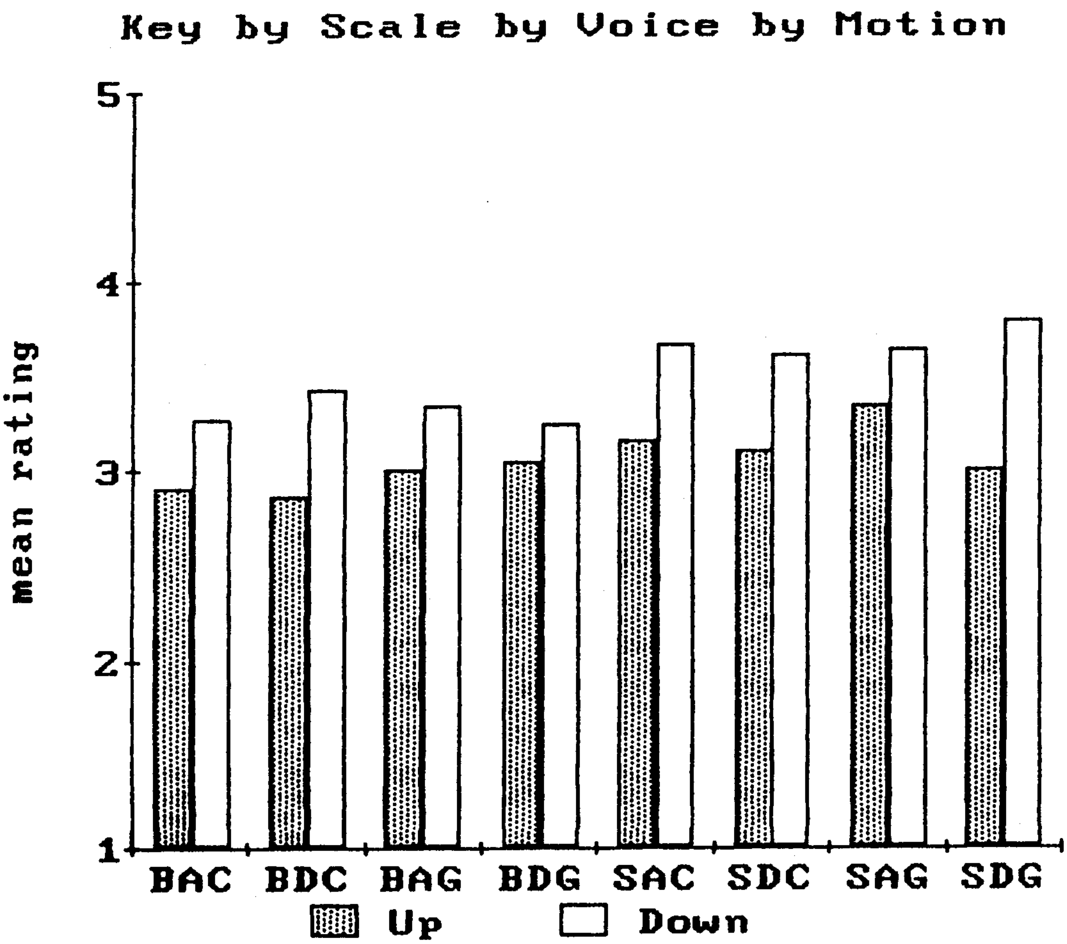


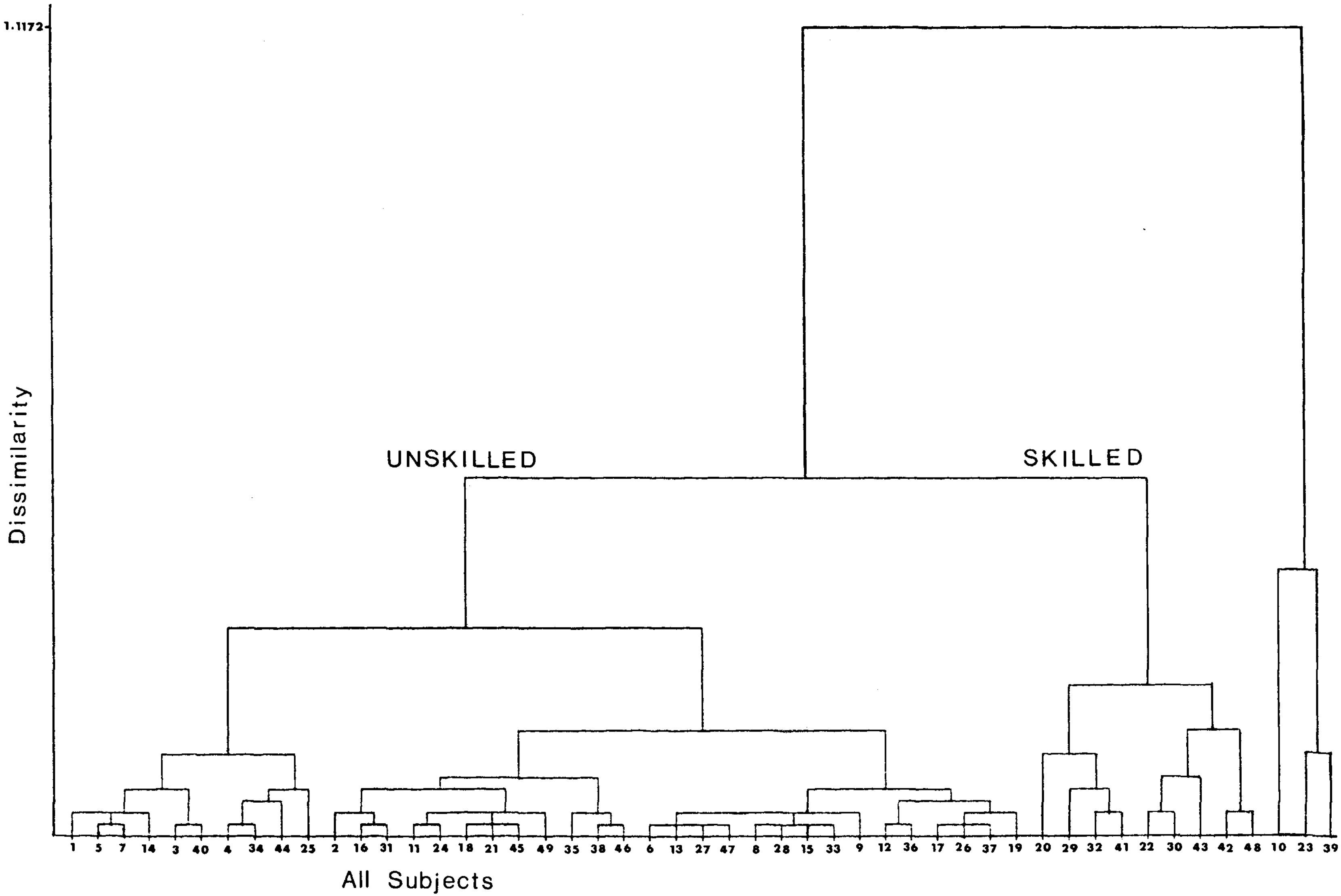
Cadence by Key by Voice by Motion
mean rating
5
4
3
2
1
Down
Up
Down
Up
Down
Up
Down
Up
C-major
G-major
C-major
G-major
G–C
C–G
Soprano
Bass

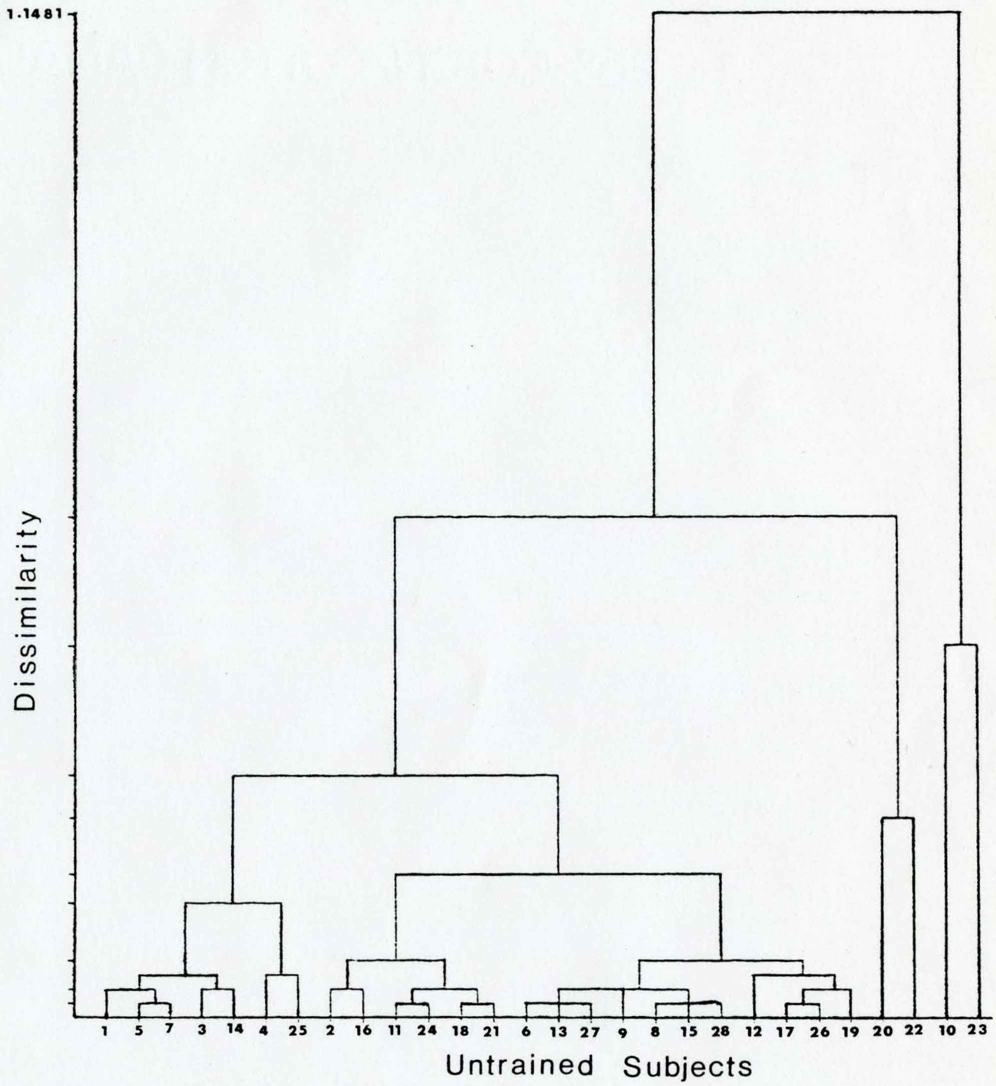
listeners for the C-G cadence type were lower overall than those given for the G-C cadence type, the influence of the other factors was also different for the two types of cadences: $F(1,47) = 6.05$; $p < 0.05$. The mean ratings for the G-C type were lowest for upward-resolving bass-led cadences, both in the C-major context (3.59) and in the G-major context (3.41). Soprano-led cadences were generally rated higher. The difference between the two types of cadences can be easily seen in Figure 19. Here, shaded bars represent bass-led cadences, while unshaded bars represent soprano-led cadences. The first eight bars on the left represent all G-C type cadences, while the eight bars on the right, which are apparently lower, represent C-G cadences. The letters "D" and "U" indicate downward- and upward-resolving cadences, respectively. The letters "G" and "C" stand for the key of the preceding scale.

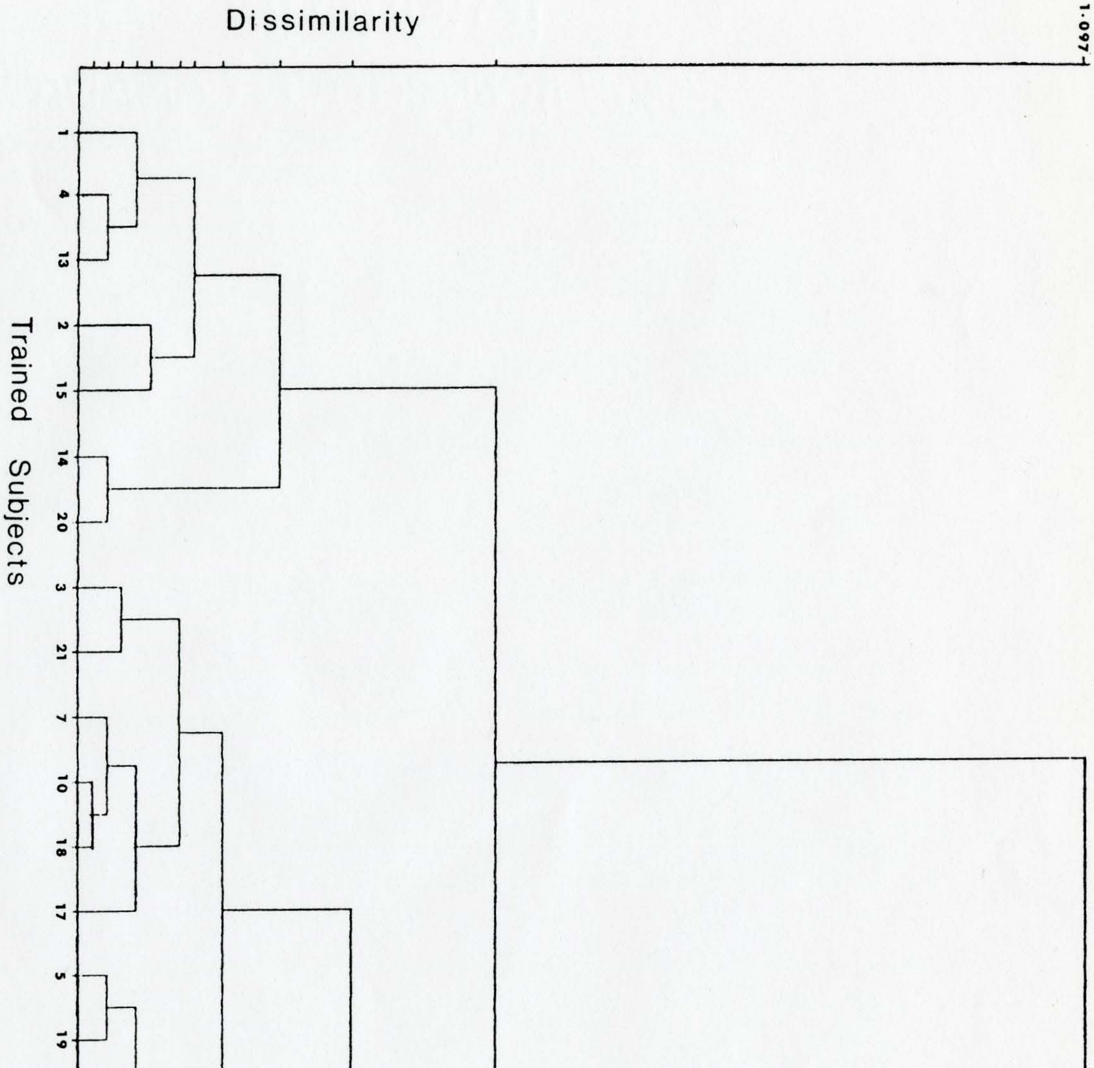
For the C-G type of cadences, the ratings graphed in Figure 19 were highest for soprano-led, downward-resolving cadences in G-major. This type of cadence also showed an unusual reversal of the trend towards soprano-led cadences. Bass-led forms of upward-resolving C-G cadences received higher mean ratings than the soprano-led forms, regardless of the key of the preceding scale.

There was a difference in the way the scale key, scale direction, and cadence resolution affected the stability ratings, according to the position of the root in


Key by Scale by Voice by Motion
mean rating
5
4
3
2
1
BAC
BDC
BAG
BDG
SAC
SDC
SAG
SDG
Up
Down


1.1172
Dissimilarity
UNSKILLED
SKILLED
1 5 7 14 3 40 4 34 44 25 2 16 31 11 24 18 21 45 49 35 38 46 6 13 27 47 8 28 15 33 9 12 36 17 26 37 19 20 29 32 41 22 30 43 42 48 10 23 39
All Subjects

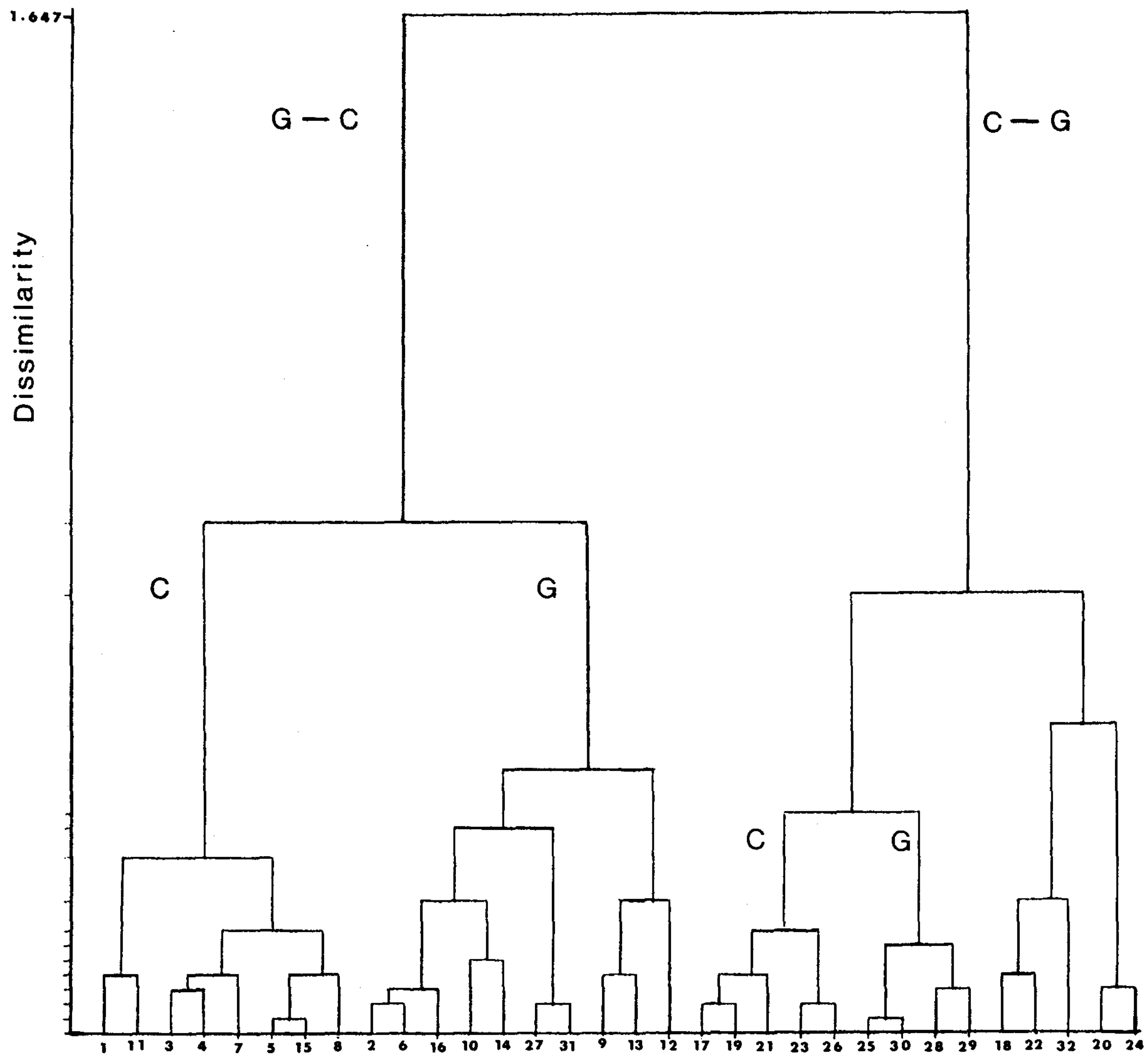

1.1481
Dissimilarity
1
5
7
3
14
4
25
2
16
11
24
18
21
6
13
27
9
8
15
28
12
17
26
19
20
22
10
23
Untrained Subjects


Dissimilarity
1.097
Trained Subjects
1
4
13
2
15
14
20
3
21
7
10
18
17
5
19
8
9
6
16
12
11

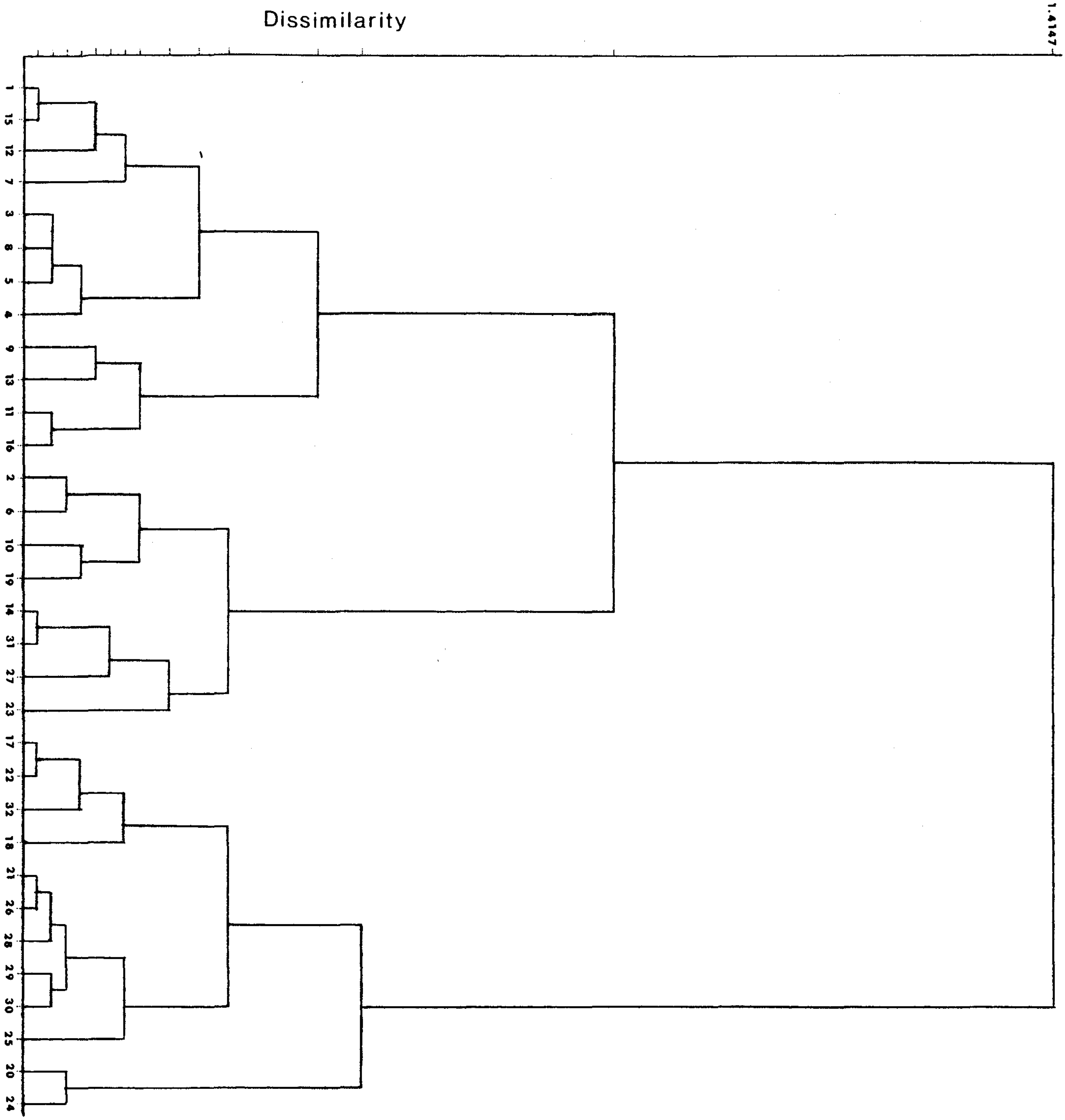
the cadence chords; F (1.47) = 26.06; $p<0.001$. The difference between stability ratings for upward and downward cadences was smallest for cadences with the root in the bass when they were preceded by a descending G-major scale. This can be observed on Figure 20 by comparing the shaded and unshaded bars at label BDG. It is possible that the descending scale context offers insufficient cues for voice-tracking. When the descending G-major scale ended on G3, the listener's attention might have been directed toward the soprano voice, and this would not be a helpful cue when the root tones of the cadence were in fact in the bass voice.

A clustering of subjects across stimuli, shown in Figure 21, divided the participants into three separate groups: a small group of 9 listeners who found many cadences stable (mean rating 3.58), a large group of 37 listeners whose mean stability ratings indicated that they found fewer stable cadences (3.09), and an outlier group with two musically trained and one untrained subject whose concept of stability may be very tolerant (4.30).

By comparison, clustering of the subjects within the groups indicates that the 28 untrained subjects, shown on Figure 22, formed a fairly cohesive cluster (mean rating 3.12), with two outliers (4.27). Looking at the trained group alone, Figure 23 shows that the 21 trained subjects formed a tightly organized grouping of 20 individuals (mean 3.27), with one outlier (4.37). The assignment of the



All Cadences – trained subjects



All Cadences – untrained subjects

subjects to separate groups according to formal training was therefore not always predictive of their performance.

Classification of the stimuli across the 21 trained subjects, also according to the hierarchical agglomerative group average method, divided the cadences into two clusters (shown in Figure 24). The "stable cadence" cluster contained all the G-C cadence type and also two C-G cadences with soprano voicing and downward motion (27 and 31). Both of these cadences were preceded by a G-major scale. The mean rating for the "stable" cluster was 4.04. The "unstable cadence" cluster contained only the remaining C-G cadences (mean 2.40).

A classification of the stimuli across the 28 untrained subjects also revealed a separation of the cadences (shown in Figure 25). The smaller subset of 12 cadences had a mean rating of 2.33 and contained most of the "unstable" C-G cadences. However, four C-G cadences were in the larger "stable" set of cadences (mean 3.74). These additions all had soprano voicing and downward motion: two were preceded by ascending scales (19, 27) and two by descending scales (23, 31). The inclusion of these "unstable" cadences within the stable cluster supports the prediction that untrained listeners are less affected by tonal context. Unlike the trained subjects, they even rated some C-G cadences, preceded by a C-major scale, as stable.

CHAPTER SIX

CONCLUSION AND REVIEW

General Discussion

The series of experiments presented here produced results consistent with the operation of two alternate listening strategies. These strategies were introduced here as alternative ways of perceiving and understanding music. It appears that musicians are able to switch between these two strategies in order to use each where it is most appropriate.

Support for these hypothesized listening strategies was obtained through a cadence-rating paradigm in which either strategy could be effective. The finding that even musically untrained listeners could rate the stability of a musical cadence in a consistent fashion is not in itself surprising. The ubiquitous presence of music in our daily lives and the development of distinctive individual musical preferences are sufficient proof that the realm of musical meaning is not restricted to highly trained specialists. The wide range of performance ability generally observed, where novices and experts can be found at opposite ends of the range, is also not in dispute here. More intriguing are the qualitative differences between the performance of moderately trained individuals and musically untrained novices who were required to assign a stability rating to a cadence.

It has been suggested (Cooper, 1957; Meyer, 1956; Hargreaves, 1986) that when two people listen to the same musical performance, the musically untrained person does not actually hear the same musical phenomenon as the musician. The explanation offered for this diversity is that the two individuals are not listening in the same way. The present dissertation has examined the nature of such supposed listening differences by focussing on trained and untrained individuals' responses to a simple musical fragment. Evidence has been assembled here to show that the differences in the way trained and untrained listeners perceive music may be ascribed to the degree of flexibility with which trained listeners switch between alternate listening strategies.

The idea that the perception of musical elements depends partly on specific formal musical training and partly on general exposure to the music of the culture is a central theme in the study of music perception and cognition (Hargreaves, 1986; Swain, 1986). According to this view, both musicians' and non-musicians' concepts of musical elements should generally be in accordance with the conventions of the music of their culture of origin. In the experiments reported here, musicians' and non-musicians' ratings of musical stability in harmonic progressions do in fact show many similarities, many of which may be explained by a shared cultural understanding of the musical concept of a cadence. More interesting, however, is the prospect that the differences between their ratings may be attributed to the use of different listening strategies.

The present thesis has argued that the way people listen to cadences can reasonably be expected to depend on the way they listen to chords. To begin with, we considered Terhardt's distinction between a basic analytic listening mode, which yields a spectral pitch percept, and a more sophisticated holistic listening mode, which yields a virtual pitch. Holistic listening was expected to result in the perception of a chordal fundamental predictable by rules of harmony. Analytic listening, on the other hand, should result in a highly variable triad pitch percept affected by the relative salience of the various chord components. An investigation to determine what aspects of the cadence are salient to listeners in two different stages of musical training seemed appropriate.

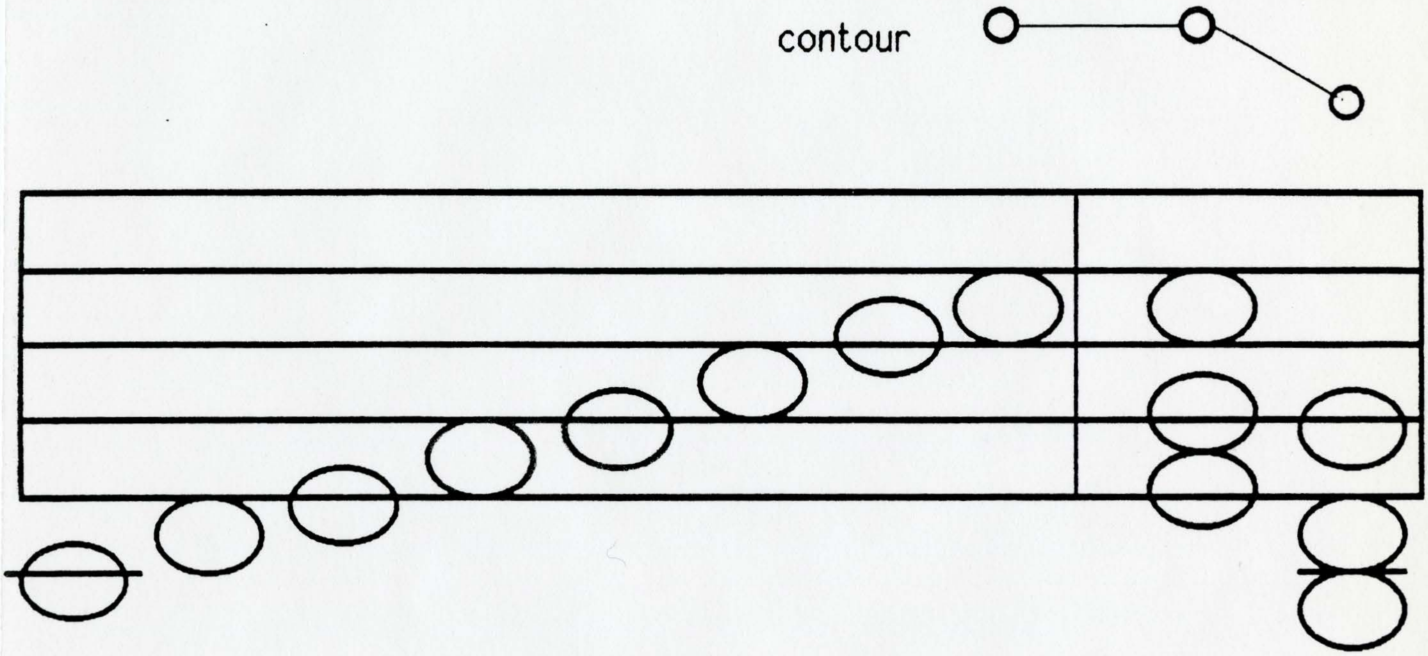
The chord percept constructed by analytic listening has been explained as one dominated by the spectral pitch of one tone in the triad (Terhardt, Stoll, and Seewann, 1982a). On a paired-comparison task, Platt and Racine (In Press) found that inexperienced listeners did not show a consistent preference for any particular triad tone when asked to indicate which tone most resembled the whole triad. Untrained listeners may not be accurately reporting their use of an analytic strategy; they may be using this strategy inadequately or inconsistently; or they may not be using an analytic strategy at all.

In the same study, listeners with some musical training showed a general preference for the highest tone of the triad. About half of a group of highly trained professionals selected the root tone as being most similar to the triad overall, regardless of the position of the root in the triad (Platt and Racine, In press). It appears that the ability to detect the root may be characteristic of the listener's perception of a triad. Both moderately and highly trained listeners appear to use analytic listening adequately, and this strategy can be more narrowly defined.

The first type of analytic listening strategy discussed in this dissertation was simple melody-tracking. This way of listening is primarily oriented towards a specific salient voice, usually the soprano. The results reported here show that the method of choice for untrained listeners appears to be melody-tracking, a rudimentary analytic listening strategy. The second analytic listening strategy suggested here was considered more advanced. This strategy, voice-tracking, allows the listener to focus on either the soprano or the bass voice in the triad sequence. It appears to be the preferred choice for listeners with

Melody-tracking

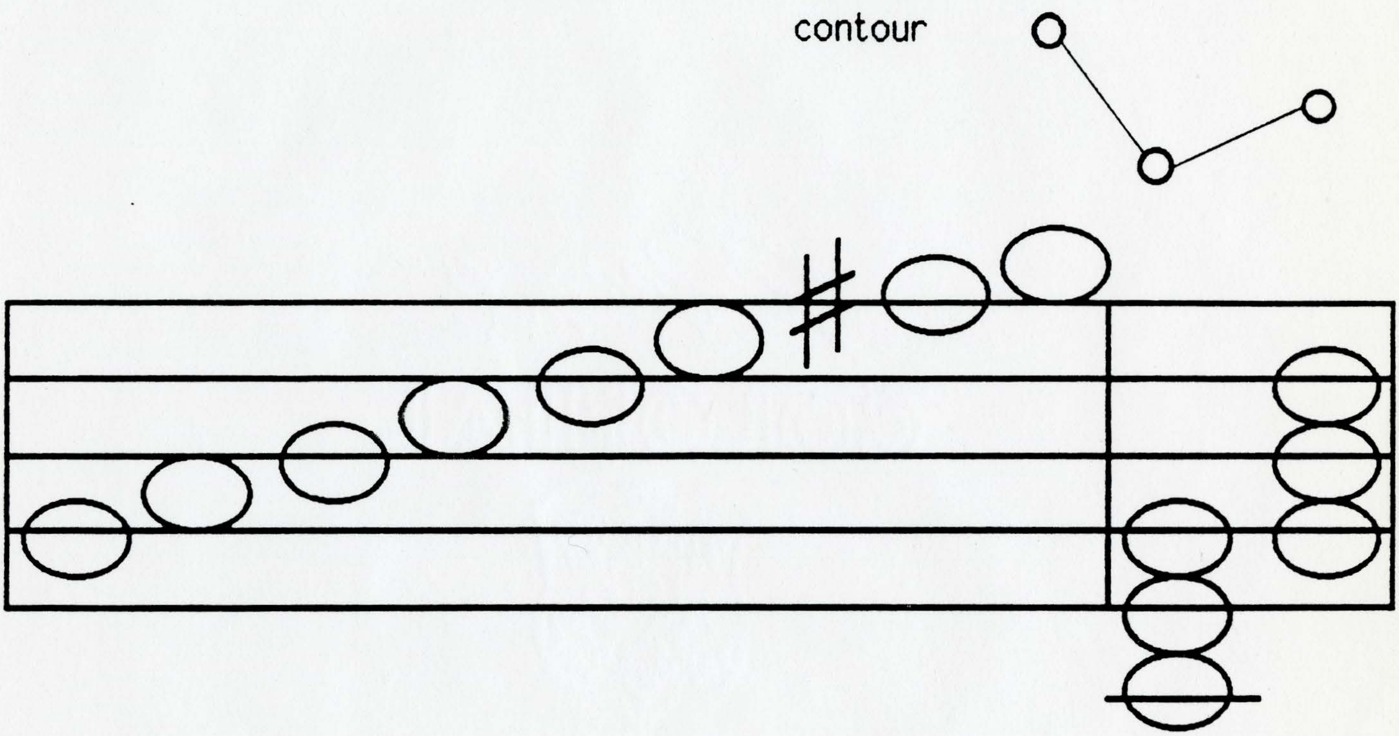
Root note in soprano position:



C-major ascending C - G

Voice-tracking

Root note in bass position:



G-major ascending C - G

some musical training.

Further aspects of the listening situation which may influence the listener's strategy are provided by the harmonic context surrounding the cadence. In the introductory chapter, two alternate ways of determining the structure of chords within a harmonic progression were suggested. One was referred to as constructive or element-focussed listening: this approach was hypothesized to be more resistant to context bias.The second approach, context-oriented or deconstructive listening was expected to be more sensitive to variations in context. In the following, this distinction may be shown to correspond to the two levels of analytic listening described above.

Melody-tracking, a rudimentary analytic perceptual strategy, could be considered a constructive, element-focussed approach, because listeners attending primarily to the melody given by the upper voice of the cadence are relatively insensitive to variations in the harmonic tonal context (see Figure 26). On the other hand, voice-tracking, a more advanced analytic strategy, may be considered a deconstructive, context-oriented approach, where the listener's perceptual focus can change in response to the harmonic context (see Figure 27).

While the melodic contour of the various voices in a series of chords certainly also plays a role for the trained listeners, the influence of melody seems not to be as exclusive for this group, who may attend to other cues in the harmonic context as well. But even a trained listener may not be aware of the reasons for a particular perceptual focus. Peretz (1988) has reported that although the performance of trained musicians often differs from that of untrained listeners when members of both groups are asked to perform typical musical tasks, these performance differences are diminished if the criteria for adequate performance are not strongly cued by the task.

Peretz and Morais (1988) found that all subjects, when asked to indicate the boundary of a musical phrase, tend to group simple melodic fragments in accordance with conventional formal laws of grouping structure. They found that the listeners' subdivision of these musical phrases correlated with conventional musical principles of proximity and separation. Even when the participants were not sure of the purpose of the experimental task, they tended to behave according to cultural norms. In this case, musicians did not perform significantly better than non-musicians.

In the experiments reported here, not all of the individuals who reported a high level of musical training showed a high level of skill on the cadence rating task. This indicates that the purpose of the task in the present paradigm was probably not clearly obvious to all listeners. The experimental paradigm may therefore be considered a test of their abilities to spontaneously employ formal criteria.

Even the musically trained listeners did not always apply formal criteria in a consistent manner.

Interpretation of the Experiments

In all three experiments presented in this thesis, both musically trained and untrained listeners who were requested to rate the stability of simple cadences performed the task with apparent ease. Untrained listeners' ratings were primarily determined by the direction in which the cadence resolved: they considered downward resolution of a cadence more stable than upward resolution. The performance of trained listeners showed an effect of the tonal context not shared by the untrained listeners. Both the key and the direction of the preceding scale affected the stability ratings that trained listeners gave to cadences. Musically trained listeners assigned the highest ratings to cadences that were preceded by a G-major ascending scale.

Musically trained listeners' ratings also consistently showed an effect of the chord component cues. In other words, trained musicians were able to adjust their responses to cadences according to the position and pitch of the tones in the triads. In the first two experiments, trained listeners gave higher stability ratings to cadences ending on chords in the first inversion position, particularly when the cadence was preceded by a G-major scale. In the final experiment, trained listeners' ratings for cadences with the root tones in the soprano voice were higher than their ratings for cadences with the root in the bass voice. These findings are mutually supportive, because when the chords of a cadence are presented in the first inversion position, the root in the soprano is the leading voice.

Untrained listeners also gave higher stability ratings to cadences when the soprano voice led the cadence instead of the bass. There was, however, a greater and more obvious difference between their high ratings for downward-resolving cadences in comparison to their ratings for cadences which resolved upwards. While only musically trained listeners consistently and correctly distinguished medial (incomplete) cadences from final (stable) ones, a cluster analysis of the stability ratings for each of the four cadence types in the second experiment showed that even untrained subjects could distinguish between a category of stable as opposed to unstable cadences. This means that the untrained subjects were able to use the concept of stability consistently.

In the third experiment, clustering of the cadence stimuli according to assigned stability ratings revealed that both trained and untrained listeners gave some of the C-G cadences in a G-major context high stability ratings. They considered this plagal cadence a stable progression. It seems that all participants were able to recognize the match of the tonic chord, G-major, and the tonic note of the G-major scale.

Only untrained listeners gave high stability ratings to a C-G cadence in a C-major context. They appear to have considered this imperfect cadence a stable progression. But we must ask why the untrained listeners' stability ratings did not adequately reflect the two different key contexts. The answer may lie in the fact that no perfect cadences were presented in G-major. The key of G-major was not successfully established simply by presenting the G-major scale and the tonic triad in G.

With regard to the effect of direction of scale context, an apparent contradiction of the findings of the first two experiments, in contrast to the third experiment, might be explained as an outcome of the expanded tonal context. In Experiment 1, the untrained listeners were probably biased by the ascending scale context; they gave higher stability ratings to cadences which resolved downward. This effect may demonstrate their use of the final scale note (G4) as a pitch-matching cue. Careful consideration of the first stimulus set reveals that five of the downward-resolving cadences began with a chord that included G4 as the uppermost note (see Figure 3: cadences 3, 7, 13, 27, and 31). Additionally, two upwards-resolving cadences include the target cue G4 (cadences 8, 32), and two downward-resolving cadences began with a note which could be confused with the target cue (cadences 19, 23). This consideration identifies a subset of nine cadences which might be likely candidates for pitch-matching, where the listener compares the pitch of the last scale note to that of a chord tone.

In contrast to this set of candidates, the stimulus set employed in Experiment 3 only offers four downward-resolving cadences with G4 in the first chord which are easy targets for pitch-matching (see Figure 12: cadences 9, 10, 11, and 12). There are only two additional bass voice cadences with a G4 pitch-matching cue: one upward (8) and one downward-resolving cadence (24). If the untrained listeners were attempting to pitch-match the soprano voice of the chords to the last note of the scale, difficulties may have arisen because there were so few cadences containing G4 available for pitch-matching. There are only six in the second stimulus set, compared to nine in the original stimulus set.

Versatility of the Paradigm

All of the experiments presented in this dissertation used simple cadences embedded in a musical context. These chord sequences can be defined by explicit conventional rules. Nevertheless, even musically untrained listeners were able to make stability judgments about musical fragments, despite their lack of any explicit verbalized knowledge of particular musical terms. Listeners generally preferred plagal to imperfect cadences, and perfect cadences to non-cadential progressions. They also seemed to prefer perfect to plagal cadences, as well as soprano-led cadences to bass-led cadences.

The untrained listeners' knowledge of music could be considered primarily based on their exposure to music in their daily lives, not on awareness of specific rules. Their procedural knowledge would allow untrained participants to acquire and retain rudimentary musical skills, despite their lack of insight into their performance. It is plausible that they should be able to use essential musical concepts effectively, as long as they have attained these concepts through casual exposure to the music typical of our culture (Sloboda, 1985). The trained listeners' knowledge of music would include the formulation and awareness of rules and verbal labels for specific music concepts and relationships.

The perspective of trained and untrained listeners alike must be guided by cues in the musical fragment. Both chordal components and harmonic context can serve as distinctive cues for the perception of harmonic progressions. The musically untrained listeners who participated in these experiments rated musical fragments in broad agreement with Common Practice, although they were unable to identify or explain these concepts. The untrained listeners showed a distinct preference for a descending cadential contour, especially when the cadence was preceded by an ascending scale. The contrast between the motion of the preceding scale and the cadence contour offers a distinct contextual cue. As shown in Fig. 26, the transition from the final note of the preceding scale to the first chord of the cadence yields a perceptual contour, which can be compared to the contour obtained from the transition between the chords in the cadence.

Trained participants, on the other hand, should have sufficient factual knowledge to be able to give an account of their own performance in standard musical terms. In post-experimental follow-up, some musically trained listeners were able to explain the purpose of the experiment and identify some of the chords, cadences, and scales they heard. In comparison to untrained listeners, trained listeners gave higher stability ratings to upward-resolving cadences as well as to cadences containing root position chords. Because trained listeners are able to focus on a melody line in the bass voice, they are able to follow the sequence of roots in the lowest chordal note position. This explains why trained listeners are able to accept a cadence with an ascending contour, as shown in Fig. 27. Despite an ascending melodic contour, the sequence of root notes, C-G, is harmonically stable according to conventional standards.

Indeed, trained listeners' ratings followed conventional standards consistently and accurately. They rated G-C cadences most stable in a C-major context, and C-G cadences most stable in a G-major context. Untrained listeners' ratings, on the other hand, did not consistently show this effect of key context. It appears that they were not always able to utilize available cues to determine a tonal centre. However, averaging across all listeners we find a bias towards cadences ending in a first inversion. If we consider the salience of the tones in a chord, this bias appears quite plausible. When a chord is played in the first inversion position, the root note is in the uppermost voice of the triad and is therefore easy to follow. Therefore, when untrained listeners are participating in the rating task, we can expect cadences with first inversion chords to be rated more stable on average than cadences containing chords in the root position.

A functional hierarchy of cues may underlie listeners' perception of stability and of a tonal centre. As predicted by Brown (1988), the perception of a tonal centre cannot be fully accounted for simply in terms of a hierarchy of isolated pitches. It is also necessary to consider the relationships between a set of pitches and the context in which they are embedded. The musical relationships of pitches may be primarily melodic, as heard in the contour of a musical pattern. The relationship between context and pitches can also be primarily temporal, and be defined by the rhythmic repetition or variation of a specific tonal pattern.

With regard to a cue hierarchy, untrained listeners seemed to be influenced more by pitch proximity cues. These cues can be detected in the cadence itself, or in the preceding scale context. Cues regarding the tonality of the scale context were generally used only imperfectly, if at all, by the untrained listener. Only musically trained listeners were able to pick up sufficient information from the context to assign a stability rating which conformed to the standards of Common Practice tonality. The trained listeners' knowledge of musical structure may allow them to deal with multiple and ambiguous cues in the most efficient and consistent manner.

Evidence of melody-tracking by both trained and untrained listeners was also found. Only trained listeners, however, proved consistently sensitive to the key of the preceding context. The key of the context scale may serve to anchor their perception of the cadence. The musically trained listeners were also affected by the scale direction, which lends further support to the melody-tracking interpretation.

Untrained listeners, on the other hand, rated the cadence more stable when the cadence's root tones were in the uppermost voice and the cadences resolved downward. The findings suggest that although untrained listeners can use some contextual and chord component cues to rate cadence stability, their preferred attentional strategy differs in kind from those utilized by experienced musicians.

Relevance of the Findings

In Chapter One, the introduction to this experimental paradigm, several questions were addressed concerning the nature of the different ways of listening to music. The first question asked whether well-defined, verbalized concepts are necessary for musical perception. If we accept the listeners' performance on this experimental task as an example of musical perception, the answer would appear to be "no". The untrained listener can give evidence of musical perception without any explicit knowledge of formal musical structure.

However, it cannot be said with certainty that the execution of this process is completely independent of verbally stated knowledge. Casual questioning of the listeners appears to support the view that even the most elementary level of musical understanding can be put into words. The musical patterns which seem to be construed as fundamental units of musical meaning are the simple authentic cadences, with their connotations of stability, closure, and downward resolution. These musical concepts are usually verbalized by untrained listeners as "completeness" or "a definite ending". Listeners are able to refer to these concepts before, during and after their performance of the task, and their judgements are likely to be affected by these verbal concepts.

In the second question, it was asked how these putatively basic musical units could form our perceptions of music. By examining listeners' ratings of only a few simple musical patterns we were able to formulate some basic rules. The simple patterns used were cadences. Both trained and untrained listeners come to accept the role and relevance of various cadences when listening to the musical works of our culture. Only serious students of music are confronted with the terms and the structure of this process of learning.

The third question addressed the role of tonal context for the function of these hypothetically basic musical units. The nature of a cadence is co-determined by the tonal context preceding it and by the composition of the chords within it. Listeners were able to infer a melody on the basis of the chord sequence presented to them. This result is complimentary to Bharucha's (1984b) finding that listeners were able to infer a harmonic chord progression on the basis of a melody.

The focus of the fourth question was the assumption that theoretical structure must be internalized during the course of learning a musical skill. While training and dedication are certainly necessary to fully master musical skills, a correct verbalization of the underlying rule may be helpful. This implies that procedural and declarative knowledge must be interwoven for mastery to occur.

The fifth and final question asked whether we could hypothesize that different strategies of perception are employed during different stages of skill acquisition. We found that the use of these different listening strategies did not always correspond to the amount of training the listener had received. However, it was also clear that over the course of serious formal training, listeners learn to apply strategies in a flexible way, while maintaining a high level of consistency in their task performance.

Future and Interdisciplinary Impact

The findings discussed in this dissertation may have implications for further studies in human perception and cognition beyond the immediate realm of the psychology of music. Many of the issues set forth here have been dealt with in the longstanding discussion regarding the role of structure as opposed to function in music theory. Others have been examined in the ongoing discussion of stimulus-oriented versus process-oriented cognitive psychological theories (Morais, 1982).

The relevance of the present findings to contemporary music theory are perhaps more easily illustrated. As Brown (1988) has pointed out, the difference between a structural and a functional model of music perception lies in the relative importance ascribed to contextual cues in each model. Both models grant that the listener is able to make harmonic decisions on the basis of partial cues within the cadence itself, provided that the musical patterns are familiar. However, only a functional model of music perception requires that the listener have additional information from contextual factors external to the cadence in order to fully define tonal structure. Butler and Brown (1984) adopted a functional approach which stresses the process by which various musical cues are evaluated within a given context. This approach does not rely on a fixed internal representation of musical concepts, but rather on the perceptual evaluation process itself. In other words, according to this view, the listener's mind does not contain a "template" for any given concept. Instead of only being measured against some internalized norm, the meaning of any musical fragment will vary according to context.

The present set of experiments offers results that are consistent with a functional model of music perception and cognition. It appears that the listener's understanding of the musical meaning of a cadence can be influenced by three interactive factors: (1) the listener's perceptual attitude or strategy, which is determined at least in part by prior experience; (2) the tonal context of the musical fragments evaluated, in this case consisting of a scale; and (3) the tonal composition of the musical fragments used, namely the arrangement of the chords played in the cadences.

These three factors determine the content and form of the musical meaning established by the listener. For the untrained listener, the meaning of a cadence appears primarily determined by the tradition of his or her culture of origin. For the trained listener, a complex hierarchy of tonal principles may gradually be established over the course of training. The musical meaning of a cadence may change in response to this set of principles, much in the same way that a child's concepts of language and behaviour are gradually modulated into those of an adult's.

Further empirical study of listeners' concepts of musical meaning may allow us to verify the existence of such a hierarchy of principles. But it is also possible that strategies take precedence over principles. Listeners may not be actually oriented towards a cognitive hierarchy, but may simply have become accustomed to using a particular perceptual style. Longitudinal studies of individual development and change might prove particularily valuable in this regard. Of particular relevance is the field of skills deficits.

The idea that behaviour, perception, and cognition are intertwined in normal task performance has been extensively considered, both theoretically (Scheerer, 1984; Prinz, 1984) and empirically. Phillips and Carr (1987) have suggested that motor and cognitive strategies may share a common basis. Both are dependent on a planning and sequencing component. Further investigation of such a hypothetical common pathway might permit an overview of the cooperation of neural structures underlying perceptual, cognitive, and motor processes, and the breakdown of this cooperation when the individual experiences a skill deficit.

The study of the phenomenon of focal dystonias in musicians, considered to exemplify a breakdown in functioning at the interface of perception and performance, is one example. Newmark and Hochberg (1987) describe the onset and extent of focal dystonias in musicians as a syndrome which entails a chronic disability, exhibited in the form of painful muscle spasm in the afflicted body part, usually restricted to a particular activity. A pianist, for example, may develop a dystonia of the right hand, observable only during performance, which does not occur during writing or during any other manual skill.

Future studies of the performance of professional musicians could examine the circumstances under which focal dystonias arise. By identifying the musical elements that correspond to particular elements of motion, it should be possible to predict specific moments of strain and eventual breakdown. These critical points could be related to the musician's technical approach. Furthermore, alternate strategies could be developed to decrease the demands of performance. These novel strategies would have to allow the same ease of expression during performance as the old, overly stressful ones. The new strategies employed would be more than a rigid collection of rules. The strategies would probably depend on the formulation of clearly defined goals. Such goals would probably have to be defined within the musician's framework of musical meanings.

Looking at the other end of the skill continuum, we could investigate strategy shifts observable in the behaviour of children who are learning a musical instrument for the first time. While learning how to play an instrument, a child also learns the language and the rules of musical meaning. By repeatedly requesting cadence ratings over the course of years we may be able to observe the shift from a rudimentary analytic strategy to an advanced analytic strategy.

A modification of the original cadence rating paradigm would allow a more detailed observation of the proposed strategies. If the participant is asked to hum along with the cadence while listening to the musical fragments, it should be possible to identify the focus of the listener's attention. This would allow us to directly observe the listener's attempts to follow the melody of the soprano or bass voice. If they are capable of correctly identifying the root of each chord in a cadence, some listeners should be able to hum along in tune with the root (Platt and Racine, In Press).

Future work may lead to a more complete understanding of the processes involved in musical listening and performance. The present examination of listeners' ratings of cadences has confirmed the possibility of two different ways of listening. Conceptualized as strategies, these ways of listening illustrate the flexibility of human perception and cognition. Examining the limits of this flexibility remains a task for future exploration.

REFERENCES

Aldenderfer, M.S. and Blashfield, R.K. (1984). Cluster Analysis. Beverly Hills, California: Sage Publications.

Anderson, J.R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Baur, J. (1985). Music Theory Through Literature, Vol. 1. Englewood Cliffs, N.J.: Prentice Hall.

Berry, W. (1987). Structural Functions in Music. New York: Dover Publications.

Benjamin, T., Horvit, M., and Nelson, R. (1975). Techniques and Materials of Tonal Music. Boston: Houghton Mifflin.

Bharucha, J.J. (1984a). Event hierarchies, tonal hierarchies, and assimilation: a reply to Deutsch and Dowling. Journal of Experimental Psychology: General, 113, 421-425.

Bharucha, J.J. (1984b). Anchoring effects in music.

Cognitive Psychology, 16, 485-518.

Bharucha, J.J. and Krumhansl, C.L. (1983). The representation of harmonic structure in music: hierarchies of stability as a function of context. Cognition, 13, 63-82.

Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. Music Perception. 5, 219-250.

Butler, D. and Brown, H. (1984). Tonal structure versus function: studies of the recognition of harmonic motion. Music Perception, 2, 6-24.

Cannel. W., and Marx, F. (1982). How to Play the Piano Despite Years of Lessons. New York: Doubleday and Company.

Castellano, M.A., Bharucha, J.J., and Krumhansl, C.L. (1984). Tonal hierarchies in the music of North India. Journal of Experimental Psychology: General, 113, 394-412.

Charness, N., Clifton, J., and MacDonald, L. (1988). Case study of a musical "mono-savant": a cognitive-psychological focus. In: L. Obler and D. Fein (Eds.), The Exceptional Brain: Neuropsychology of Talent and Other Abilities. New York: The Guilford Press.

Clifton, T. (1983). Music as Heard: a study in applied phenomenology. New Haven: Yale University Press.

Cook, N. (1987). A Guide to Musical Analysis. London: J.M. Dent and Sons, Ltd.

Cooper, G. (1957). Learning to Listen. Chicago: The University of Chicago Press.

Copland, A. (1939). What to Listen For in Music. New York: McGraw Hill.

Cuddy, L.L. and Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. Perception and Psychophysics, 41, 609-620.

Deliege, I. (1987) Grouping conditions in listening to music: an approach to Lerdahl and Jackendoff's grouping preference rules. Music Perception, 4, 325-360.

Divenyi, P.L., Efron R., and Yund E.W. (1977). Ear dominance in dichotic chords and ear superiority in frequency discrimination. Journal of the Acoustical Society of America, 62, 624-632.

Dowling, W.J. (1986). Context effects on melody recognition: scale-step versus interval representations. Music Perception, 3, 281-296.

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

Dowling, W.J., K. M. Lung, and S. Herrbold (1987). Aiming attention and time in the perception of interleaved melodies. Perception and Psychophysics, 41, 642-656.

Edwards, B. (1979). Drawing on the Right Side of the Brain. Los Angeles: J.P. Tarcher, Inc.

Efron, R. (1985). The central auditory system and issues related to hemispheric specialization. In: M.L. Pinheiro and F.E. Musiek (Eds.) Assessment of Central Auditory Dysfunction: Foundations and Clinical Correlates. Baltimore: Williams and Wilkins.

Everitt, B. (1980). Cluster Analysis. New York: Halsted Press.

Fraisse, P. (1982). Rhythm and tempo. In: D. Deutsch (Ed.) The Psychology of Music. New York: Academic Press.

Frances, R. (1984). La Perception de la Musique. Paris: Librairie Philosophique J. Vrin.

Fienberg, S.E. (1980). The analysis of cross-classified categorical data. Cambridge: M.I.T.

Gardner, H. (1973). The Arts and Human Development. New York: Wiley.

Gaede, S.E., Parsons, O.A., and Bertera, J.H. (1978). Hemispheric differences in music perception: Aptitude vs experience. Neuropsychologia, 16, 369-373.

Gordon, A.D. (1981). Classification: Methods for the exploratory analysis of multivariate data. New York: Chapman and Hall.

Gordon, H.W. (1983). Music and the right hemisphere. In: A.W. Young (Ed.) Functions of the Right Cerebral Hemisphere. New York: Academic Press.

Grossman, M., Shapiro, B.E., and Gardner, H. (1981) Dissociable musical processing strategies after localized brain damage. Neuropsychologia, 19, 425-433.

Hargreaves, D. (1986). The Developmental Psychology of Music. Cambridge: Cambridge University Press.

Henschen, S.E. (1926). On the function of the right hemisphere of the brain in relation to the left in speech, music and calculation. Brain, 49, 110-123.

Hindemith, P. (1943). Traditional Harmony. New York: Schott and Co., Ltd.

Hindemith, P. (1969). A Composer's World. London: Peter Smith.

Houtsma, A.J.M. and Goldstein, J.L. (1972). The central origin of the pitch of complex tones: evidence from musical interval recognition. Journal of the Acoustical Society of America, 51, 520-529.

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

Jones, M.R. (1982). Music as a stimulus for psychological motion: Part II. An expectancy model.

Psychomusicology, 2, 1-13.

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

Judd, T. (1988). The varieties of musical talent. In: L.K. Obler and D. Fein (Eds.) The Exceptional Brain: Neuropsychology of Talent and Special Abilities. New York: Guilford Press.

Keppel, G. (1982). Design and Analysis: a Researcher's Handbook. Englewood Cliffs: Prentice-Hall.

Kerman, J. (1985). Contemplating Music: Challenges to Musicology. Cambridge: Harvard University Press.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. Canadian Journal of Psychology, 15, 166-171.

Kitson, C.H. (1946). Elementary Harmony. Oxford: Clarendon Press.

Krumhansl, C.L. and 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. and 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.

Krumhansl, C.L. (1983). Perceptual structures for tonal music. Music Perception, 1, 28-62.

Langer, S. (1967). Mind: An Essay on Human Feeling. Baltimore: Johns Hopkins University Press.

Lehrdahl, F. and Jackendoff, R. (1983). A Generative Theory of Tonal Music. Cambridge: The MIT Press.

Lucci, D., Fein, D.. Holevas, A., and Kaplan, E. (1988). Paul: a musically gifted boy. In: L. Obler and D. Fein (Eds.), The Exceptional Brain: Neuropsychology of Talent and Other Abilities. New York: The Guilford Press.

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

Meyer, L.B. (1967). Music, the Arts, and Ideas. Chicago: The University of Chicago Press.

Morais, J. (1982). The two sides of cognition. In: J.Mehler and E.C.T. Walker (Eds.), Perspectives on Mental Representation. New Jersey: Hillsdale.

Moore. B.C.J. (1982). An Introduction to the Psychology of Hearing. London: Academic Press.

Neisser, U. (1967). Cognitive Psychology. San Fransisco: W.H. Freeman.

Neumann, O. (1984). Automatic processing: a review of recent findings and a plea for an old theory. In: W. Prinz and A.F. Sanders (Eds.) Cognition and Motor Processes. New York: Springer-Verlag.

Newmark, J. and Hochberg, F.H. (1987). Isolated painless manual incoordination in musicians. Journal of Neurology, Neurosurgery, and Psychiatry, 50, 291-295.

Nisbett. R. and Ross, L. (1980). Human Inference: Strategies and Shortcomings of Social Judgement. Englewood

Cliffs: Prentice-Hall.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia, 9, 97-113.

Palmer, C. and 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. and 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.

Peretz, I. and Morais, J. (1987). Analytic processing in the classification of melodies as same or different. Neuropsychologia, 25, 645-652.

Peretz, I. and Morais, J. (1988). Is Music Special? Unpublished Manuscript, Universite de Montreal.

Phillips, A.G. and Carr, G.D. (1987). Cognition and the basal ganglia: a possible substrate for procedural knowledge. The Canadian Journal of

Neurological Sciences, 14, 381-385.

Platt, J. and Racine, R. (1990). In Press.

Plomp, R. (1977). Aspects of Tone Sensation. New York: Academic Press.

Pohl, P. (1983). Central Auditory Processing. V. Ear advantages for acoustic stimuli in baboons. Brain and Language, 20, 44-53.

Posner, M.I. and Cohen, Y. (1980) Attention and the control of movements. In: G.E. Stelmach and J. Requin (Eds.) Tutorials in Motor Behaviour. Amsterdam: North-Holland.

Posner, M.I. and Snyder, C.R.R. (1975) Attention and cognitive control. In: R. Solso (Ed.) Information Processing and Cognition: The Loyola Symposium. Potomac, MD: Erlbaum.

Priesing, D. and Tecklin, L. (1959). Language of the Piano: A Workbook in Theory and Keyboard Harmony. New York: Carl Fischer, Inc.

Prinz, W. (1984). Modes of linkage between perception and action. In: W. Prinz and A.F. Sanders

(Eds.) Cognition and Motor Processes. New York: Springer-Verlag.

Roberts, L.A. and Shaw, M.L. (1984). Perceived structure of musical triads. Musical Perception, 2, 95-124.

Rosen, C. (1972). The Classical Style. New York: W. W. Norton.

Rosner, B.S. and Meyer, L.B. (1986). The perceptual roles of melodic process, contour, and form. Music Perception, 4, 1-40.

Rowell, L. (1983). Thinking About Music. Amherst: University of Massachusetts Press.

Sacks, O. (1987). The man who mistook his wife for a hat. New York: Harper and Row.

Scheerer, E. (1984). Motor theories of cognitive structure: a historical review. In: W. Prinz and A.F. Sanders (Eds.) Cognition and Motor Processes. New York: Springer-Verlag.

Schweiger, A. (1985). Harmony of the spheres and the hemispheres: the arts and hemispheric specialization. In: D.F. Benson and E. Zaidel (Eds.) The Dual Brain:

Hemispheric Specialization in Humans. New York: The Guilford Press.

Schweiger, A. and Maltzman, I. (1985). Behavioural and electrodermal measures of lateralization for music perception in musicians and non-musicians. Biological Psychology, 20, 129-145.

Shapiro, B.E., Grossman, M., and Gardner, H. (1986). Selective musical processing deficits in brain damaged populations. Neuropsychologia, 19, 161-169.

Seashore, C.L. (1938). Psychology of Music. New York: McGraw Hill.

Serafine, M.L. (1983). Cognition in music. Cognition, 14, 119-183.

Sidtis, J.J. and Bryden, M.P. (1978). Asymmetrical perception of language and music: Evidence for independent processing strategies. Neuropsychologia, 16, 627-632.

Sloboda, J.A. (1985). The Musical Mind: The cognitive psychology of music. New York: Oxford University Press.

Squire, L.R. (1986). Mechanisms of memory. Science, 232, 1612-1619.

Swain, J.P. (1986). The need for limits in hierarchical theories of music. Music Perception, 4, 121-148.

Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds. Perception and Psychophysics, 23, 483-492.

Terhardt, E., Stoll, G., and Seewann, M. (1982a). Pitch of complex signals according to virtual-pitch theory: tests, examples, and predictions. Journal of the Acoustical Society of America, 71, 671-678.

Terhardt, E., Stoll, G., and Seewann, M. (1982b). Algorithm for extraction of pitch and pitch salience from complex tonal signals. Journal of the Acoustical Society of America, 71, 670-688.

Trehub, S.E., Bull, D., and Thorpe, L.A. (1984). Infants' perception of melodies: the role of melodic contour. Child Development, 55, 821-830.

Wexler, E., Halwes, T., and Heninger, G.R. (1981). Use of a statistical inference criterion in drawing inferences about hemispheric dominance for language function from dichotic listening data. Brain and Language, 13, 13-18.

Wilson, F.R. (1987). Tone deaf and all thumbs? Toronto: Random House (Vintage Books).

Wood, A. (1975). The Physics of Music. London: Chapman and Hall (Science Paperbacks).

Zatorre, R.J. (1984). Musical perception and cerebral function: a critical review. Music Perception, 2, 196-221.

Zatorre, R.J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. Neuropsychologia, 23, 31-41.

Zuckerkandl, V. (1956). Sound and Symbol. Princeton, New Jersey: Princeton University Press.