1999

Music perception and cognition as a basis for an ear-training approach

Zila Shwartz Medini
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Medini, Zila Shwartz, "Music perception and cognition as a basis for an ear-training approach" (1999). *Master's Theses*. 1948.

DOI: <https://doi.org/10.31979/etd.buy8-5uh4>

https://scholarworks.sjsu.edu/etd_theses/1948

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI[®]
800-521-0600

MUSIC PERCEPTION AND COGNITION
AS A BASIS FOR
AN EAR-TRAINING APPROACH

A Thesis

Presented to

The Faculty of the School of Music and Dance

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Zila Shwartz Medini

December 1999

UMI Number: 1397746

UMI Microform 1397746

Copyright 2000 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

©1999

Zila Shwartz Medini

ALL RIGHTS RESERVED

APPROVED FOR THE SCHOOL OF MUSIC AND DANCE

A handwritten signature in cursive script, reading "Theodore D. Lucas", written over a horizontal line.

Dr. Theodore D. Lucas (Chair)

A handwritten signature in cursive script, reading "Brian Belet", written over a horizontal line.

Dr. Brian Belet

A handwritten signature in cursive script, reading "Brent Heisinger", written over a horizontal line.

Dr. Brent Heisinger (Emeritus)

APPROVED FOR THE UNIVERSITY

A handwritten signature in cursive script, reading "William Fish", written over a horizontal line.

ABSTRACT

MUSIC PERCEPTION AND COGNITION AS A BASIS FOR AN EAR-TRAINING APPROACH

by Zila Shwartz Medini

While the research of music perception and cognition focuses on isolated aspects and small controllable units, the focus of Western tonal music is on relationships, both on small scale and on large scale. The purpose of this thesis is to investigate whether, and to what extent, teachers of college-level ear training can benefit from such research. After a brief and general discussion of perception, research is reviewed, focusing on the aspects that are most pertaining to the musical syntax of the Common Practice: pitch and time. Research has shown some perceptual principles that pertain to various musical aspects. Some perceptual categories are more stable in memory, and perceived more directly and accurately (such the beat and the tonic); pattern recognition and the extraction of meaning (or tonal implications) enhances memory; attention can be directed in different ways. After stating the goals of ear training, these findings and others are used for the development of an efficient and meaningful ear training approach.

Acknowledgements

I would like to thank the faculty of the School of Music and Dance at San José State University for making this project possible. Their warm support and open mindness enabled me to pursue my interest and tread my own path.

I am indebted to Professor Bat-Sheva Rubinstein, from the Rubin Academy of Music in Tel-Aviv University. Her ear training approach was an inspiration for me to explore the realm of music perception.

Special thanks to Dr. William George, whose door was always open and who more then once was willing to spend his time digging through his books and tracing hard to find sources.

My committee, Dr. Theodore Lucas, Dr. Brian Belet and Dr. Brent Heisinger were always ready to read or discuss and their feedback and comments were greatly appreciated.

Writing this thesis was a long process. I am especially grateful to Dr. Lucas who read every draft and every version and whose continuous support helped me persist and bring this work to completion.

Finally, I would like to thank my family: my husband, Yotam, who not only helped with the technical aspects of writing, but also cooked and baby-sat; my kids Dori, Amalia and Hillel, who kept sneaking into the computer room to ask: “are you done, Mom?” Now I am done.

Table of Contents

1	Introduction	1
1.1	Thesis Overview	4
1.2	Perception and Cognition	7
2	Single-Pitch Perception	15
2.1	Pitch Retention by Nonpossessors of Absolute Pitch	16
2.2	Absolute Pitch	18
2.2.1	Mechanism of Absolute Pitch	19
2.2.2	Speed and Accuracy of Absolute Pitch	20
2.2.3	Perceptual Principles Pertinent to Absolute Pitch	22
2.2.4	Types of Errors in Judgments of Absolute Pitch	26
2.2.5	Additional Absolute Pitch-Skills	29
2.2.6	Acquisition of Absolute Pitch	32
2.2.7	Benefits of Absolute Pitch	40
2.2.8	Drawbacks of Absolute Pitch	42
2.2.9	Summary	43
3	Perception of Melodies	45
3.1	Research Methods	46
3.2	Interval Perception	49
3.3	Perception of Three-Tone Sequences	54
3.4	Non-Interval Aspects of Melody Perception	55
3.5	The Recall Method	62
4	Perception of Simultaneities	77
4.1	Perception of “Vertical” Simultaneities	78
4.1.1	Consonance and Dissonance	78
4.1.2	Perception of Harmonic Intervals	90
4.1.3	Chord Perception	104
4.2	Perception of “Horizontal” Simultaneities	121

4.2.1	Channelling Attention	121
4.2.2	Attending Musical Simultaneities	132
5	The Influence of Timbre on Pitch Perception	134
5.1	Physical Attributes of Sound Influencing Timbre	136
5.2	Interaction of Timbre and Pitch	138
5.3	Timbre as an Organizing Factor	142
6	Perception of Tonality	147
6.1	Tonal-Hierarchy Theory	148
6.1.1	Psychological Distances	149
6.1.2	Key Relations	153
6.1.3	Criticism	155
6.2	Intervallic-Rivalry Theory	158
6.3	Combining the Two Theories	165
7	Perception of Musical Time	167
7.1	Introduction	167
7.2	Definitions	170
7.3	Perception of Temporal Aspects of Music	178
7.3.1	Perception of Durations	178
7.3.2	Perception of Pulses	181
7.3.3	Perception of Meter	184
7.3.4	Perception of Rhythmic Groups	190
7.3.5	Perception of Hyper-Measures and Large-Scale Rhythmic Groups	198
8	Summary and Conclusions	201
8.1	Summary of Research	201
8.2	Goals of Ear Training	208
8.3	Research Based Ear Training Approach	213
8.3.1	Professor Rubinstein's Aural Drill	214
8.3.2	Isolated Intervals and Chords	220
8.3.3	Sight Singing	223
8.3.4	Dictation	232
8.4	Summary	238
	Bibliography	240

List of Figures

1.1	Law of proximity	12
1.2	Law of similarity	12
3.1	Recognition of three-tone patterns	55
3.2	Transposed atonal melody with altered intervals	56
3.3	A familiar tune with a distorted contour.	56
3.4	A dominant tone replaced by a leading tone	57
3.5	The rhythm of a familiar song	60
3.6	Dowling's stimuli and their metric grouping	62
3.7	Sloboda and Parker's stimulus	64
3.8	Phrase 3, with a half-measure shift	64
3.9	Contour of the stimulus	65
3.10	A different melody with a similar harmonic structure	66
3.11	Recalls 1 and 6 of subject 4	68
3.12	Recall 3.1: Octave skip is eliminated	69
3.13	Recalls that were interrupted when the meter was broken	71
3.14	The assumed available information in recall 2.1	73
3.15	Recall 3.1 and its correct information	73
3.16	Recalls 5.1 and 5.2	75
3.17	Recall 7.3	76
4.1	The first 16 harmonic partials	82
4.2	The two diminished triads with enharmonic interpretation	112
4.3	Terhardt's model of subharmonics	115
4.4	Possible roots for C major	116
4.5	Possible roots for C minor	116
4.6	A m3 as an indirect root support	117
4.7	Perception of simultaneous ascending and descending scale	125
4.8	Alternative melodic lines	126
4.9	Interleaved melodies	128
4.10	Spatial location experiment	131

5.1	Possible segregations of interleaved melodic lines	143
6.1	Effects of temporal order on the perception of tonality	157
6.2	Position finding facilitated by the Tritone	159
6.3	Tonic-identifying rare intervals	160
6.4	Profiles of the diminished triad	162
6.5	Profiles of the diminished triad and major triad, superimposed	163
6.6	Effect of temporal order on tonic judgment	164
6.7	Effect of temporal order on tonal ambiguity	164
6.8	Effect of metric position on tonality	165
6.9	Effect of metric position with the Elimination of temporal order . . .	165
7.1	Rhythms of equal spaces but different durations	175
7.2	Rhythmic patterns created by various means of accentuations	177
7.3	Durations versus points of attack	179
7.4	Bach, <i>Brandenburg Concerto No. 5</i> , Adagio	186
7.5	Beethoven, <i>Symphony No. 9</i> , second movement	190
7.6	Beethoven, <i>Symphony No. 7</i> , Scherzo	191
7.7	Effect of tempo on the perception of rhythmic patterns	192
7.8	Povel's durations ratios	196
8.1	Chunking into memory bits	206
8.2	Modes of presentation of pitch combinations	218
8.3	Pitch retention in aural drill	219
8.4	Tendencies in C major	220
8.5	Modifications of a sight singing text	225
8.6	Movable syllables for melodic tendencies and harmonic functions . . .	231
8.7	Features extraction in taking a dictation	237

List of Tables

2.1	Absolute Pitch judgments errors, Group A	27
2.2	Absolute Pitch judgments errors, Group C	28
4.1	Tenney's Consonance–Dissonance Concepts	80
4.2	Average number of responses for one and two semitones	94
4.3	Dissonance–consonance error types	97
4.4	Error types (Ortmann)	98
4.5	Same–scale–step versus one–semitone errors	100
4.6	One–semitone, with no same scale step interval	100
4.7	Discrimination test results compared with the two models	106
4.8	Similarity test results compared with the interval size model	107
4.9	Ranking comparison of the discrimination test results	110
4.10	Ranking of root ambiguity for triads and 7th chords.	117
4.11	Average masking ranges of some pitches.	126
6.1	Correlations of C major to other major keys	153
6.2	Correlations of C minor to other minor keys	154
6.3	Correlations of C major to other minor keys	155

Chapter 1

Introduction

The purpose of this study is to review research of music perception and cognition and to determine whether and to what extent it is relevant to teaching college level ear training for Western tonal music. The scope of the thesis is limited in two aspects; the music and the listeners. The music dealt with here is Western Common Practice music, while world musics, atonal music and twentieth century music are mentioned only for comparison. As for the listeners, the focus is on college level music majors, and the discussion of developmental aspects of music perception is limited. Again, nonmusicians are mentioned for comparison, and of special interest are the differences in strategies rather than the expected worse performance on tasks that are based on learned skills. Musicians and nonmusicians differ in their perception of music to the

point that they process music in different brain hemispheres.¹

Research of music perception and cognition is conducted by scholars from various disciplines, and mostly under interdisciplinary fields such as psychoacoustics and psychomusicology. The interaction of these disciplines — sciences such as physics and acoustics, with psychology, which the scientific community is reluctant to accept as a science, and with the art of music — is problematic. The validity of the results of scientific experiments depends on the design of the stimuli. Scientists strive to isolate components and aspects and to avoid interactions, but the art of music involves relations, contexts and interactions. Thus, most stimuli used in experiments are musically poor, of little interest to musicians, and fail to demonstrate mental processes occurring during listening to “real” music.

The goals of the research do not necessarily lay in the realm of music and thus the findings are not always musically relevant. Auditory perception is often researched as an expansion of findings in the realm of perception in other modalities, such as vision. The auditory stimuli used in such experiments may, or may not be of any musical significance, but often times the conclusions are erroneously expanded to music.

In the wide spectrum of the world’s musics, very few phenomena are regarded as universal, or at least pertaining to most musical cultures. The first is octave

¹T. Bever and R. J. Chiariello, “Cerebral Dominance in Musicians and Nonmusicians,” *Science*, 185 (1974): 537-539.

equivalence, which was shown to exist even in the pitch perception of white rats.² The second musical phenomenon is the division of the octave into discrete pitches rather than perceiving pitch on a continuum, and the third one is the existence of four to seven structurally important pitches in an octave.

Even those basic principles are not always taken into account in the design of experiments, especially earlier ones. Experiments using a pitch continuum were designed by scholars such as Wedell³ and Cuddy,⁴ while Paulson *et al.* divided the pitch continuum into equal differences of 100 cps,⁵ which means different frequency ratios in different registers.⁶ Even experiments using discrete pitches in intervals based on frequency ratios, do not always treat these pitches as a part of any musical system, and thus their conclusions may be valid as long as they are not held relevant to music perception.

Musicians are not incapable of providing their own insights about music perception and cognition. Anecdotal accounts of music perception often come from ear training teachers and serve as a basis for developing various methods. There are

²H. R. Blackwell and H. Schlosberg, "Octave Generalization, Pitch Discrimination, and Loudness Thresholds in the White Rat," *Journal of Experimental Psychology*, 33 (1943): 407-419.

³C. H. Wedell, "The Nature of the Absolute Judgment of Pitch," *Journal of Experimental Psychology*, 17 (1934): 485-503.

⁴L. L. Cuddy, "Training the Absolute Identification of Pitch," *Perception and Psychophysics*, 8 (1970): 256-259.

⁵Cycles per second.

⁶D. G. Paulson, R. Orlando and A. M. Schoelkopf, "Experimental Analysis and Manipulation of Auditory Generalization in Three Developmental Retardates by Discriminated-operant Procedures," *IMRID Papers and Reports*, 4, no. 13 (1967) (Nashville, Tennessee: Institute on Mental Retardation and Intellectual Development, Geroge Peabody College for Teachers).

numerous books presenting methods and approaches based on teachers' individual experience. This variety reflects not only individual differences in music perception, but also the teachers' individual beliefs about the process of perception and cognition. Both music perception research and ear training approaches could benefit from more communication and interaction. Mental processes assumed by ear training teachers could serve as hypotheses for experiments, while evidence found by research could help teachers develop efficient and meaningful approaches.

Since perception is a process of organization, some analysis methods may be regarded as models of perception. It should be taken into account, though, that these are *possible* models, rather than true or proved description of mental processes. Since it is possible to control and direct the focus of attention as well as some mental organization processes, mastering a specific method of analysis may alter one's perception of a musical piece. However, not all analysis methods are widely accepted, or proved, as being perceptually valid.

1.1 Thesis Overview

This thesis attempts to provide an overview of research in music perception in as much as it is relevant to tonal ear training. The last chapter (Chapter 8) points out information that can be used by ear training teachers, and suggests an approach that follows perceptual principles found by this research. Section 1.2 sets the grounds for

the following chapters by a brief discussion of perception and cognition in general, and specifically auditory perception. Chapters 2 through 7 focus on the perception of pitch aspects in Common Practice music. While separating pitch aspects from temporal aspects is necessary for the discussion, this separation is not a perceptual phenomenon; on the contrary, pitch and time interact in many ways, as do other aspects such as timbre and loudness.

The chapters proceed from the smallest musical unit, the single pitch (Chapter 2), to two-tone combinations and short melodies (Chapter 3). While a single pitch has little musical meaning, in the smallest pitch-combination the two tones create a context for each other and their combination already carry some tonal meaning. Thus, the perception of a context-free single pitch is relatively easy to explore as an isolated stimulus subject to controlled experiments. As the stimuli get more complex and musically meaningful, designing a controlled experiment gets more complicated. Thus, a compromise is needed, between a poor insignificant stimulus with full control over all aspects on one hand, and a musically meaningful stimulus with some uncontrolled interferences, on the other hand. The type of compromise and the direction it leans toward, depends by and large on the discipline under which the experiments are conducted. Thus, psychoacoustical experiments are usually highly controlled. The tendency toward less control and more meaningful stimuli is best represented by the

work of Sloboda and Parker (see Section 3.5).⁷

Simultaneous musical events are discussed in Chapter 4. Again, the chapter proceeds from isolated tone combinations such as simultaneous intervals and chords, to simultaneous events which unfold over time. The later type of events present new aspects, such as streaming, or selecting which components are to be grouped together, and the selection and direction of attention. Although the grouping process is also a part of the perception of single melodic lines, in the case of melodies different groupings imply different spans of time, while in the case of simultaneous events, the same time span may allow more than one way of grouping.

As was mentioned earlier, timbre interacts with pitch perception. This interaction is of special importance in this thesis for two reasons. First, some of the experiments cited used sine tones as stimuli, while others used specific musical timbres, usually generated by synthesizers. As shown in Chapter 5, these choices are very likely to have an impact on the tests results. Second, ear training is largely performed on the piano; it is likely that skills that are acquired during such training are not going to be as useful in a musician's life when s/he encounters a variety of timbres, as they were in the classroom.

The discussion of tonality (Chapter 6) no longer deals with the perception of specific stimuli, but rather with cognitive processes involving the integration of

⁷J. A. Sloboda and D. H. H. Parker, "Immediate Recall of Melodies," in *Musical Structure and Cognition*, eds. P. Howell, I. Cross and R. West (London: Academic Press, 1985).

musical events with an existing schema of the tonal system. These processes lead both to the establishment of tonality, and to the establishment of tonal hierarchy within a previously established tonality.

Of all aspects, time is the hardest one to keep in a separate discussion. Music has no meaning without unfolding over time, and temporal considerations are mentioned in other chapters. The discussion of musical time requires some definitions, since musicians use terms such as *rhythm*, *meter*, *beat* and *pulse* in a variety of meanings. The perception of these temporal aspects is discussed in Chapter 7, as well as the perception of rhythmic groups.

The last chapter (Chapter 8) attempts to bridge between music perception research and the ear training classroom. A summary of the research is followed by a discussion of the goals of ear training, and how research can help achieving these goals. A general approach to tonal ear training (rather than a specific method) which takes into account conclusions drawn from this research is suggested.

1.2 Perception and Cognition

The cognitive process is described by psychologists as consisting of perception, encoding, storage and retrieval. There is, though, a pre-cognitive process, in which sensory information is stored briefly, without undergoing any cognitive processes. This type of memory storage is called, in the case of music, echoic memory, and it lasts about

two seconds.⁸ While sensation can occur without awareness, the process of perception involves attention — the selection of what is attended to and what is ignored. The cognitive process involves categorizing, labeling and comparing to existing mental representations. It may result in learning as changes and adjustments to the existing mental representations are necessary. Since the focus of this thesis is on musicians and their learning process, the cognitive processes are of greater interest here than the sensory ones.

While sensation and perception occur in the presence of the stimulus, cognitive processes may last longer, or occur later, and thus involve memory. Memory is of a special importance in the case of music which works in the space of time, as put by J. Sloboda: “To perceive an event musically ... is to relate it to past events.”⁹ In the case of echoic memory, no relations can be perceived, unless the stimuli are less than two seconds apart.

Psychologists debate on whether the distinction between short-term and long-term memory is valid, but they do agree that there are no clear-cut boundaries between these memory types in terms of their capacity, duration and type of processes involved. For clarity of discussion, this distinction between short-term and long-term memory will be kept here, but it should be noted that processes described below may

⁸R. G. Crowder, “The Role of One’s Own Voice in Immediate Memory,” *Cognitive Psychology*, 1 (1970): 157-178.

⁹J. A. Sloboda, *The Musical Mind: The Cognitive Psychology of Music* (New York: Oxford University Press, 1985), 175.

not be strictly limited to one type of memory or the other.

Short-term memory is a “working buffer” which keeps information as part of the psychological present while the information is being submitted to cognitive processes and then discarded to make space for more information. The capacity of short-term memory is measured in information bits rather than in time, although it does have a time limit. It is believed to hold between five and nine bits of information. The size of an information bit depends both on the stimulus and on the level of expertise.

C. Brower describes the dependence on the musical material:

The duration of the psychological present depends in part on the kind of information presented, the strength of its organization, and the rate at which events occur. The more recursively organized the material, the more efficiently we are able to chunk it, and the greater are our chances of reaching a grouping boundary before we exhaust the resources of working memory.¹⁰

An experienced musician is likely to perform this grouping and organization more efficiently and group larger chunks of information into one bit.

The storage in long-term memory is practically unlimited in capacity and in duration. It stores memory traces of past events as well as concepts and abstractions, or as termed by E. Tulving, it serves as an *episodic memory* and a *semantic memory*.¹¹

Parncutt explains that sensory material becomes more stable in memory the more

¹⁰Candace Brower, “Memory and the Perception of Rhythm,” *Music Theory Spectrum*, 15, no. 1 (1993): 23.

¹¹Endel Tulving, *Elements of Episodic Memory* (Oxford: Clarendon Press, 1983).

it goes through semantic processing.¹² Thus, material stored in long-term memory underwent the highest level of cognitive processes.

Cognitive processes, then, are processes such as noticing and selection of attention; integrations with previously stored information such as categorizing, labeling, activation of and changes in mental representations; organization such as hierarchical grouping and other groupings, abstraction and extraction of meaning. Musical meaning is regarded in this thesis as embodied, non-referential meaning, based on L. B. Meyer's theory of meaning resulting from expectations.¹³ Meyer has adopted Morris' definition of meaning: "Anything acquires meaning if it is connected with, or indicates, or refer to, something beyond itself."¹⁴ In music, as in language, semantics is related to syntax, meaning is related to form. However, since musical events can only indicate more musical events, syntax and form become meaning.

Expectations are based on experience, and in the case of Common Practice music, as in natural language, exposure results in learning of the musical syntax and recognition of well formed formulae as well as deviations from such formulae. It is these syntactical formulae, created by systematic transformations of pitch over time that is of interest in this thesis. Thus, the thesis focuses on pitch aspects and temporal

¹²R. Parncutt, *Harmony: A Psychoacoustical Approach*, Springer Series in Information Sciences (New York: Springer-Verlag, 1989), 28.

¹³Leonard B. Meyer, *Emotion and Meaning in Music* (Chicago: University of Chicago Press, 1956).

¹⁴Morris Cohen, *A Preface to Logic* (New York: Henry Holt & Co., 1944), 47.

aspects of Common Practice music. Other aspects, such as timbre and loudness are mentioned for the purpose of discussing their interaction with the perception of pitch and time.

Organization of material and pattern recognition improves with learning, but according to the *Gestalt* theory some level of organization occurs spontaneously and is regarded as innate and universal. This type of perception is described by a series of perceptual principles presented by scholars such as Wertheimer¹⁵ and Koffka.¹⁶ These principles were developed specifically for visual perception, but some of them are applicable to aural perception, and indeed, they were incorporated in theoretical works such as Meyer's *Emotion and Meaning in Music* and Lerdahl and Jackendoff's *Generative Theory of Tonal Music*.¹⁷

A basic *Gestalt* principle is the law of *Prägnanz* which states that a form or an organization will be as good as the prevailing conditions allow it to be. It means that the mind strives to find the simplest and most complete organization available. Other *Gestalt* principles that are applicable to music are the laws of *proximity*, *similarity* and *good continuation*. The laws of proximity and similarity are demonstrated for visual perception in Figures 1.1 and 1.2 respectively. In Figure 1.1, the proximity

¹⁵M. Wertheimer, "Untersuchungen zur Lehre von der Gestalt," *Psychologische Forschung*, 2 (1923): 301-350, cited in: K. Koffka, *Principles of Gestalt Psychology*, 4th ed. (First published in England in 1935. London: Routledge & Kegan Paul, 1955).

¹⁶K. Koffka, *Principles of Gestalt Psychology*, 4th ed. (First published in England in 1935. London: Routledge & Kegan Paul, 1955).

¹⁷F. Lerdahl and R. Jackendoff, *A Generative Theory of Tonal Music* (Cambridge, Massachusetts: MIT Press, 1983).

OXXO OXXO OXXO OXXO
 OXXO OXXO OXXO OXXO
 OXXO OXXO OXXO OXXO
 OXXO OXXO OXXO OXXO
 OXXO OXXO OXXO OXXO
 OXXO OXXO OXXO OXXO
 OXXO OXXO OXXO OXXO

Figure 1.1: Law of proximity

O X X O O X X O O X X O O X X O
 O X X O O X X O O X X O O X X O
 O X X O O X X O O X X O O X X O
 O X X O O X X O O X X O O X X O
 O X X O O X X O O X X O O X X O
 O X X O O X X O O X X O O X X O
 O X X O O X X O O X X O O X X O

Figure 1.2: Law of similarity

of some of the components results in the emergence of the pattern OXXO, while in Figure 1.2 the similarity of the components results in the vertical organization of XX and OO. The law of *good continuation* is explained by K. Koffka: “any curve will proceed in its own natural way, a circle as a circle, an ellipse as an ellipse, and so forth.”¹⁸ All these principles will be mentioned later in this thesis in regard to their musical manifestations.

The *Gestalt* theory was often criticized for underestimating the role of learning

¹⁸Koffka, *Principles of Gestalt Psychology*, 153.

in perception. The spontaneous character of perception described by these principles, should be restricted in two ways. First, though the pattern perceived spontaneously in Figure 1.1 is OXXO, and in Figure 1.2 separate vertical patterns of OO and XX are likely to be perceived, it is possible to alter these perceptions voluntarily. Thus, with some concentration, it is possible to see the pattern OXXO in Figure 1.2, and the vertical patterns OO and XX in Figure 1.1. Second, learning can, and does alter perception. In Figures 1.1 and 1.2, one principle was clearly effective, but we regularly encounter less clear patterns on which more than one principle operates. Consider the visual stimulus **IEECMMOTT**; it is likely to be grouped by similarity as **I EE C MM O TT**, or **IEE CMM OTT**. When presented with wider spaces, such as in **IEEC MMO TT** it will probably be grouped by proximity. Altering the order of the letters to **COMMITTEE** will result in the immediate perception as a whole of the word “committee” by the experienced in this language, while the inexperienced is likely to group by similarity to **CO MM I TT EE**.

One last *Gestalt* principle, involved in the perception of the above example, is the principle stating that the whole is more than the sum of its parts. The meaning of the written word **COMMITTEE** cannot be explained solely by means of its component letters. The understanding of this principle is of vital importance to the understanding of the implications of the various aspects of research that are discussed in this thesis. Though the perception of a single pitch is not irrelevant to

the perception of melodies. the later cannot be explained solely on the basis of the perception of its component pitches, and the same is true for most perceptual musical wholes and their components.

Chapter 2

Single-Pitch Perception

Though a single pitch has very little musical meaning, the ability to retain a pitch is valuable for a musician. This ability is also relatively easy to explore, because the pitch is free of context — a rather disturbing element in controlled experiments.

The inherent human ability to retain a pitch is very poor. We take it for granted that we can instantly identify and name colors, flavors, textures — but only few people (absolute pitch possessors) have the same ability regarding pitch identification. For most people, and for most musicians, memory for pitch deteriorates rapidly, and no representation for it is stored in a long-term memory. Neurological experiments that compared brain waves of absolute pitch possessors with those of nonpossessors, showed that while nonpossessors produced certain brain waves during the task of identifying a pitch, absolute pitch possessors produced no such brain waves. This

led to the conclusion that most people process pitch identification in a temporary “buffer,” or a working memory, in which they compare the pitch to a former pitch (thus, using relative pitch).¹

2.1 Pitch Retention by Nonpossessors of Absolute Pitch

The deterioration rate of the memory for pitch is influenced by the different kinds of interference. A series of experiments held by Diana Deutsch researched mainly interferences in memory, or in other words — what makes us forget. The first interference is, of course, time. Pitch memory, though, deteriorates very slowly when the time interval is silent; speech interference does not impact it significantly, but interference of other pitches does.

In an experiment held in 1970, Deutsch presented subjects with pairs of tones separated by a six seconds interval.² The subjects were asked to judge whether the second tone was similar to the first one. In the first condition, the time interval was silent, and only subjects that obtained a 100% score proceeded to the second condition. Then, eight tones were interpolated, and the subjects were instructed to

¹Edwin C. Hantz *et al.*, “Effects of Musical Training and Absolute Pitch on the Neural Processing of Melodic Intervals: A P3 Event-related Potential Study,” *Music Perception*, 10 (1992): 25-41.

²D. Deutsch, “The Deterioration of Pitch Response in Memory” (Unpublished doctoral diss., University of California, San Diego, 1970), cited in: D. Deutsch, “The Processing of Pitch Combinations: Memory Systems,” in *The Psychology of Music*, ed. D. Deutsch, Academic Press Series in Cognition and Perception (New York: Academic Press, 1982).

ignore them, and again, judge the first and last tones. The correct answers were not far from chance (40% errors).

These results raised the question why those interpolated tones interfered with the memory, and Deutsch suggested three hypotheses: (1) the interpolated tones distracted the subjects' attention from the test tones, (2) the interpolated tones overloaded the memory and (3) pitch memory is a specialized system, in which only pitch information may distract the attention.

Deutsch's next experiment was set to determine which of the above was true: the first two hypotheses imply that any other interference, such as speech, would cause the same deterioration in the memory for pitch. But when spoken numbers were interpolated, there was only minimal deterioration of the memory for the test tones, even when subjects were instructed to listen to the numbers and recall them. Deutsch's conclusion was that the third hypothesis was true — that pitch information is retained in a specialized system.³

In another series of experiments held by Deutsch, a tone similar to the first or last test tones appeared in various serial positions. Deutsch concluded that when the critical tone was identical to the first tone and appeared in the second serial position, the memory for the first tone was better. Deutsch called this phenomena "consolidation," and concluded that consolidation is independent of time, but depends

³D. Deutsch, "Tones and Numbers: Specifity of Interference in Short-term Memory," *Science*, 168 (1970): 1604-1605.

on serial position.⁴

2.2 Absolute Pitch

The memory mechanism for a single pitch is entirely different in the case of absolute pitch possessors, and the research of this phenomenon addresses entirely different issues. This research will be discussed here in some length, which does not imply that possessing absolute pitch is the ultimate goal of ear training. The research is, however, of special interest to this study. The speed and accuracy of response to a single pitch is a very handy tool for a musician, especially if it can be extended to more complex stimuli (such as intervals, chords, chord successions, etc.). The experiments investigating the acquisition of absolute pitch may suggest ways of simulating and extending this skill to develop aural skills that work as quickly and accurately as absolute pitch.

Absolute pitch is regarded by many musicians as a rare and desirable ability, and has occupied the minds of scholars for many years, with the research going as far back as 1883.⁵ It was studied by scholars of various disciplines: psychologists, neurologists, musicologists and musicians, with over eighty sources available. While

⁴D. Deutsch, "Mapping of Interactions in the Pitch Memory Store," *Science*, 175 (1972): 1020-1022.

⁵C. Stumpf, *Tonpsychologie: Vol. I* [The Psychology of Music: Vol. I] (Leipzig: S. Hirzel Verlag, 1883).

musicians have a lot to say about the nature of absolute pitch and its applications to music, they do not usually hold experiments, and thus, very little of the literature on the subject is written by musicians.

Many questions and debates were brought up by this wide research, some of which are not yet resolved. Experiments were set to investigate the mechanism of absolute pitch and its difference from relative pitch, and maybe come up with a definition of who is a “genuine” absolute pitch possessor. The nature of absolute pitch was investigated in an attempt to determine whether it works as pitch discrimination on a pitch continuum, or as pitch identification of distinct categories. Other experiments investigated the skills involved in absolute pitch. Can absolute pitch possessors produce certain pitches at will, or just identify them? Can they identify chords and keys as quickly and directly as they do single pitches? But more than all other questions, researchers debated the nature–nurture question, and attempted to create a method to train nonpossessors to acquire absolute pitch.

2.2.1 Mechanism of Absolute Pitch

The debate about the mechanism of absolute pitch revolved around the question whether absolute pitch possessors have inner representation for one referential tone, or whether they have representations for each tone (or pitch class). It seems that with the more recent neurological researches, there is evidence in favor of the later

possibility.⁶ This type of experiment measures the appearance of an event-related potential called P₃₀₀ (positive amplitude, and an average of 300 milliseconds of latency).⁷ The relatively late appearance of the P₃₀₀ is believed to reflect the time in which a cognitive process is taking place.

Klein *et al.* found that the P₃₀₀ was missing in absolute pitch possessors while processing tasks of pitch identification.⁸ The P₃₀₀ was not missing in nonpossessors. They concluded that the P₃₀₀ was activated by a “working memory” (a short-term memory in which information is processed by comparing the stimuli), and that this comparison was not processed by absolute pitch possessors who seemed to have representations for all the pitches. This also explains the instant response of absolute pitch possessors to pitch-identification tasks.

2.2.2 Speed and Accuracy of Absolute Pitch

There is, then, a real difference between an absolute pitch possessor, and a nonpossessor with good memory for pitch. The criteria for possessing absolute pitch are the degrees of speed and accuracy. These criteria were researched by K. Miyazaki, who presented some of the most interesting, carefully designed and well reported recent

⁶M. Klein, M. G. H. Coles and E. Donchin, “People with Absolute Pitch Process Tones Without Producing a P300,” *Science*, 223 (1984): 1306-1309; Hantz *et al.*; M. Tervaniemi *et al.*, “Absolute Pitch and Event-related Brain Potentials,” *Music Perception*, 1, no. 3 (spring 1993): 305-316.

⁷Event-related potentials are small changes in the electrical brain activity caused by the processing of a stimulus.

⁸Klein *et al.*, “People with Absolute Pitch”

experiments in this field.⁹ After the first experiment, Miyazaki classified the subjects into three groups: (1) a group of ten subjects who achieved more than 90% correct answers (eliminating the task of octave discrimination — the chance level for identifying one of the twelve pitch-classes is about 8.5%, or 1:12), with the average response time of 0.605 seconds, (2) a group of twelve subjects who achieved 70%–90% correct answers, with the average response time of 0.717 seconds, and (3) a group of ten subjects who achieved 70% correct answers with the average response time of 1.167 seconds (Grebel'nik had also set an accuracy of 63% correct answers as a standard for “genuine” absolute pitch¹⁰). These results give us some idea of the accuracy of absolute pitch, and also show clearly the correlation between speed and accuracy. Miyazaki found that the correct answers were faster than the incorrect ones. Probably, the delay in the identification indicates that some other methods (such as relative pitch) were being employed.

Miyazaki's criterion for choosing her subjects was probably at least 70% correct answers. Terhardt and Seewann also refer to 70% to 80% as a minimum for a potential absolute pitch.¹¹ Summarizing an experiment with 135 talented and trained musicians

⁹K. Miyazaki, “The Speed of Musical Pitch Identification by Absolute Pitch Possessors,” *Music Perception*, 8 (1990): 177-188.

¹⁰S. G. Grebel'nik, “Formirovanie u Doshkol'nikov Absoliutnovo Muzykal'no, Slukha [The Formation of Absolute Pitch in Pre-school Children],” *Voprosy Psikhologii*, 2 (March-April 1984): 90-98, cited in: M. Rush, *An Experimental Investigation of the Effectiveness of Training on Absolute Pitch in Adult Musicians* (Ph.D. diss., Ohio State University, 1989).

¹¹E. Terhardt and M. Seewann, “Aural Key Identification and its Relationship to Absolute Pitch,” *Music Perception*, 1 (1983): 63-83.

(eleven of them were absolute pitch possessors), they wrote: “Non-AP-possessors are not really ignorant of absolute pitch; rather, many of them can be regarded as ‘potential AP possessors.’ In a categorical sense, the only difference between these non-AP-possessors and ‘real AP possessors’ is that the former are not aware of their ability.”¹²

2.2.3 Perceptual Principles Pertinent to Absolute Pitch

A great majority of the research concerning absolute pitch was done from a psychological (rather than musical) point of view. This may explain the number of experiments that regard pitch-perception as working on a pitch-continuum rather than dividing the continuum into distinct pitch classes, and treat sounds as aural stimuli not related to any musical system.¹³ Thus, subjects were asked to identify frequencies, regardless of musical pitches, and respond with cps rather than pitch names. This method ignores some aspects of music perception such as the principle of octave generalization, the categorical nature of pitch perception, and the perceptual equivalence of intervals and their inversions.

¹²Terhardt and Seewann, “Aural Key Identification,” 82.

¹³Wedell, “Absolute Judgment;” Cuddy, “Training.”

Octave Generalization

Octave Generalization is a cross cultural principle which was widely investigated, with the conclusion that humans of all musical cultures — and some animals too¹⁴ — perceive tones in the interval of an octave as equivalent. Ignoring this principle while measuring the errors in identification by frequencies leads to regarding an error of an octave as a large one, and an error of a semitone as a small one. On the other hand, experiments based on musical-scale methods show that absolute pitch possessors make more octave-identification errors than any other type of errors.¹⁵ One of Miyazaki's experiments, which required identification of the pitch class and the specific octave, led her to the conclusion that "absolute-pitch subjects perceive pitch classes immediately but they perceive pitch height as indirectly and inaccurately as those who have no absolute pitch do."¹⁶

Categorical Pitch Perception

It is true that there is no evidence of discontinuity in pitch discrimination along the pitch continuum, but there is much evidence of the categorical nature of both relative and absolute pitch. As noted by H. Helmholtz, in all cultures the octave is divided

¹⁴Blackwell and Schlosberg "Octave Generalization."

¹⁵A. Bachem, "Time Factors in Relative and Absolute Pitch Determination," *Journal of the Acoustical Society of America*, 26 (1954): 751-753.

¹⁶Miyazaki, "Speed of Musical Pitch Identification," 183.

into discrete steps, so that the melodic movement in a melody will be apparent.¹⁷ This concept was carried further by G. A. Miller who argued that the human mind is capable of only 7 ± 2 categorical judgments, and indeed this is the number of scale steps in most musical cultures.¹⁸

It is possible that verbal labeling of pitches increases the tendency to categorize, and thus absolute pitch possessors may categorize “out-of-tune” pitches to the nearest pitch class that they have labeled as a prototype.¹⁹ Indeed, research shows that absolute pitch possessors do not have a superior ability to discriminate pitches within the categories. Siegel and Siegel showed that while absolute pitch possessors demonstrated good discrimination between categories, their pitch-discrimination within the categories was poor.²⁰ Tervaniemi *et al.* carried out a neurological experiment using brain waves, and they too concluded that absolute pitch possessors were not performing better than nonpossessors in discrimination of pitches which were out of tune with the Western musical scale.²¹ Their final conclusion was that pitch discrimination (on the pitch continuum) and pitch identification (categorization) are based on different brain mechanisms.

¹⁷H. L. F. von Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, trans. A. Ellis, New York: Dover, 1954 (Originally published as *Die Lehre von der Tonempfindungen als physiologische Grundlage für die Theorie der Music*, Braunschweig, 1863).

¹⁸G. A. Miller, “The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information,” *Psychological Review*, 63 (1956): 81-97.

¹⁹Sloboda, *Musical Mind*, 27.

²⁰J. A. Siegel and W. Siegel, “Absolute Identification of Notes and Intervals by Musicians,” *Perception and Psychophysics*, 21 (1977): 143-152.

²¹Tervaniemi *et al.*, “Absolute Pitch and Event-related Brain Potentials.”

An interesting exception of the tendency to categorize pitch was found in an experiment of training three retarded teenagers to respond to the frequency of 1455 cps (about F#⁶). Paulson, Orlando and Schoelkopf reported that two of the subjects hardly responded to other pitches, and one of them, who was further tested in intervals of 100 cps, kept responding only to the 1455 cps tone.²² Their conclusion did not relate to the surprising absolute pitch identification ability of the subjects; Paulson *et al.* concluded that their subject lacked the ability of generalization.

Based on *Gestalt* principles and on writings of Garner²³ and Rosch,²⁴ Krumhansl asserts that "particular perceptual and conceptual objects have special psychological status ... within categories certain members are normative, unique, self consistent, simple, typical."²⁵ This assertion is a basis to Krumhansl's studies of hierarchies of musical pitch, which, among their many implications, impact the speed and accuracy of their identification.

The categorical nature of absolute pitch and Krumhansl's concept of hierarchy are reflected in the types of errors made by subjects and reported by Miyazaki in 1990.²⁶ The most common errors seem to be proximity errors (Miyazaki regards semitone errors as proximity errors); the pitch classes G and C are most easily identi-

²²Paulson *et al.*, "Experimental Analysis."

²³W. R. Garner, "Good Patterns Have Few Alternatives," *American Scientist*, 58 (1970): 34-42.

²⁴E. Rosch, "Cognitive Reference Points," *Cognitive Psychology*, 7 (1975): 532-547.

²⁵C. L. Krumhansl, *Cognitive Foundations of Musical Pitch* (New York: Oxford University Press, 1990), 17.

²⁶Miyazaki, "Speed of Musical Pitch Identification."

fied, and then D, E and F; A varies in terms of speed and accuracy between third and sixth in this row. Then come the “black-key” pitch classes, with B somewhere after one to three “black” pitch classes. These findings of the “white-key” pitch classes being easier to identify than the “black-key” pitch classes are supported by previous findings. Miyazaki also noted that black-key pitch classes are more often “corrected” to the adjacent white-key pitch class than vice versa.

Equivalence of Inversions

Although the stimuli were presented in over five octaves (from C^2 to B^6), the data provided by Miyazaki do not indicate whether an error is of a specific interval or its inversion. Therefore, only the first six intervals were used to present the results, regarding the other six as perceptually equivalent inversions (see Deutsch, 1982²⁷). Empirical evidence to the phenomenon of inversions equivalence will be discussed in Section 4.1.2.

2.2.4 Types of Errors in Judgments of Absolute Pitch

The data for groups A (90% accuracy) and C (70% accuracy) are displayed by Miyazaki in two charts showing for each pitch class, the number of correct answers,

²⁷D. Deutsch, “The Processing of Pitch Combinations: Memory Systems,” in *The Psychology of Music*, ed. D. Deutsch, Academic Press Series in Cognition and Perception (New York: Academic Press, 1982), 273.

and also the number of each other wrong name that was assigned to it. Table 2.1 and Table 2.2 are adaptations of Miyazaki's Figures 2 and 3 respectively, with the data presented to show which intervals between the correct answer and the wrong answer were most common.

Table 2.1: Absolute Pitch judgments errors, Group A

Pitch Class	Correct Responses	Minor 2nd	Major 2nd	Minor 3rd	Major 3rd	Perfect 4th	Aug. 4th	Total Errors
B	196	4	0	0	0	0	0	4
A#	177	10	10	0	0	3	0	23
A	189	5	6	0	0	0	0	11
G#	176	12	9	0	0	3	0	24
G	193	2	4	0	1	0	0	7
F#	175	16	2	1	3	3	0	25
F	191	6	2	1	0	0	0	9
E	188	9	1	0	0	2	0	12
D#	188	8	4	0	0	0	0	12
D	192	5	2	1	0	0	0	8
C#	187	8	3	0	0	2	0	13
C	190	5	3	1	1	0	0	10
Total	2242	90	46	4	5	13	0	158
% of Total Errors	-	57.0%	29.0%	2.5%	3.0%	8.0%	0.0	100.0%

Source: Miyazaki 1990, p. 180

As revealed by the "Total Errors" column in Table 2.1, the decrease of number of errors is not completely correlated with the increase in interval width. The number of errors for a M2 is 51% percent of the errors of a m2; but the decrease from the M2 to the m3 is of 91%. This may imply that both errors of m2 and M2 may be

accounted for as proximity errors. The 3rds, both minor and major provide about the same number of errors, but there is a significant increase in errors for the P4 (and its equivalent P5). Out of the 2242 trials, none of the errors was at an interval of a TT.

Table 2.2: Absolute Pitch judgments errors, Group C

Pitch Class	Correct Responses	Minor 2nd	Major 2nd	Minor 3rd	Major 3rd	Perfect 4th	Aug. 4th	Total Errors
B	94	73	13	8	4	3	0	101
A#	96	63	6	6	7	11	6	99
A	156	19	13	2	3	2	0	39
G#	58	94	25	9	3	6	0	137
G	169	14	8	1	1	1	1	26
F#	80	82	18	9	5	0	1	115
F	104	63	16	4	3	5	0	91
E	153	25	5	4	1	5	2	42
D#	73	98	9	2	3	9	1	122
D	163	21	6	3	1	1	0	32
C#	65	106	14	1	1	6	2	130
C	151	19	17	0	0	8	0	44
Total	1362	677	150	49	32	57	13	978
% of Total Errors	-	69.0%	15.5%	5.0%	3.0%	6.0%	1.5%	100.0%

Regarding group C as possessing less of the “absolute pitchness,” its results show the same pattern, but to a lesser extent. The decrease in the number of errors in group C is more gradual — lessening the difference between the proximity errors and other kinds of errors — which may imply that possessing less of “absolute pitchness”

increases the dependence on pitch height.

These results may be explained by chroma, overtones, or the tonal system. It is clear, however, that absolute pitch relies on other clues than pitch height.

2.2.5 Additional Absolute Pitch–Skills

Active Absolute Pitch

So far absolute pitch was treated as the ability to identify single pitches quickly and accurately. Though singing was part of some of the experimental training methods,²⁸ none of the experiments was designed to investigate “active absolute pitch” — the ability to produce a specific pitch at will. It is true that the physical sensation of voice production gives musicians some reference to the pitch they are about to produce; but this has nothing to do with the way absolute pitch works. Whether and to what extent active absolute pitch is an implied skill of absolute pitch possessors is yet to be investigated.

Identification of Keys

Another skill that may be implied by possessing absolute pitch is the ability to identify tonal keys. The aural identification of a key may be regarded as the ability to name the tonic of a piece without a reference tone. While some absolute pitch possessors

²⁸Grebel'nik, “Formirovanie u Doshkol'nikov.”

claim they have this ability (Bachem,²⁹ Vernon³⁰), others claim that the tonic is not available to them as a primary characteristic, but may be inferred (Corliss,³¹ and also Brady³² who actually simulated absolute pitch, as will be described later). Bachem claimed that “Particular characteristics of certain keys, e.g., the brilliancy of A major, the softness of D-flat major, can only be appreciated fully through absolute pitch.”³³ As nonpossessors of absolute pitch also express some ideas of key-characteristics, it is difficult to determine whether or not they are ‘appreciated fully.’ N. Spender writes that modulations provide absolute pitch possessors with “sensory pleasure,” while for nonpossessors this is merely an aesthetic and intellectual experience.³⁴ Probably, any novice with no absolute pitch may derive sensory pleasure from a modulation of a semitone (a “shift”). It is true, though, that many “smooth” diatonic modulations go unnoticed by nonpossessors musicians. These cases may be regarded as poor absolute pitch ability of those musicians, and should not exclude the possibility that other nonpossessors musicians would perceive a shift in the pitch of the tonic.

Terhardt and Ward found that absolute pitch possessors do recognize keys bet-

²⁹A. Bachem, “Absolute Pitch,” *Journal of the Acoustical Society of America*, 27 (1955): 1180-1185.

³⁰P. E. Vernon, “Absolute Pitch: A Case Study,” *British Journal of Psychology*, 68 (1977): 485-489.

³¹E. L. Corliss, “Remark on ‘Fixed-scale Mechanism of Absolute Pitch’,” *Journal of the Acoustical Society of America*, 53 (1973): 1737-1739.

³²P. Brady, “Teach Yourself Absolute Pitch in 365 Hard Lessons,” *Journal of the Acoustical Society of America*, 48 (1970): 88.

³³Bachem, “Absolute Pitch,” 1185.

³⁴Natasha Spender, “Absolute Pitch,” *The New Grove Dictionary of Music and Musicians* (New York: Macmillan, 1980).

ter then nonpossessors, but that nonpossessors also had variable degrees of key recognition.³⁵ A year later, Terhardt and Seewann³⁶ investigated the aural key identification by musicians, both possessors and nonpossessors of absolute pitch, presenting two questions: whether absolute pitch possessors identify keys better than nonpossessors, and whether the process of identification differs between the two groups. The subjects were presented with a score, and were asked whether the piece heard was in the “right” key, and if not, they were asked to specify the transposition. The nonpossessors showed a significant ability of key recognition. The possibility that they were doing so by cues other than pitch (such as key character) was excluded by the fact that the number of errors decreased as the interval of the transposition got wider. Terhardt and Seewann concluded that nonpossessor musicians have some degree of absolute pitch which is well above chance level, without being aware of this ability. Furthermore, they found these results sufficient to exclude the existence of primary key characteristics.

As to the process of identification, it was found that the performance level of the nonpossessors depended on the unambiguity of the tonality of the piece, while the absolute pitch possessors were not influenced by this. Terhardt and Seewann concluded that while absolute pitch possessors relied on identification of single pitches,

³⁵E. Terhardt and W. D. Ward, “Recognition of Musical Key: Exploratory Study,” *Journal of the Acoustical Society of America*, 72 (1982): 26-33.

³⁶Terhardt and Seewann, “Aural Key Identification.”

the nonpossessors relied on a “holistic feeling for key”³⁷ deducted from a series of tones.

Identification of Intervals and Chords

While exploring the possibility of training people to acquire absolute pitch, some attempts were made to expand this ability to identifying intervals and chords. Regardless of whether they succeeded or not, it seems that absolute pitch possessors are gifted initially with the ability to identify single pitches and do not always apply it further. Moreover, sometimes the dependence on absolute pitch impairs the ability to identify relations between pitches. Such a case was demonstrated by Ken'ichi Miyazaki³⁸ who showed that when asked to identify melodic intervals starting at out-of-tune C, the performance of absolute pitch possessors was less accurate than the performance of nonpossessors.

2.2.6 Acquisition of Absolute Pitch

The greatest debate about absolute pitch is the one concerning the “nature–nurture” question.³⁹ It seems that scholars who are themselves possessors of absolute pitch (e.g., Bachem, Corliss) tend to believe that this is an inherited ability, while most

³⁷Terhardt and Seewann, “Aural Key Identification,” 80.

³⁸K. Miyazaki, “Perception of Musical Intervals by Absolute Pitch Possessors,” *Music Perception*, 9 (1992): 413–426.

³⁹W. D. Ward and E. M. Burns, “Absolute Pitch,” in *The Psychology of Music*, Academic Press Series in Cognition and Perception, ed. D. Deutsch (New York: Academic Press, 1982).

scholars try to find some evidence that absolute pitch is learned. These views are reflected in four different theories, with the heredity theory standing alone on one hand, and three learning theories on the other hand: learning, unlearning and imprinting.⁴⁰

The Heredity Theory

The main advocate of the heredity theory, A. Bachem, could not provide clear evidence to support her theory, but so far no one was able to exclude completely a genetic component in the absolute pitch. To do so, one would have to create a training method that would succeed for all the subjects, or, accepting the imprinting theory (see below) — for all young children. This genetic component may be a general musical talent (combined with the right kind of training at the right age), or a more specific ability connected with the absolute identification of pitch. Whatever it is, this question remains yet unanswered.

The Learning Theory

The learning theory suggests that, with the right training, almost any person can acquire absolute pitch. This approach matches the current general trend of favoring learning over heredity. Again, to make this suggestion acceptable, one has to come up with an effective training method that will work for most people.

⁴⁰Ward and Burns, "Absolute Pitch."

Most attempts of scholars to train themselves to acquire absolute pitch were not successful. Meyer reported achieving 60-64% accuracy, but also that when the training stopped, his ability deteriorated.⁴¹ Other attempts were not more successful (Mull,⁴² Wedell⁴³). Experiments of training other subjects showed a remarkable improvement in the pitch identification ability, but none of them could be considered as acquiring absolute pitch.

One experiment, though, showed some outstanding results. In 1970, P. Brady trained himself for two months and, allowing an error of one semitone, achieved 96% accuracy.⁴⁴ Brady reported that after the training stopped, he retained his ability, and in 1975 he participated in an experiment held by J. B. Carroll,⁴⁵ together with four "genuine" absolute pitch possessors. His score was indistinguishable from that of the other possessors.

It should be noted, though, that reported in the form of scientific experiments, many of the training experiments would hardly be accepted by the scientific community. Experiments were held with an extremely small number of subjects (sometimes one); data were reported selectively, and sometimes only interpreted results were pre-

⁴¹M. Meyer, "Is the Memory of Absolute Pitch Capable of Development by Training?" *Psychological Review*, 6 (1899): 514-516.

⁴²H. K. Mull, "The Acquisition of Absolute Pitch," *American Journal of Psychology*, 36 (1925): 469-493.

⁴³Wedell, "Nature of the Absolute Judgment of Pitch."

⁴⁴Brady, "Teach Yourself Absolute Pitch."

⁴⁵J. B. Carroll, "Speed and Accuracy of Absolute Pitch Judgments: Some Latter-day Results," (Research Bulletin No. ETS-RB-75-35, Princeton, New Jersey: Educational Testing Service, ERIC Document Reproduction service No. ED 150 153, n.d.).

sented without any display of the raw data; the training periods span from two days to one year; and not all the experiments included both a pre-test and a post-test. In view of all the above, it seems that there is not much point in citing the claimed results of the methods, especially for most of them which could claim only for “remarkable improvement” but not for acquiring absolute pitch.

The various training methods reflect the different views on the nature of absolute pitch. Some do not use any musical elements, some use one referential tone, some train an identification of a series of tones, and some construct a musical-tonal method.

In the nonmusical methods, the pitch continuum was divided to nonmusical categories, determined by the purpose of the experiment. Such experiments were held by C. H. Wedell who attempted to determine the smallest frequency difference which may be learned as absolute pitch,⁴⁶ and L. L. Cuddy, who attempted to determine whether the referential or serial training works better (training the subjects to identify one reference tone, or training to identify each tone of a series).⁴⁷

The musical-referential method is based on the view that absolute pitch works by comparing all pitches (using relative pitch) to one referential pitch which is stored in a long-term memory. It involves training the subjects to remember a single pitch.

⁴⁶Wedell, “Nature of the Absolute Judgment of Pitch.”

⁴⁷Cuddy, “Training the Absolute Identification of Pitch.”

Mull held such an experiment, measuring the answers in microtones.⁴⁸ Cuddy's first series of experiments was set to train the subjects to identify one single pitch — A⁴.⁴⁹ Other pitches were played too, but the subjects were instructed to concentrate on the identification of A⁴, and feed-back was given only for that pitch. These experiments may have scientific value in the realm of perception, but with the recent findings as to P₃₀₀ and the inner representations of all pitch classes, they are of no use in the attempts to simulate absolute pitch.

Lundin was probably the first to employ a method with some tonal meaning, using a tonal-serial method.⁵⁰ First, he introduced the diatonic ("white key") tones, and then added the chromatic tones following the standard order of the sharps. This method matches in part Miyazaki's findings,⁵¹ only Lundin used it almost thirty years earlier. Brady's outstanding experiment was held in 1970, and was programmed on a tonal basis too. He started with a series of tones in which C occurred most frequently; then he gradually reduced the occurrences of C until it got to 1/12 of the stimuli tones. His method might have been regarded as training of a single referential tone, only he sounded other tones and thus was experienced with their relations to the C during the training; this might be one facet of a "relative absolute pitch" — an absolute pitch

⁴⁸Mull, "The Acquisition of Absolute Pitch."

⁴⁹L. L. Cuddy, "Practice Effects in the Absolute Judgment of Pitch," *Journal of the Acoustical Society of America*, 43 (1968): 1069-1076.

⁵⁰R. W. Lundin, "Can Perfect Pitch Be Learned?" *Music Education Journal*, (1963): 49-51.

⁵¹Miyazaki, "Speed of Musical Pitch Identification."

relative to a tonic. Brady commented on his work: "One never identifies tones from 'interval from previous note' techniques, unless the previous note is C. For example, the sequence G-A should not sound like a whole step; the G should sound like C's dominant, and A like C's minor."⁵² Interestingly, though, Brady reported that he could not identify the key of a piece, which may indicate that he trained himself to the tonality of C major.

Brady's impressive results attracted other scholars' interest. As mentioned above, Carroll tested him in 1975; in 1980, Joseph M. Waters, a student at the University of Minnesota, following Brady's method, tried to reproduce his success and improve the method. Unfortunately, this is an unpublished paper, cited in Rush's dissertation.⁵³ Waters started with training only C and another pitch class. In each of the following sessions, he added another pitch class, keeping the occurrences of C at 50%. Only when all pitch classes were introduced, did he start reducing the occurrences of C to chance level. This method was applied to a small group of subjects, and to Waters himself. Additionally, Waters, as a sole participant of his own experiment, constructed "tonality training" in which he drilled for each pitch class a certain tonal sequence. Waters reported a remarkable improvement in his absolute pitch, but

⁵²P. Brady, "Fixed-scale Mechanism of Absolute Pitch," *Journal of the Acoustical Society of America*, 48 (1970): 884.

⁵³J. Waters, "[Development of] Absolute Pitch" (Unpublished paper, University of Minnesota, n.d.); cited in M. Rush, *An Experimental Investigation of the Effectiveness of Training on Absolute Pitch in Adult Musicians* (Ph.D. diss., Ohio State University, 1989).

claimed that for each identification task he used C, which was well memorized, as a reference tone.

It seems that the strongest results were achieved by Brady's method followed by Waters, and it may be worthwhile to continue this research with a more controlled and well reported experiment. If this training resulted in memorization of C, repeating the same method of training twelve times on the twelve tones, may result in acquiring absolute pitch (see Professor Rubinstein's aural drill in Section 8.3.1).

Paul Hindemith supported the learning theory, but his position was somewhat different and relied strongly on nonpitch cues such as timbre:

The ear of every normally gifted musician can be trained to become what is supposed to be "absolute" in its judgment, the more so since there is no such absoluteness, based on any physical attribute of the ear. What we call "absolute pitch" is merely a highly developed ability to compare quickly an audible impression with acoustic archetypes stored in our memory. These archetypes are always taken from earlier musical experience and are closely related to all kinds of external qualities of tones and of tone-production.⁵⁴

Two cultural facts seem to support the learning theory and Hindemith's view: one is the existence of musically-illiterate cultures, in which absolute pitch has no meaning (one would expect that an inherent ability would have some universal use or meaning). The second one is the existence of tonal languages (such as Chinese) in which meaning is attached to pitch (that is, relative pitch, but not absolute pitch).

⁵⁴P. Hindemith, *Elementary Training for Musicians* (New York: Associated Music Publishing, 1946), 206.

An existence of absolute pitch innate in the human brain might have prevented the evolution of such languages.

The “Unlearning” Theory

The unlearning theory was presented in 1901 by O. Abraham, who pointed out that most musical activities do not encourage the use of absolute pitch.⁵⁵ He suggested that most people were born with the potential for developing absolute pitch, but this potential was neglected or suppressed; only some lucky people who, in their early childhood, were presented with absolute pitch tasks, retained this ability.

The Imprinting Theory

The imprinting theory is related to the unlearning theory, and suggests that absolute pitch can be acquired in early childhood, and that people who do not possess absolute pitch were probably not presented with absolute pitch tasks in the critical period. Scholars supporting this theory often compare the development of categories for absolute pitch to that of the categories for phonemes: “...for a particular language, many acoustically different sounds known as allophones are taken as exemplars of one phoneme. Similarly, in music, acoustically different sounds are categorized by a par-

⁵⁵O. Abraham, “Das absolute Tonebewusstsein,” *Sammelbde. International Musikges*, 3 (1901): 1-86, cited in: Paul Fraisse, “Rhythm and Tempo,” in *The psychology of music*, ed. D. Deutsch, Academic Press Series in Cognition and Perception (New York: Academic Press, 1982).

ticular note name.”⁵⁶ As it is known that the categories for language are acquired in early childhood, it may be assumed that the same is true for musical categories, but as Cohen and Baird point out, the musical categories are not as necessary as the linguistic ones, and this might be the reason that they are not universally acquired.

Indeed, the correlation of early-childhood training and the possession of absolute pitch was shown in several experiments (Sergeant,⁵⁷ Miyazaki⁵⁸), and the results made this theory very plausible. The critical period is believed to be between the age of three and the age of five or six. This is also believed to be the age when a child can acquire an accent-free foreign language.⁵⁹

2.2.7 Benefits of Absolute Pitch

Although many musicians regard absolute pitch as a desirable ability, its contribution for the perception of musical relations and structure is limited, and in some cases, it may even become a disability. *Gestalt* psychologists noted the primacy of relational structures in perception. This means that absolute pitch does not provide better perception and understanding of musical patterns and structures (such as intervals,

⁵⁶A. J. Cohen and K. Baird, “Acquisition of Absolute Pitch: The Question of Critical Periods,” *Psychomusicology*, 9 (1990): 31.

⁵⁷D. Sergeant, “Experimental Investigation of Absolute Pitch,” *Journal of Research in Music Education*, 17 (1969): 135-143.

⁵⁸K. Miyazaki, “Musical Pitch Identification by Absolute Pitch Possessors,” *Perception and Psychophysics*, 44 (1988): 501-512; Miyazaki, “Absolute Pitch Identification: Effects of Timbre and Pitch Region,” *Music Perception*, 7 (1989): 1-14.

⁵⁹Ward and Burns, “Absolute Pitch.”

chords, melodies, harmony and tonality). If a musician had to choose between relative and absolute pitch, it is obvious that relative pitch is indispensable for understanding music. It is then left to find out what more absolute pitch can contribute to our music perception, as well as the ways in which it may limit the perception.

The advantage of the speed and accuracy of absolute pitch is clear in the case of atonal music tasks; as for tonal music, the speed of identification typical to absolute pitch possessors is very appealing, especially since it indicates that no extra steps are being taken while on task. For nonpossessors, on the other hand, there is a typical tendency to take excessive steps during tasks such as melodic dictation or sight-singing. A musician might calculate the interval from the previous tone, divide the interval into smaller intervals for easier identification, use familiar tunes to identify any of those intervals, and so forth. These extra steps are time consuming and extraneous to the musical material. Furthermore, one wrong step is enough to lead to a totally wrong answer.

Another advantage, mentioned earlier, is the perception of modulations, which is undoubtedly different for possessors and nonpossessors. Whether or not key-characteristics are real perceptual entities, and whether or not absolute pitch possessors are more sensitive to them than nonpossessors is not yet determined. Regardless, the mere ability to have the initial tonic readily available in the memory while or after a modulation takes place, definitely puts absolute pitch possessors in an advantage

over nonpossessors.

It should be pointed out, though, that absolute pitch is relative in many ways:

- (1) Individual possessors have various degrees of absolute pitch and, as was shown above, even nonpossessors, especially musicians, possess some degree of absolute pitch.
- (2) The categorical nature of absolute pitch contradicts the very use of the term “absolute,”
- (3) Brady acquired absolute pitch relying on relativity of all tones to one single standard,
- (4) Absolute Pitch possessors must use relative pitch to understand musical relations, and
- (5) The possessors’ account of their own ability varies.

2.2.8 Drawbacks of Absolute Pitch

Possessing absolute pitch seems to involve some inflexibility. Transposition tasks are extremely difficult for absolute pitch possessors, as well as adjusting to a different tuning (such as Baroque tunings). The categorical nature of absolute pitch makes it more difficult for possessors to identify “out-of-tune” intervals. While this will present a nonpossessor with the simple task of categorizing the interval, a possessor might first categorize both tones, and then categorize the interval, which might lead him/her to the wrong conclusion.⁶⁰

Regarding their inflexibility, some absolute pitch possessors report a great distress when, at around the age of fifty, a shift of one or two semitones upwards occurs.⁶¹

⁶⁰Miyazaki, “Perception of Musical Intervals.”

⁶¹Vernon, “Absolute Pitch.”

This is explained by the aging of the ear, and some physical changes in the basilar membrane.⁶²

The greatest disadvantage of absolute pitch, though, is that possessors seem to depend greatly on their absolute pitch, even in tasks which require understanding of tonal relationships. In my experience as an ear training teacher, I encountered absolute pitch possessors who have “survived” an extensive musical education, but failed to sing a root of a chord, a tonic of a melodic sequence, or a resolution of a dissonance. Thus, although they may show superior ability in taking a melodic dictation, this ability may not necessarily reflect a better understanding of music. As was noted by Miyazaki, “it may be argued that those children who have acquired AP must be trained in RP [relative pitch], through some training program specially designed for them.”⁶³

2.2.9 Summary

The perception of a single pitch, then, is very different for absolute pitch possessors and nonpossessors. The nonpossessors’ memory for a single pitch is very poor but may be improved significantly by training. Investigation of the mechanism of absolute pitch showed a strong correlation between speed and accuracy, both being significantly higher for possessors than nonpossessors. Yet, when the stimuli consists of more

⁶²Ward and Burns, “Absolute Pitch.”

⁶³Miyazaki, “Perception of Musical Intervals,” 424.

than a single pitch, with patterns and structures being involved, nonpossessors may perform as well, if not better than absolute pitch possessors.

Chapter 3

Perception of Melodies

Research on the perception of melodies is entirely different from that of a single pitch, because melodies, having context and carrying meaning, are more complex stimuli. Mastering the simplest tonal melody requires understanding of tonal relationships. Trotter¹ and Wing² showed that subjects with good pitch discrimination did not necessarily have good ability to master a melody.

The influence of meaning on perception (in L. B. Meyer's sense of meaning as expectations) was shown by Unyk and Carlsen who constructed a "melodic expectancy profile" for individual subjects, and then presented them with melodic dictation that fulfilled or denied their expectations.³ They found that dictation that followed the

¹J. R. Trotter, "The Psychophysics of Melodic Interval: Definitions, Techniques, Theory and Problems," *Australian Journal of Psychology*, 19 (1967): 13-25.

²H. Wing, "Tests of Musical Ability and Appreciation," *British Journal of Psychology*, 27 (1948): Monograph Supplements.

³A. Unyk and J. Carlsen, "The Influence of Expectancy on Melodic Perception," *Psychomusicol-*

subjects' expectations was easier to take. Francès has found that detecting alterations in melodies was harder for atonal melodies than for tonal ones.⁴ Regarding perception as a process of organization goes along with these findings: when there are expectations, organization "happens;" when expectations are not met, the stimuli is perceived as "random" and an extra intellectual effort is needed to organize it.

3.1 Research Methods

Most of the studies focus on one, or a few aspects of melodies, such as intervals, tonality and contour (pitch aspects), rhythm, meter and serial position (temporal aspects). Focusing on one aspect makes it easier to construct controlled experiments. Different aspects were found to have different roles in the perception of melodies, and their prominence would change depending on the length of the melody and also on the type of memory involved (short-term versus long-term memory). It is convenient, for the sake of research and discussion, to study each of these aspects separately, but in "real music" all of these aspects interact and influence each other, and it is almost impossible to discuss the perception of one aspect without relating to any of the other aspects.

The most easily controlled method of research is the *Recognition Method*, which

ogy, 7 (1987): 3-23.

⁴R. Francès, *The Perception of Music*, trans. W. J. Dowling, Hillsdale, New Jersey: L. Erlbaum, 1988.

presents the subjects basically with a “yes or no,” or “same or different” question, and thus provides data which is quite simple to analyze. The short-term memory task is detecting alterations of the standard melody, while the long-term memory task is recognizing a familiar melody despite the alterations. The *Recall Method*, on the other hand, requires of the subjects to reproduce the melody, by singing, whistling, playing, etc. This method provides highly complex data. The errors in the recall carry meanings, and are much more difficult to formulate into percentages and statistics. Yet, there is much to be said in favor of this method, mainly because recalling a melody is a higher cognitive process than mere recognition. Furthermore, the data provided are indeed complex, but while it means complications in analysis, it provides us with valuable information and are able to account for all aspects of the melodies.

Experiments using the recognition method focus mainly on the role of different pitch aspects of melodies, such as scale, contour, intervals, and compositional devices (such as tonal answers) in our ability to recognize a melody as “same.” Although these aspects influence our memory, none of them can account for the way melodies are stored in our memory. In language, a sequence of phonemes produce words, a sequence of words produce a phrase. None of these, though, are stored in our memory in that sequence for the purpose of reproducing the meaning carried by the phrase. What is stored is an abstraction of the meaning, which can be reproduced by

paraphrasing, or even in another language, without ever repeating the same sequence of phonemes. Music, however, does not denote anything but the music itself, and thus, there is no external reference that may be stored. The abstraction, then, is in the musical meaning, which is one's set of expectations based on one's musical culture. In Common Practice music, the abstraction will be the tonal system as it is manifested in a specific melody. One equivalence of paraphrasing in music may be transposing. A transposed melody preserves the tonal meaning of the original melody without using the same pitches. Another way of paraphrasing is reproducing the harmonic succession using different pitches.

Dowling has commented on the powerful quality of this abstraction: "Tonal scales constitute one of the most durable families of perceptual-motor schemata that have been observed in psychology, ranking perhaps only after the schemata of natural language in their stability and resistance to change in adult life."⁵ Our long-term memory for melodies is so powerful that tonal melodies are often used as mnemonics.

Having said all that, it is still worth while looking at the way the different components affect melody recognition. There is a wide agreement that the most basic component of melodies, the actual pitches (or in physical terms, the frequencies), are not the basis for recognition of melodies.⁶ To support this agreement, it is enough to

⁵W. J. Dowling, "Scale and Contour: Two Components of a Theory of Memory for Melodies," *Psychology Review*, 85 (1978), 345.

⁶D. Deutsch, "Octave Generalization and Tune Recognition," *Perception and Psychophysics*, 11 (1972): 411-412; J. Edworthy, "Towards a Pitch-contour Continuum Theory of Memory for

point that a transposed melody may not share a single pitch with the original melody and yet be recognized as “same.”

The smallest melodic cell, then, is a sequence of two tones, which is relatively easy to investigate, since the intervallic component of the melody is isolated. Other pitch aspects can be easily controlled as well; contour can only be ascending, descending, or repeating the same pitch, and tonality takes at least three tones to be unequivocally established (see David Butler's theory of Rare Intervals⁷). As for temporal aspects, even with accents, meter or rhythm cannot be established with two tones, and the serial position can be one of two.

3.2 Interval Perception

Much like the perception of single pitches by absolute pitch possessors, and like the perception of speech phonemes, the perception of intervals is mostly categorical. It may be surprising for a musician to read in Ward's summary of intonation studies how imprecise intervals get during performance, and how these imprecisions are not detected by musicians, who accept a variety of pitch ratios as being the same inter-

Melodies,” in *Acquisition of Symbolic Skills*, eds. J. Sloboda and D. Rogers (New York: Plenum Press, 1983)

⁷David Butler, “Describing the Perception of Tonality in Music: A Critique of the Tonal Hierarchy Theory and a Proposal for a Theory of Intervallic Rivalry,” *Music Perception*, 6 (1989): 219-242.

val.⁸ On the other hand, musicians' precision in adjusting intervals was shown to be relatively superior to the adjusting of other stimuli, both auditory (e.g. loudness) and nonauditory (e.g. length of lines).⁹

Results of interval recognition experiments obtained similar results to those of speech experiments of voiced and unvoiced stops.¹⁰ Overlap between categories is small, and the boundaries of the categories are clear, with peaks of recognition around the boundaries, and troughs within the categories. Speech phonemes were regarded as unique auditory stimuli that are perceived categorically, and while speech is processed in the left hemisphere, music was believed to be processed in the right hemisphere.¹¹ It was shown by Bever and Chiariello that while this was true for nonmusicians, experienced musicians did process music in the left hemisphere.¹²

Compared to Miller's concept of five to nine possible categories,¹³ Burns and Ward hold that trained musicians have thirty categories for intervals (which is all the intervals within a range of about two and a half octaves).¹⁴ In this sense, the

⁸W. D. Ward, "Musical Perception," in *Foundations of Modern Auditory Theory*, ed. J. V. Tobias (New York: Academic Press, 1970).

⁹S. S. Stevens and M. Guirao, "Subjective Scaling of Length and Area and the Matching of Length to Loudness and Brightness," *Journal of Experimental Psychology*, 66 (1963): 177-186.

¹⁰Siegel and Siegel; M. Studdert-Kennedy *et al.*, "The Motor Theory of Speech Perception: A Reply to Lane's Critical Review," *Psychological Review*, 77 (1970): 234-249.

¹¹M. Studdert-Kennedy and D. Shankweiler, "Hemispheric Specialization for Speech Perception," *Journal of the Acoustical Society of America*, 48 (1970): 579-594.

¹²Bever and Chiariello, "Cerebral Dominance."

¹³Miller, "The Magical Number Seven."

¹⁴E. M. Burns and W. D. Ward, "Categorical Perception — Phenomenon or Epiphenomenon: Evidence from Experiments in the Perception of Melodic Musical Intervals," *Journal of the Acoustical Society of America*, 63 (1978), 456-468.

perception of intervals by trained musicians is absolute; they categorize intervals with no comparison reference. Investigating the existence of such categories, Miller *et al.* have demonstrated peaks of discrimination level in a speech-like stimuli continuum.¹⁵ These results were interpreted as showing the existence of some natural boundaries between categories. There is no evidence, though, for such natural categories for musical intervals. Moreover, training proved to be crucial to the level of intervals identification, since performance of nonmusicians in such tasks was found to be poor.¹⁶

It is clear, however, that there are learned interval categories, as shown by the performance of trained musicians. Miyazaki has found that some of the categories were perceptually superior to others, both in speed and in accuracy of identification.¹⁷ Absolute Pitch possessors were found to identify faster major seconds, major thirds and perfect fifths than any other intervals, while the accuracy of the major seconds, perfect fifths and perfect octaves was better than those of any other intervals. Results were not that consistent for nonpossessors, but they were found to perform better in the identification of perfect fifths and perfect octaves, as well as smaller intervals. It is important to note that these intervals also have a special musical status, being tonally more important than intervals such as sixths and sevenths.

¹⁵J. Miller *et al.*, "Discrimination and Labeling of Noise-buzz Sequences with Varying Noise-level Times: An Example of Categorical Perception," *Journal of the Acoustical Society of America*, 60 (1976): 410.

¹⁶L. L. Cuddy and A. J. Cohen, "Recognition of Transposed Melodic Sequences," *Quarterly Journal of Experimental Psychology*, 28 (1976): 255-270.

¹⁷Miyazaki, "Perception of Musical Intervals."

Having learned categories does not mean that the perception of intervals is limited to categorizing any given interval to one of those learned categories, without any ability of intra-categories discrimination. In fact, categorical perception and continuous perception are not dichotomous, but rather, themselves lie on a continuum on which better intra-categories discrimination is a “less categorical perception” while better identification of categories is a “more categorical perception.” Musicians were found to alter the interval sizes in performance consistently depending on the tonal context, or specifically, on the harmonic tension.¹⁸ While this type of performance shows the categorical perception of these intervals, the consistent alterations attest within-categories discrimination.

The results of a series of experiments support the assumption that the degree of categorical perception depends on the degree of the stimuli uncertainty.¹⁹ When subjects were asked to judge between a smaller number of interval categories, they made finer judgments (such as “a flat m3”). The assumption is that in “real” music (as opposed to laboratory music) there is a high degree of uncertainty, and therefore interval judgments are categorical.

Although interval identification is a typical relative-pitch task, research has shown that absolute pitch possessors may fail to employ relative-pitch strategy and

¹⁸A. Rakowski, “Intonation Variants of Musical Intervals in Isolation and in Musical Contexts,” *Psychology of Music*, 18 (1990) 60-72.

¹⁹Burns and Ward, “Categorical Perception — Phenomenon or Epiphenomenon.”

instead rely on their absolute pitch ability. It was found that even in the simplest task of determining which of two notes is higher, the P₃₀₀ amplitude was reduced, which suggests that subjects were using absolute pitch.²⁰ Using absolute pitch in an interval identification task would mean categorizing each tone separately and then calculating the interval size. Such strategy may reduce the accuracy in such cases as when both tones are out of tune in opposite directions (e.g., a sharp D and a flat F# may be categorized as a m3, while an absolute pitch possessor would categorize the D and the F# first, and calculate a M3 between these tones.) The impact on speed of identification was shown by Miyazaki in two different experiments: one experiment used an out-of-tune reference tone, and the other one used a tonal context which is either out-of-tune, or in a less familiar key (F# major).²¹ In all these cases, identification by absolute pitch possessors was slower. These findings were affirmed by further experiments of Hantz *et al.* that have shown that in a task of identifying major and minor thirds, absolute pitch possessors did not produce P₃₀₀ brain waves, or in other words, they were not employing relative pitch strategy.

²⁰Hantz *et al.*, "Effects of Musical Training."

²¹Miyazaki, "Perception of Musical Intervals."

3.3 Perception of Three-Tone Sequences

A logical step in applying theories of intervals perception to the perception of melodies is experimenting with longer sequences of intervals. In a series of experiments, held by Cuddy and Cohen, using three tone sequences, three models were suggested, all based on the assumption that melody recognition is based on the recognition of its component intervals.²² None of these models were supported by the test results, and in fact some of the results suggested just the opposite, that melodies are more than the sum of their component intervals.

Cuddy and Cohen were indeed able to show that trained musicians do abstract nonadjacent intervals, but the three tone sequences proved to be easier to recognize than the two tone ones. The facilitation of the task with the added interval may imply that intervals were not the main source of information. It seemed that adding a third tone added more nonintervallic information, an assumption that is supported by the findings that patterns A and B of these experiments were easier to recognize than patterns C, D and E (Figure 3.1). Although all patterns have the same intervallic content, patterns A and B are more prominent in tonal music.

²²Cuddy and Cohen, "Recognition."

Source: Cuddy and Cohen (1976), p. 158.

Figure 3.1: Recognition of three-tone patterns

3.4 Non-Interval Aspects of Melody Perception

The perception of melodies, then, cannot be explained solely on the basis of interval perception. Not only are melodies more than the sum of their component pitches, they are also more than the sum of their component intervals. Other, more meaningful aspects of melody perception are contour and tonality. The perceptual organization of all these pitch aspects are is greatly impacted by temporal aspects such as meter, rhythm and serial position.

Contour

As opposed to intervals, contour is a basic melodic aspect which seems to be available both to musicians and nonmusicians. Dowling and Fujitani found that transposed atonal melodies with the same contour were recognized as the same as the original melody, although they had different intervals (Figure 3.2).²³ They concluded that contour, and not intervals, was the main component in the recognition of short

²³W. J. Dowling and D. S. Fujitani, "Contour, Interval, and Pitch Recognition in Memory for Melodies," *Journal of the Acoustical Society of America*, 49 (1971): 524-531.

Figure 3.2: Transposed atonal melody with altered intervals

Figure 3.3: A familiar tune with a distorted contour.

atonal melodies. Further experiments with short tonal melodies supported these conclusion.²⁴

Although the importance of contour was shown to depend on the length of the melody and the type of memory involved (short-term versus long-term), Dowling holds that regardless of its salience, contour is an independent component in melody recognition.²⁵ Figure 3.3 is a melody with a distorted contour, and although most of the intervals are wrong, the directions (ups and downs) are preserved, and the melody is easily recognizable.

²⁴Edworthy, "Towards a Pitch-contour Continuum Theory."

²⁵Dowling, "Scale and Contour."

Figure 3.4: A dominant tone replaced by a leading tone

Tonality

Melodies are often defined as “a sequence of intervals,” but intervals do not seem to be the key to melody recognition of tonal melodies, or, as was shown by Dowling and Fujitani, in the case of short-term memory, not even for atonal melodies. Dowling maintains that in tonal melodies tonality is more important than intervals: “The function of mode is not to fix a set of intervals in semitones as belonging to a melody. If it were, tonal answers would not be confused with exact transpositions. The mode is simply a framework on which the contour may be hung.”²⁶ Dowling pointed out that altering the size of the intervals would not be obvious when these alterations are diatonic. He suggested a system that accounts for diatonic steps, diatonic skips of thirds, and larger diatonic leaps. Still, this system does not account for such hard-to-detect changes as replacing the dominant tone with the leading tone (Figure 3.4).

The importance of the role of tonality depends on the extent to which it is

²⁶Dowling, “Scale and Contour,” 350.

established, and in this sense, tonality is expected to play a more important role in longer melodies. This assumption is indeed affirmed by findings such as those of J. Edworthy, which showed that although contour is the most salient component of short melodies, as they get longer, tonality takes over.²⁷ She then made an interesting musicological implication, suggesting that this explains why “contour-preserving compositional devices” produce very short themes.

As was mentioned earlier, another factor in the salience of these components (except for the length of the melody) is the type of memory in question. While in short-term memory changes in intervals may go undetected, in long-term memory, the smallest deviation from the exact intervals is easily detected. Again, this does not necessarily imply that melodies are stored in long-term memory as a sequence of intervals, but that changes of interval sizes may alter the tonal meaning of melodies, and tonality is long lasting in long-term memory.

Many of the above short-term recognition studies used transposed melodies as comparison and target stimuli. Francès has shown that melodies are easily recognized when the transposition is of an octave, and hardest when the melody is transposed to a remote key.²⁸ Cuddy, Cohen and Miller have shown that recognition of transposed melodies is more accurate for a related key transposition (such as the dominant key)

²⁷Edworthy, “Towards a Pitch-contour Continuum Theory.”

²⁸Francès, *The Perception of Music*.

than for a remote key transposition (such as the tritone).²⁹ Dowling also noted that comparison melodies were most easily recognized in short-term memory when they started at the same pitch as the standard melody.³⁰

Temporal Aspects

Although temporal aspects such as meter, rhythm and serial position are parts of the context of melodies that interact with the context created by other pitches, most of the experiments do not refer to any of these aspects. One example of the interaction of rhythm and pitch was shown in an experiment held by Jones *et al.*, in which rhythmic context directed the subjects' attention to certain tones, and an alteration in those tones was recognized more easily than alterations of other tones.³¹

Rhythm influences the perception of a melody not only by interacting with pitch, but also as an independent aspect. It is possible to recognize familiar melodies (stored in long-term memory) by their rhythm only provided the rhythm is sufficiently unique. In a completely anecdotal experiment, most people I have asked were able to recognize the song in Figure 3.5. Dowling showed that recognition accuracy for the rhythm of a familiar melody was as high as 33%.³²

²⁹L. L. Cuddy A. J. Cohen and J. Miller. "Melody Recognition: The Experimental Application of Rules," *Canadian Journal of Psychology*, 33 (1971): 148-157.

³⁰Dowling, "Scale and Contour."

³¹M. Jones, M. Boltz and G. Kidd, "Controlled Attending as a Function of Melodic and Temporal Context," *Perception and Psychophysics*, 32 (1982): 211-218.

³²Dowling, "Scale and Contour."

recalled more accurately, and recall for intervals in the middle of a melody was the least accurate.³⁵

In general, only experiments using the *Recall Method* were able to account for these aspects. Except for the absence of valuable information, ignoring temporal aspects may distort the results of the experiments. W. J. Dowling, for example, was asking subjects to respond with “same” or “different” to melodies in which the duration of the quarter notes was controlled by a timer.³⁶ The melodies were altered in terms of transpositions, intervals and contour, assuming that the temporal aspects remained constant. The following two melodies, for example, were judged as “different” by the subjects, and in the analysis of the results, this was accounted for on the basis of changes in intervals and contour (Figure 3.6a,b).

There are, though, two temporal aspects that may have supported the subjects in their judgment: (1) one of the changes in contour is in the last serial position, which favors the “different” conclusion, (2) while the standard melody lends itself to a double metric grouping (Figure 3.6c), the comparison lends itself to a triple one (Figure 3.6d), again, favoring the “different” conclusion.

³⁵Jack Taylor, “Perception of Melodic Intervals Within Melodic Context” (Unpublished doctoral dissertation, University of Washington, 1971), cited in: Peggy A. Long, “Relationships Between Pitch Memory in Short Melodies, and Selected Factors,” *Journal of Research in Music Education*, 25, no. 4 (1977): 272-282.

³⁶Dowling, “Scale and Contour.”

Source: Dowling (1978), p. 346.

Figure 3.6: Dowling's stimuli and their metric grouping

3.5 The Recall Method

The recall method, as presented by Sloboda and Parker, reflects a totally different approach.³⁷ Using “real” melodies as stimuli does not allow a controlled experiment and a close look at each component separately. It results in complex data that may be less accurate than in other experiments, but are very insightful. Sloboda and Parker’s experiment yielded an enormous amount of data, only part of which was interpreted by them, and the rest was left for the reader to use and derive his/her own conclusions. One of the causes for the complexity of the data is that it involves musical meaning and structure, which provides us with some clues about the relationship between

³⁷Sloboda and Parker, “Immediate Recall of Melodies.”

musical knowledge and memory. It was this involvement of high level structure that was the impetus for Sloboda and Parker's investigation.

The experiment was constructed somewhat loosely, giving up some control over some factors for the benefit of the complex data that was collected. A careful consideration was given to each loose end, evaluating the deterioration in precision versus the benefits of the information. Eight female subjects (four musicians and four non-musicians) were presented with several melodies. Each of the melodies was played six times and the subjects were asked to recall each melody by singing immediately after each hearing. They found that the first recall interfered with the subsequent recalls, and more than that — the recalls of the first melody interfered with the recalls of the subsequent ones to such an extent that Sloboda and Parker discarded all the data, keeping only the results of the recalls for the first melody.

The melody, described by Sloboda and Parker as an A_1A_2B form, is actually not a complete form, but a phrase modulating from F major to D minor (Figure 3.7). They did not refer to the metric complexity of the phrase, due mostly to the irregular use of the anacrusis (preceding the second phrase). Anacrusis tend to “set a trend” and it is likely that once a phrase starts with an anacrusis, the subsequent phrases will start with one too.

An additional trigger to add an anacrusis to the third phrase is the octave skip, a melodic pattern typical of an anacrusis. Moreover, in the third phrase the meter

Source: Sloboda and Parker (1985), p. 148.

Figure 3.7: Sloboda and Parker's stimulus

Figure 3.8: Phrase 3, with a half-measure shift

becomes somewhat ambiguous; the dotted quarter note, formerly placed on the down beat of phrases 1 and 2, is now on the third beat. the dominant harmonies are on the down beats, and the tonic harmonies are on the third beats, altogether creating a half-measure shift (Figure 3.8).

The analysis of the data required some unified measures that would enable the scoring of the results. Sloboda and Parker measured each musical component separately, by counting the number of correct items in each component. In measuring the contour results, for example, the recalls were matched to the original contour (Figure 3.9) by finding the phrase, or phrases, that was most similar in contour to the recall,³⁸ and then scored by the number of the matching notes. Other measured

³⁸Contour in this experiment was regarded as the exact up and down interval movements, while contour in the sense used by Dowling (i.e., same directions but not necessarily same intervals) was referred to as "crude contour."

Source: Sloboda and Parker (1985), p. 150.

Figure 3.9: Contour of the stimulus

components were meter (number of beats in which $\frac{4}{4}$ time was maintained), rhythm (number of half bars matching in rhythm to half bars of corresponding positions in the stimulus), phrasing (number of breaths at phrase ends and at other places), phrase structure (more similarity between phrases 2 and 1 than between phrases 3 and 1) and harmonic structure (number of half bars beginning with a tone that is part of the original harmony for that half bar). All these numbers were then calculated, under each category, to the percentage of the total number of notes/beats in each subject's recalls.

With the relatively small number of subjects, it is hard to relate to "trends," and even more so to musicians versus nonmusicians trends, which would bring the number of subjects in a group down to four. On the other hand, every single subject and every recall is worth relating to, each subject being 25% of his/her group. In discussing the results, Sloboda and Parker's analysis of the data will be presented first, followed by additional interpretation of the raw data.

Figure 3.10: Subject 6 Recall 5: A different melody with a similar harmonic structure

Contour

The contour analysis showed a wide range of scores, from 0% to 88%. Sloboda and Parker were able to explain the lowest results of 0% and 2%, which narrowed the scores' range to 28% to 88%. Subject 6 (0%) produced an entirely different melody, based on similar harmonic structure (Figure 3.10).³⁹ As for subject 4 (2%), her opening phrase resembles phrase 3 and not phrase 1 and the appropriate score comparing the phrase this way would be 57%.

The question is whether this wide range means that the scoring of the contour is irrelevant to the perception of the melody, or whether it is simply that the subjects vary widely in their abilities.

³⁹This recall transcript and the ones that will follow are adapted from Sloboda and Parker, who provided the six recalls of all eight subjects. They are numbered by subject number and recall number (e.g., number 6.5 means subject number 6, recall number 5).

Meter and Rhythm

Meter was shown in this experiment to be the most fundamental underlying structure, the percentage scores average being as high as 88%. All the subjects were able to perceive the $\frac{4}{4}$ time and maintain it for at least two measures. Rhythm, though, proved to be independent of meter. Within the $\frac{4}{4}$ time frame, subjects created rhythms that were entirely different from those of the stimulus. In percentage, an average of only 46% of the rhythms were similar to those of the stimulus, the highest score being 66%.

The metric phrase structure of the recalls regarding anacruses was not analyzed by Sloboda and Parker, but a look at the transcripts shows that, as expected, the irregular occurrence of the anacrusis, the octave skip at the beginning of phrase 3, and the longer note on the third beat of phrase 3, seemed to have caused confusion. Since Sloboda and Parker have not indicated where they assumed bar lines in their metric analysis, dashed bar lines were added here by looking at all the recalls of each subject, and attempting to assume where the metric accents were perceived to be, rather than counting four beats from any given point. This process sometimes resulted in measures with more or less than four beats per measure, because it is possible that once a subject has established a meter, s/he may stray from it and resume it later. Most recalls seemed to give a clear idea of the metric perception of the subject, although it may be that listening to the taped recalls would have led to

Stimulus

The figure shows musical notation for a stimulus and three recall attempts. The stimulus is a single staff of music. Below it are three staves labeled 'a. Recall without a shift', 'b. Recall with a shift', and 'c. Recall: first without a shift, and then with a shift'. Each recall staff has a number to its left (4.1, 4.1, and 4.6 respectively). The notation includes various musical symbols: notes, rests, and beams. A legend at the bottom explains the symbols: a curved line under a group of notes means 'in one breath', an 'x' above a note means 'approximate duration', a wavy line above a group of notes means 'quickenings of tempo', and an arrow above a group of notes means 'slowing of tempo'.

a. Recall without a shift

b. Recall with a shift

c. Recall: first without a shift, and then with a shift

— = in one breath x = approximate duration ~~~ = quickenings of tempo → = slowing of tempo

Source: Sloboda and Parker (1985), p. 164.

Figure 3.11: Recalls 1 and 6 of subject 4

different assumptions.

The most ambivalent were the recalls of subject 4, who seemed to have indeed shifted by half a measure. The long notes may indicate a down beat (Figure 3.11a), although an interpretation with no shift is also possible (Figure 3.11b).

The last recall of subject 4, though, seems to indicate that a metric shift did occur, since it seems more obvious that phrase 3 starts with an anacrusis (Figure 3.11c).

A tally of the recalls reveals that out of the 48 recalls of phrase 1, eleven had an added anacrusis, and out of the 37 recalls of phrase 3, twelve recalls had an added anacrusis. Six more recalls of phrase 3 which did not add an anacrusis, eliminated the octave skip (see Figure 3.12 for one such example). As for phrase 2, 25 out

Figure 3.12: Recall 3.1: Octave skip is eliminated

of the 38 recalls had no anacrusis. It should be noted though, that this cannot be fully attributed to the tendency to keep all phrases consistently with or without an anacrusis, but also to the more general tendency to assimilate phrases 1 and 2.

Breath and Phrase Structure

The scores for breaths yielded the highest percentage, an average of 96% breaths at phrase endings, with musicians and nonmusicians yielding similar scores. The phrase structure analysis, though, did show some difference between musicians and nonmusicians. All subjects perceived, at least to some degree, the similarity between phrases A_1 and A_2 . Another common tendency was to assimilate phrase B to the first two phrases. This tendency was interpreted by Sloboda and Parker as the tendency of memory to simplify structures, and this is where musicians differed from nonmusicians. While the similarity score of phrases A_2 and B of the stimulus is 13, musicians' average score was 26, and nonmusicians' 52 (a higher score denotes a higher degree

of assimilation).

Harmony

Much like the results of melodic contour, the scores for harmony varied widely from one subject to another. Comparing individual subjects' scores for the different aspects, though, did not show significant correlation. An especially wide difference was found for subject 6 who scored lowest of all subjects in contour (0%) and second to the highest in harmony (93%). The lack of significant correlation between the scores of contour and harmony suggests that the harmonic perception is independent of the perception of the tones constituting the harmony. The scores for harmony also varied widely from one phrase to another, the mean score for the third phrase being the lowest, probably due to the modulation. A significant difference was found between musicians' and nonmusicians' scores, suggesting that harmony is a higher level structure whose perception depends on training.

Sloboda and Parker treated transpositions made by the subjects as equivalent to the original key and did not keep records of them. Though musically equivalent, the transpositions do have an impact on the perception of the tonality. Keeping the original key may help the subjects with a previously established tonality, while by singing in transposition, they have to re-establish the key.

Sloboda and Parker concluded that since scores for meter were so much higher

Stimulus

Recall

2.1

Recall

5.4

x = approximate duration o = approximate pitch

Source: Sloboda and Parker (1985), p. 163-4.

Figure 3.13: Recalls that were interrupted when the meter was broken

than those for harmony, meter may be represented independently of harmony. Harmony is one of the clues for meter, though, and it could be that subjects did perceive metric clues through harmony, yet were not able to reproduce that harmony. Another question regarding the interaction of meter and harmony, worth further research, is whether the perception of harmony or contour is independent of the perception of meter. Recalls 1.2 and 5.4 are two examples in which when meter was lost, the subjects stopped the recall (Figure 3.13).

These kinds of interaction may need to be investigated in a more controlled experiment. As was shown above, controlled experiments are easier to design using the recognition method, but it may be possible to increase the control over different aspects using the recall method by modifying the stimulus in ways that may enhance one component or another.

Some Remarks on the Method

Control over the different musical aspects was only one of the impediments presented by the recall method. It is assumed that the recalls represent the subjects' perception, but vocal control may also play a role. The linear representation forced by the singing does not necessarily reflect the way the melody is perceived. By singing, the subjects create another stimulus, which interferes with the perception of the original stimulus.

Singing the recall confines the subjects to a linear output of the recall, regardless of the sequence in which s/he perceives the different parts and aspects of the melody, and moreover, as Sloboda and Parker pointed out, the first recall interferes with the subsequent ones. Ideally, the subjects would have recalled whatever they have perceived, and omit whatever was not yet perceived. Since the human mind strives to create complete patterns, the subjects filled in the missing information with their own "compositions" and in many cases these compositions became more prominent in their memory than the stimulus.

Looking at the first recall of subject 2, for example (see Figure 3.13, recall 2.1), one may speculate that the subject has perceived the meter, the descending line of the first three beats, and the similarity of phrases 1 and 2, and filled in the rest of the information. The correct pitches were not perceived yet and the outcome was an atonal melody. Figure 3.14 is a notated representation of what is assumed to be the available information in recall 2.1.

Figure 3.14: The assumed available information in recall 2.1

Source: Sloboda and Parker (1985), p. 163.

Figure 3.15: Recall 3.1 and its correct information

Subject 3, on the other hand, had much more information in her first recall: she had the complete time frame (three phrases of two measures each, in $\frac{4}{4}$ time), and the phrase structure (first two phrases in D minor, the third in F major; higher similarity between phrases 1 and 2 than between phrases 1 and 3; an anacrusis to phrase 2). Figure 3.15 shows her first recall (Figure 3.15b), and a notated representation of the correct information contained in that recall (Figure 3.15c).

Using notation by the subject is impossible for the nonmusicians, and being a

complex skill even for musicians, it would involve many more parameters and complicate the interpretation of the data. Moreover, Sloboda and Parker's experiment is designed to investigate memory, and recording part of the information midway would interfere with this investigation. It may be possible, though, to design an experiment in which recalls with partial information are avoided. One way of doing so is letting the subjects listen to the melody six times without singing, and then sing the recall only after the sixth hearing. To ensure active listening, the subjects may be instructed to sing internally with the stimulus (in which case it is likely that the audible stimulus will override any deviations in the internal singing).

Another possibility is to let the subjects sing more than once after each hearing. This will not eliminate the problem of the recall becoming a part of the stimulus, but it may help overcoming the linear nature of the recall. After one recall, a subject may realize that some of the information contradicts other information that is already available, that some structural principles that s/he had perceived were not manifested in the recall, or that elements that were perceived were not put in the correct order. The last case was highly evident in Sloboda and Parker's experiment when five of the subjects used material from the last phrase in the first phrase (a phenomenon referred to by Sloboda and Parker as the "recency effect,"⁴⁰ and may be related to earlier finding regarding the perceptual superiority of the last part of the stimulus⁴¹).

⁴⁰Sloboda and Parker, 151.

⁴¹Taylor, "Perception of Melodic Intervals."

The figure displays musical notation on a single staff. At the top, a line of music is labeled 'Stimulus'. Below it, two recall attempts are shown. Recall 5.1 is a short sequence of notes. Recall 5.2 is a longer sequence, with a bracket underneath the first four measures labeled '= in one breath' and a 'v' above the fifth measure labeled 'v = slight pause'. A legend below the notation explains these symbols.

Stimulus

Recall

5.1

Recall

5.2

= in one breath v = slight pause

Source: Sloboda and Parker (1985), p. 164.

Figure 3.16: Recalls 5.1 and 5.2

Subjects 2 and 4 consistently replaced the musical material of all phrases by the material of phrase 3, but for subjects 5, 7 and 8, although their first recall started with influences of phrase 3 in phrase 1, this was fixed in the second recall (see, for example, subject 5 in Figure 3.16).

In such cases, an additional recall (or recalls) may improve the results without an additional hearing. This may actually have been the case in recall 7.3 (Figure 3.17). It is the only recall that contains almost seven full measures, and it may be that the subject was “stuck” after measure 5, and that measure 6 is a repetition and an attempt to improve the recall of measure 5.

One of the benefits of the recall method is its resemblance to “real” musical activities. It may be worth conducting more experiments along this line of listening to “real” melodies, but eliminating, or minimizing the obstacle of the singing. Subjects may recall aspects of the melody verbally, answering questions such as: how many

Stimulus

Recall

7.3

(7.3)

— = in one breath o = approximate pitch ~~~~ = quickenning of tempo

Source: Sloboda and Parker (1985), p. 165.

Figure 3.17: Recall 7.3

phrases did you hear; how many measures in each phrase; what was the tonality of the phrases; what was the meter; where there anacruses preceding any of the phrases; did any of the phrases resemble each other; can you clap the rhythm; can you sing the tonic of the first and last phrase; can you sing the first tone of each phrase; can you sing the last tone of each phrase; can you sing the first tone of each measure, and so on. Of course, some of the questions can be answered only by trained musicians, but, as was shown by Sloboda and Parker's experiment, aspects like phrasing were perceived also by nonmusicians. Such an experiment may also give more clues about the perception of specific musical aspects.

Chapter 4

Perception of Simultaneities

When two or more stimuli are competing for one's attention, it is most likely that one will assert itself as figure and the others as ground. The selection of the figure can happen involuntarily, but it is also possible to direct one's attention and alter the spontaneous figure/ground organization.

In the case of music, chords in isolation are treated as a single stimulus. Although some component tones may be more salient than others (mostly the highest pitch), unless a special effort is made in directing the attention, none of the tones is likely to become a figure. Thus, chords can be regarded as vertical simultaneities.

A more complex form of simultaneity is a harmonized melody, in which both vertical and horizontal aspects come into play. Francès has noted that in such cases, it is the richer form, or melody, that will become a figure, and the other melodic lines

will become a part of the ground.¹ Thus, one melodic line is treated horizontally, while the others are primarily vertical events rather than melodic lines.

4.1 Perception of “Vertical” Simultaneities

The mere existence of voice leading conventions attests that even the harmony, or the ground part, is not purely vertical. In the case of polyphony, there is not one single part that readily presents itself as a figure, nor do all parts blend into the ground. Rather, more than one figure is present at one time (or one span of time) and the listener’s attention shifts from one figure to another. However, even polyphony, the most “horizontal” form of musical simultaneity, has a very important vertical aspect.

4.1.1 Consonance and Dissonance

The term consonance denotes so many different sonorities, depending on discipline, time, culture, style, and views, that it can hardly be used without further definitions. James Tenney has attempted to untangle a semantic chaos in his historic view of consonance and dissonance and has listed the different “consonant–dissonant concepts” that have emerged over the history of Western music as *CDC1* through *CDC5*.²

Tenney maintains that each of these *CDC*’s is related to a musical practice

¹Francès, *The Perception of Music*.

²James Tenney, *A History of ‘Consonance’ and ‘Dissonance’* (New York: Excelsior Press, 1988).

prevailing at one time, and that most of these concepts were carried over, in one way or another, and added to the emerging new concept. Table 4.1 summarizes Tenney's perspective of the different concepts. *CDC4a* was referred to by Tenney as "a variant" of *CDC4*, which is represented by Riemann's theory.³ It can be seen that some of the concepts are context free (such as *CDC2*), while others are context dependent (such as *CDC3* and *CDC4a*).

Tenney believes that, with further research, it is possible to find acoustical correlates to each of the above concepts. Though physically based theories have existed for many years, there is no agreement about such correlates. The opposite view is that consonance and dissonance are products of the Western musical language and since this is not the only musical language in the world, it is not necessary, nor is it possible, to derive its syntax from nature.

Physically Based Theories

Theory of Simple Frequency Ratios Historically the term consonance goes back to the time of Pythagoras and the simple frequency ratios, but the link of this physically based theory to the musical consonance of the Common Practice style is far from being simple and direct.

Following simple frequency ratios, perfect consonances should be preferred over

³Hugo Riemann, "Die Natur der Harmonik," *Sammlung musikalischer Vorträge*, 40 (1882), cited in: James Tenney, *A History of 'Consonance' and 'Dissonance'*, New York: Excelsior Press, 1988.

Table 4.1: Tenney's Consonance–Dissonance Concepts

<i>CDC</i>	Time	Era	Entitive referents	Qualitive referents	Consonance Definition	Suggested Names
1	?B.C.	pre-poly-phonic	melodic intervals	relations between pitches	directly tunable	monophonic, melodic
2	10th century	early poly-phonic	simul-taneous dyads	sonorous character	sounds like a single tone	diaphonic
3	14th century	contra-puntal & figured bass	simul-taneous dyads	sonorous character	clarity in the lower voice of a contrapuntal texture	polyphonic, contra-puntal
4	18th century	triadic-tonal	indi-vidual tones	tendency toward motion	stability as a triadic component	triadic
4a	18th century	triadic-tonal		stability	tonic	functional
5	19th century		any sound	sonorous character	smoothness	timbral

Source: Tenney (1988), p. 115.

imperfect ones, which is not the case in Common Practice music. Also, any intonation that does not follow the just tuning system creates frequency ratios that are far from being simple. Not only is the tempered tuning system widely used in Common Practice music, but also, as was discussed in Section 3.2, musicians use a variety of frequency ratios for musical intervals. Were these intervals rendered their consonant nature by simple frequency ratios, deviations from these ratios would have been expected to create dissonances.

Natural Law Theory Another physical basis for a theory of consonance is the “Natural Law” theory which attempts to construct the theory of Common Practice harmony on the acoustical basis of the harmonic partials (See Section 5.1).⁴ Although complex tones are usually perceived holistically (and not analytically), the human ear does perceive the tone’s spectrum, at least as timbre. It is also possible to consciously focus on the partials and hear about the first six (see Figure 4.1). Yet, as tempting as it may be to regard the development of Common Practice harmony as natural, universal and inevitable as nature itself, the “Natural Law” theory presents too many discrepancies to hold.

The very first problem is that Common Practice harmony is *not* universal; the only universal aspect of consonance that may be related to partials is the octave. Following the order of the harmonic partials leads, again, to the expectation that perfect consonances will be preferred over imperfect ones. Moreover, if the major triad is indeed derived from partials 4,5,6, then one may expect that the $\text{}^6_4$ chord, derived from the more prominent partials 3,4,5 would enjoy a superior status to that of the triad, and not be treated as a dissonance. However, the major problem that this theory presents, is the lack of explanation for the status of the minor triad, the second most important triad, which cannot be easily derived from the partials. This

⁴N. Cazden, “The Definition of Consonance and Dissonance,” *International Review of Aesthetics and Sociology of Music*, 11 (1980): 123-168.

Figure 4.1: The first 16 harmonic partials

problem was addressed by Terhardt⁵ and later by Parncutt,⁶ as will be discussed later in Section 4.1.3.

Beats Theory The psychoacoustical explanation of consonance is based mostly on beats; Helmholtz maintained that a sonority is consonant when the harmonics of the tones coincide (which happens with pairs of tones in simple frequency ratios).⁷ Otherwise, adjacent harmonics within a certain frequency range create fast beats which give rise to a sensation of roughness, or dissonance. The width of this range, called critical bandwidth, varies by register (it is about three semitones around the register of A#⁴, narrower in lower registers, and wider in higher registers). This explanation cannot account for musical consonance either, mainly because musical consonance and dissonance are not restricted to simultaneities; it applies also to melodic sequences, and historically, melody preceded harmony. Again, any deviation

⁵E. Terhardt, "The Impact of Computers on Music: An Outline," in *Music, Mind and Brain*, ed. M. Clynes (New York: Plenum Press, 1982).

⁶R. Parncutt, "Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord," *Music Perception*, 6 (1988): 65-94.

⁷Helmholtz, *Sensations of Tone*.

from just tuning creates beats,⁸ and it was shown that such deviations do not affect the judgments of consonance.⁹ Furthermore, roughness is affected not only by frequency ratios, but also by the spectrum of the sound¹⁰ and by the register. The musical implication is that an interval may be consonant in a specific register or on a specific instrument, and dissonant in a different register or on a different instrument. Another reason why this type of psychoacoustical explanation fails to explain the musical consonance is that measuring consonance by roughness, or the presence of beats, leads to a concept of consonance and dissonance lying on a continuum. While in music some dissonant sonorities may be considered more dissonant than others, there is no continuum between the two concepts; a sonority is either dissonant or consonant. Finally, the beats theory cannot explain the musical phenomenon of enharmonic spelling; the same physical frequency ratio may be spelled enharmonically in more than one way, and regarded, accordingly, as a dissonance or a consonance.

In psychoacoustical research these problems are addressed by separating the discussion of what is called by researchers a sensory consonance,¹¹ or a tonal consonance¹² from that of the musical consonance (i.e., a consonance in a musical context).

⁸J. Vos, "Purity Ratings of Tempered Fifths and Major Thirds," *Music Perception*, 3 (1986): 221-258.

⁹J. Peterson and F. Smith, "The Range and Modifiability of Consonance in Certain Musical Intervals," *American Journal of Psychology*, 4 (1930): 561-572.

¹⁰Vos, "Purity Ratings."

¹¹E. Terhardt, "The Concept of Musical Consonance: A Link Between Music and Psychoacoustics," *Music Perception*, 1 (1984): 276-295.

¹²R. Plomp and W. J. M. Levelt, "Tonal Consonance and Critical Bandwidth," *Journal of the Acoustical Society of America*, 38 (1965): 548-560.

The research then focuses on the sensory consonance whose aspects are more measurable, but are of little relevance to musical syntax.

Musical Theories

Musicians' view of dissonance is not based on roughness versus pleasantness.¹³ Piston regards the concepts of consonance and dissonance as a function of stability: "A consonant interval is one which sounds stable and complete, whereas the characteristic of the dissonant interval is its restlessness and its need for resolution into a consonant interval."¹⁴ Later, he specifically rejects the notion of pleasantness as a basis for consonance: "it is the dissonant element which furnishes much of the sense of movement and rhythmic energy ... the essential quality of dissonance is its sense of movement and not, as sometimes erroneously assumed, its degree of unpleasantness to the ear."¹⁵ In a more recent harmony text book, Aldwell and Schachter define consonance and dissonance with no reference to pleasantness: "We call the stable intervals *consonances* or *consonant intervals*; the unstable ones are *dissonant*."¹⁶ One may argue that it is the unpleasantness that renders dissonances their instability. Dissonance, though, is a part of the musical syntax and it is impossible to construct syntax on

¹³M. Guernsey, "The Rôle of Consonance and Dissonance in Music," *American Journal of Psychology*, 40 (1928): 173-204.

¹⁴W. Piston, *Harmony*, 4th ed. (New York: Norton, 1978), 6.

¹⁵Piston, *Harmony*, 7.

¹⁶Edward Aldwell and Carl Schachter, *Harmony and Voice Leading*, 2nd ed. (New York: Harcourt Brace Jovanovich, Publishers, 1989), 27.

such a subjective concept. A dominant seventh chord, for example may be judged as a pleasant sonority (except for a case of musicians with prior knowledge that this is an “official dissonance” and that dissonance means an unpleasant sonority). Its instability, though, is clear to anyone who is exposed to the syntax of Western tonal music. As put by Cazden, the concept of musical consonance and dissonance does not refer to what these sonorities are, but rather to “how sonorities act, to what they do, to the perceptual expectations they arouse.”¹⁷ Gardner and Pickford’s experiment of dissonant judgment in a musical context, supports this view.¹⁸

Cazden’s view, though, is that any sonority may be heard in a specific context as either consonance or dissonance. By defining consonance as “stability” and dissonance as “instability,” Cazden also equates consonance with the tonic function, and dissonance with the dominant function. Thus, a major triad which functions as a dominant is, according to Cazden, a dissonance, and even a single tone, in a context that implies dominant harmony, is a dissonance. Cazden does not give examples, though, of typically dissonant sonorities (such as a tritone or a dominant seventh chord) being manipulated by their context to become consonants. He also does not refer to the function of subdominant, which is unstable, but does not point to a specific resolution the way dominant does.

¹⁷ Cazden, “Definition of Consonance and Dissonance,” 166.

¹⁸ A. D. Gardner and R. W. Pickford, “Relation Between Dissonance and Context,” *Nature*, 152 (1943): 358.

Research does show, though, that isolated intervals are judged as consonant or dissonant in the musical sense. In a series of experiments held with experienced musicians, Cohen *et al.* have found correlation between the amplitude of P_{300} and the degree of dissonance that matched the consonance concept of music rather than that of the sensory consonance.¹⁹ The amplitude for the imperfect consonance (m3) was smaller than that of the perfect consonance (P5), and largest for the dissonant interval (m2). The latency of the P_{300} was shortest for the dissonance and longest for the imperfect consonance.

The Semantics of Dissonance and Consonance

This variety of meanings given to the concepts of consonance and dissonance stems from their long term, interdisciplinary use. There is definitely a need for more specific terminology to describe the variety of concepts to which dissonance or consonance may refer. Some discipline-specific terms were introduced in the past, but they were used mainly by scholars and became part of the jargon of small specialized groups. Musicians and psychoacousticians still use the terms consonance and dissonance as if their sole meaning lies within their own discipline.

As was mentioned above, in psychoacoustic research, the “pleasant consonance” is also called “tonal consonance” or “sensory consonance,” as opposed to “musical

¹⁹D. Cohen *et al.*, “Response to Intervals as Revealed by Brainwave Measurement and Verbal Means,” *Journal of New Music Research*, 23, no. 3 (1994): 265-290.

consonance.” “Euphony” was also suggested for the psychoacoustic consonance.²⁰ Though terms such as “concordance” and “discordance” were introduced for the musical concepts,²¹ musicians simply use “consonance” and “dissonance.” There is no specific term for a “tonic implying consonance” or a “dominant type dissonance,” nor is there a term for a dissonance that, without a tonal context, does not imply a specific resolution. While it is true that a single interval may sound dissonant or consonant in different contexts, without context some sonorities are definitely dissonant, while others would sound primarily consonant, although they may be manipulated to sound unstable, or dynamic within a specific tonal context.

It is beyond the scope of this thesis to introduce new terminology, especially when even musicians do not agree on which of the variety of musical phenomena may be termed dissonant, consonant or neither. The following is a partial list of Common Practice musical phenomena that are referred to as consonant or dissonant.

1. Sonorities out of the context of a musical piece

- (a) A sonority that implies stability (such as octaves, fourths and fifths, thirds and sixths, major and minor triads)
- (b) A sonority whose only source of instability is the missing root (such as

²⁰Cazden, “Definition of Consonance and Dissonance.”

²¹C. Stumpf, “Konsonanz und Konkordanz,” *Beiträge zur Akustik und Musikwissenschaft*, 6 (1911): 116-150, cited in: N. Cazden, “The Definition of Consonance and Dissonance,” *International Review of Aesthetics and Sociology of Music*, 11 (1980): 123-168.

inversions of the major and minor triads)

- (c) A sonority that implies instability and a dominant-to-tonic resolution (such as the dominant seventh chord, the tritone, the minor seventh, the diminished triad and its inversions)
- (d) A sonority that implies instability, but does not imply a dominant-to-tonic resolution (such as the seconds and the major seventh, the augmented triad and the major and minor seventh chords)

2. Sonorities in the context of a musical piece

- (a) A sonority that implies a tonic function (may be a single tone, or any sonority of groups 1(a) or 1(b) above)
- (b) A sonority that implies a dominant function (almost any sonority which is a part of the Common Practice syntax may imply dominant function, except for the minor triad)
- (c) A sonority that implies subdominant function

This list is far from capturing the complexity of the consonance/dissonance concept. Many of the examples given for each group are not always interpreted in the same way. Most of them were given on the basis of the most common interpretation of that sonority. The tritone, for example has two resolutions, but since they are mutually exclusive, normally the listener will perceive one or the other resolution

as “implied;” A V_7 sonority, in a specific context and with an altered enharmonic spelling, will be interpreted by most theorists as a subdominant function (such as in the case of an Augmented Sixth Chord); A tonic-function sonority with a suspension is a dissonant sonority on the “pleasantness scale,” but its status on the “stability scale” may be a cause for a semantic argument.

Summary and Conclusions

It is clear, then, that there is a cultural component in the concept of the Common Practice consonance. On the other hand, the obvious links to some physical phenomena, as problematic as they may be, cannot be ignored. Cazden regards these links as “*limiting natural conditions*.”²² Terhardt’s *virtual pitch* theory, though, seems to suggest a more meaningful link between partials and the musical syntax.²³

Terhardt maintains that the human auditory system learns and stores the spectra of speech sounds as part of the process of acquiring a language. This learning process then enables the extraction of a virtual pitch even when some partial, or even the fundamental of a tone, are missing. In this process, the auditory system is exposed to intervals such as octaves and fifths. The particular harmonic language that developed in the music of the Common Practice style is just one possible language.

²²Cazden, “Definition of Consonance and Dissonance,” 161.

²³Terhardt, “Concept of Musical Consonance.”

4.1.2 Perception of Harmonic Intervals

The perception of melodic intervals and its categorical nature were discussed in Section 3.2, where they were compared to speech phonemes. Intervals, though, as opposed to phonemes, do carry meaning even out of any musical context. They can imply consonance or dissonance, major or minor; some intervals may imply certain tonalities; some intervals may render more stability to one of their component tones than the other (e.g., in a P5, the lower tone is normally perceived as more stable, while in a P4, the higher tone is perceived as more stable).

Although interval recognition and identification are readily measurable, there is not enough research to compare performance of melodic and harmonic intervals. In fact, there is surprisingly little research on harmonic intervals from a musical (rather than psychoacoustical) point of view. Different studies examined different aspects of interval perception but there is not enough data to draw conclusions as for the interaction of these aspects. It is clear, however, that musicians' perception of intervals is different than that of nonmusicians. Interval recognition is extremely hard for nonmusicians. As was mentioned in Section 3.2, musicians process music in the left brain hemisphere, and this was shown again in a harmonic intervals experiment held by Dalia Cohen *et al.*²⁴ Cohen *et al.* also found that musicians' response to a series of harmonic intervals was based mostly on the harmonic content, followed by the

²⁴Cohen *et al.*, "Response to Intervals."

melodic interval between the two harmonic intervals, and last by intensity (loudness). Nonmusicians' response was based on the reverse order of the above aspects, i.e., mostly by intensity, less by melodic content, and least by harmonic content.

Dimensions of Interval Perception

Psychoacoustic research tends to assume that intervals are recognized by their sizes.²⁵ However, there are many other aspects that influence interval recognition and the types of error made. Recognition is influenced by the degree of consonance or dissonance, same-scale-steps relations (e.g., m6 and M6), intervals inversions, tonal meaning and context.

Unfortunately, the data provided by the existing studies is insufficient to make fine classifications (such as examining the percentage of errors between dissonant intervals and their inversions) since these would reduce the results to numbers too small to base any conclusions on. Furthermore, such classifications alter the interactions with other dimensions. Consider examining the percentage of errors between dissonant intervals and other dissonant intervals, excluding errors of same-scale-steps. The possible errors will be, in such a case: m2-TT, m2-m7, m2-M7, M2-TT, M2-m7, M2-M7, TT-m7 and TT-M7. All of these pairs of intervals are four semitones apart

²⁵R. Plomp, W. A. Wagenaar and A. M. Mimpfen, "Musical Interval Recognition with Simultaneous Tones," *Acustica*, 29 (1973): 101-109; E. Terhardt *et al.*, "Tonhöhenmehrdeutigkeit, Tonverwandtschaft und Identifikation von Sukzessivintervallen," *Acustica*, 61 (1986): 57-66.

or more, which reduces the effect of interval size to a minimum. This problem may be overcome by using compound intervals (and thus eliminating the effect of size for all other intervals too), but this was not done in any of the studies cited here.

Interval Size Plomp and Levelt showed a case in which intervals' consonance was judged by their width, but this case is of little musical relevance.²⁶ The subjects were nonmusicians (the data of 38 of the 88 subjects were discarded for inconsistency) and the intervals were presented using simple tones (as opposed to complex tones). The subjects were asked to rate the intervals by their degree of dissonance, and the results were that intervals that exceeded the critical bandwidth were judged as consonant, and those that were within the critical bandwidth were judged as dissonant.

In an experiment published a year later, in which Plomp took part, complex tones were introduced, and the authors came to the conclusion that there was also a strong correlation with the complexity of the frequency ratios.²⁷ Since the conclusions were drawn from the data showing which intervals were judged as most similar, a few remarks on the method should be made.

First, the subjects were presented with groups of three intervals, and in each group they were asked to rate the intervals' similarity: two pairs were assigned the rating "most similar," and the other two were assigned the rating "least similar."

²⁶Plomp and Levelt, "Tonal Consonance."

²⁷W. J. M. Levelt, J. P. van de Geer and R. Plomp, "Triadic Comparisons of Musical Intervals," *British Journal of Mathematical & Statistical Psychology*, 19 (1966): 163-179.

In interpreting the data, it should be noted that this method forces the subjects to assume similarity even in cases they may regard none of the intervals as similar at all.

A second concern about the validity of the data interpretation is that the intervals were not constructed along a musical scale; although most of the intervals were musical intervals (twelve musical intervals, and three “in between” intervals), they were presented with both notes equally spaced around 500 Hz (a sharp B⁴). This setting may influence the similarity judgments in two ways: (1) a smaller difference in width would result in a smaller pitch difference between the upper tones and between the lower tones, and may result in a “more similar” judgment. (2) an increased interval width of an even number of semitones will result in an interval of the same chromatic scale, while an increased width of an odd number of semitones may be perceived as “out of tune” interval, and judged as less similar.

A third experiment conducted by Plomp *et al.* with trained musicians, brought them to the conclusion that both simple and complex tones are confused with intervals of similar size rather than similar frequency-ratio simplicity.²⁸ The interpretation of their data, though, shows clear effects of other dimensions mentioned above. Some of these effects were detected by the authors of the experiment, but referred to as “anomalities.” Furthermore, in a graph representing the average number of responses

²⁸Plomp, Wagenaar and Mimpfen, “Musical Interval Recognition.”

Table 4.2: Average number of responses for one and two semitones

Duration	Average no. of responses	
	1 semitone	2 semitones
120 ms	41	2
60 ms	39	2
30 ms	30	5
15 ms	15	8

Source: Plomp *et al.* (1973), p. 107.

for each distance (e.g., one semitone, two semitones etc.) it is clear that the main, and maybe only meaningful effect is that of an error of one semitone. The experiment had four different conditions of durations, and this phenomenon was clearest in the longest duration (120 ms), which is the duration that has the best chance to give rise to other effects than size. Table 4.2 shows the average number of responses for one semitone and two semitones (the average number of responses for any distance of more than two semitones was always smaller than that of the two semitones).

Plomp *et al.* report that the subjects “have tried to ‘hear out’ the two pitches in order to decide upon their interval width.”²⁹ This may be due to their specific training or to the way they were instructed at this specific experiment, but it may have affected the results to represent more accurately the perception of melodic, rather than harmonic intervals. Another strategy used by some students to identify melodic intervals (though not necessarily used by the subjects in this particular experiment)

²⁹Plomp, Wagenaar and Mimpén, “Musical Interval Recognition,” 107.

is “filling out” all the semitones between the two tones. Such strategy may easily result in errors based on interval size.

Consonance–Dissonance For the purpose of this discussion, two different approaches to consonance and dissonance will be taken, both relating to out-of-context harmonic intervals. The first approach regards m3, M3, P4, P5, m6, M6 and P8 as consonances (which may be further classified to perfect and imperfect consonances), and the rest of the intervals as dissonances. The second approach follows Ortmann’s classification, as follows:³⁰

- marked consonances: M3, P4, P5, M6 and P8
- mixed fusion (minor consonances and mild dissonances): m3, TT, m6 and m7;
these may be further classified as:
 - consonant mixed fusion: m3, m6
 - dissonant mixed fusion: TT, m7
- marked dissonances: m2, M2, M7

Again, with the amount of data available, some of the finer classifications may not be possible. In such cases, the discussion will be confined to assumptions and suggestions for further research.

³⁰Otto Ortmann, “Problems in the Elements of Ear-dictation,” *Research Studies in Music*, 2 (October 1934), 66-67.

The conclusions are drawn mostly from data provided by Plomp *et al.*³¹ and Ortmann.³² Although Ortmann's study is quite dated, he has provided an extensive amount of data and was able to demonstrate some consistent types of errors.

Most of the data provided by Plomp *et al.* are presented in a way that demonstrates the effect of the intervals size, but further interpretation shows a clear consonance–dissonance effect. Of the 690 identification errors made, 222 are between consonances, 198 are between dissonances, 143 are between consonances and dissonances, and 128 are between dissonances and consonances. It should be noted that there are seven consonances and only five dissonances. Since each interval was presented 180 times, this effects both the number of stimuli in each category and the number of options for confusion. Counting each possible pair in each category, the numbers were averaged, as shown in Table 4.3.

There is a marked difference between the dissonances–dissonances category and all other categories, which may point to the conspicuity of the dissonances; they are more often mistaken to be other dissonances than the consonances are mistaken to be other consonances. Further classification of the consonances to perfect and imperfect brings the average of the imperfect consonances closer to that of the dissonances, and that of the perfect consonances closer to that of the consonances–dissonances category. Two reservations should be made to this conclusion; (1) one must keep in

³¹Plomp, Wagenaar and Mimpfen, "Musical Interval Recognition."

³²Ortmann, "Problems in the Elements of Ear–dictation."

Table 4.3: Dissonance–consonance error types

Type of Error	a. Number of Errors	b. Number of Possible Pairs	Average (a divided by b)
Dissonances–Dissonances	198	20	9.90
Consonances–Consonances	222	42	5.29
Consonances–Dissonances	143	35	4.09
Dissonances–Consonances	128	35	3.66
Imperfect Consonances	100	12	8.33
Perfect Consonances	33	6	5.50

Source: Plomp *et al.* (1973).

mind that the more classifications are made, the less amount of data are available and the less reliable the conclusions are. (2) In the imperfect consonant group other dimensions are very likely to have an impact on the results, since each interval in this category has its counterpart in terms of same scale step, inversion and major–minor. In the perfect consonances category, on the other hand, there are no same scale step intervals, the inversion of the P8 is not considered an interval, and the major–minor relationship is irrelevant.

The brainwave–based experiment (see Section 4.1.1) supports the concept of the conspicuity of the dissonances; the P_{300} amplitude was largest in response to the dissonant interval (m2), and its latency shortest.³³ The conclusion was that the dissonant interval was the most conspicuous, and it was supported by subjects'

³³Cohen *et al.*, "Response to Intervals."

Table 4.4: Error types (Ortmann)

Type of Error	Number of Errors
Marked Dissonances	134
Marked Consonances	253
Mixed Fusion	572
Total	959

Source: Ortmann (1934).

report that this was the easiest interval to identify. The results for the perfect interval (medium amplitude and latency) do not match the conclusions drawn here based on the study of Plomp *et al.*, but this is not surprising considering the reservations mentioned earlier about these conclusions.

Ortmann's classification to consonant, dissonant and mixed-fusion intervals brought him to similar conclusions. As can be seen in Table 4.4, the least amount of within-type errors was made with the marked dissonances, followed by marked consonances. The greatest amount of errors is found in the category of the mixed fusion intervals which consists of mild dissonances and minor consonances. Further interpretation of Ortmann's data reveals that the smallest amount of errors (averaged by number of possible pairs) are made between any group of consonances (all consonances, marked consonances or consonant mixed fusion) and marked consonances (and vice versa).

Scale Steps Plomp *et al.* has shown that the one-semitone error is the most common one. However, separating same-scale-step errors from other one-semitone errors reveals that intervals are confused more often with the same scale step than with the interval which is one-semitone apart in the other direction. Table 4.5 demonstrates these differences both in numbers and in percentage. As can be seen in the last column ("Difference in %"), overall, the difference between same-scale-step errors and one-semitone errors decreases as the size of the intervals increases. Nevertheless, even with the largest interval that has a same-scale-step (M7), there are fewer one-semitone errors than same-scale-step errors (Since a one-semitone error is not possible in the case of the m2, this type of error was averaged for seven different intervals, rather than eight).

As for intervals that do not share a scale step with another interval, (P4, TT, P5, P8), it can be seen that except for the P8, their number of errors in each direction (to a smaller or a larger interval) is closer to the average number of the one-semitone error than to that of the same-scale-step error (Table 4.6). The higher number in the case of the P8 is in accordance with the tendency mentioned above of the larger size intervals to show a smaller difference between the two types of errors, but it may also be attributed to the fact that the one-semitone larger interval, the m9, was not a part of the test, and thus all the errors were made in the direction of the smaller interval.

Table 4.5: Same-scale-step versus one-semitone errors

Interval	Total no. of errors	Same scale-step	% of errors	One semitone	% of errors	Difference in %
m2	70.9	38.0	54	/	/	/
M2	74.2	36.9	50	6.0	8	42
m3	56.2	24.3	43	7.0	12	31
M3	54.5	24.7	45	5.4	10	35
m6	55.9	21.4	38	12.0	21	17
M6	54.0	21.7	40	12.6	23	17
m7	59.5	23.2	39	12.0	20	19
M7	66.0	21.9	33	15.8	24	9
Total	491.2	212.1	43	70.8	17	
Average		26.51	5	10.11	2	24

Source: Plomp *et al.* (1973).

Table 4.6: One-semitone errors for intervals that do not have another same-scale-step interval

Interval	smaller interval	larger interval	Total	Average (smaller/larger)
P4	5.4	9.3	14.7	7.4
TT	9.2	10.8	20.0	10.0
P5	11.0	12.2	23.2	11.6
P8	19.9	/	19.9	<i>9.9^a</i>
Total	45.5	32.3	77.8	
	<i>35.5</i>	<i>42.2</i>		<i>38.9</i>
Average (per interval)	11.4	10.8	19.5	
	<i>8.9</i>	<i>10.6</i>		<i>9.7</i>

Source: Plomp *et al.* (1973).^aItalic numbers denote errors of P8 which were divided by 2.

Separating Ortmann's data of one-semitone errors in the same manner, shows that same-scale-step errors were made over four times more often than other one-semitone errors.

Inversions All studies reported here used the middle octave (C^4 – C^5) for the stimuli. It may be assumed that using a lower register may give rise to more inversions errors than reported here. Yet, some of the data do show an inversion-equivalence effect.

Data from both Plomp *et al.* and Ortmann's studies show that P4 and P5 are confused significantly more often with each other than with the adjacent tritone. This phenomenon may be explained on the basis of perfect consonance and on the basis of inversion. If perfect consonance is the reason for the increased confusion, one should expect to see some of the same effect in the case of the P8. This effect may be expected to be reduced, though, due to the big difference in size between the P8 on one hand and the P4 and P5 on the other hand. In fact, there seem to be no increase in confusion between these two intervals (P4, P5) and P8.

Other intervals, mostly the dissonances, also show an increased number of errors with their inversion. In the case of the m2, m7 and M2, their inversions have the greatest number of errors after the same-scale-step and one-semitone errors. The M2 has the same number of errors with its inversion as the number of errors with

m3 (one-semitone) and with P4. The only consonance that shows some effect of the inversion is the m6 that has the third largest number of errors with its inversion, following the same-scale-step and one-semitone errors.

Stronger evidence of the perceptual equivalence of intervals was provided by Deutsch and Roll who presented musically untrained listeners with pairs of tones, accompanied by a lower tone in the other ear, which they were instructed to ignore.³⁴ Their task was to determine whether the two attended tones were the same, or different. It was found that “different” tones tended to be judged as the same not only when presented with the same intervals, but also when presented with intervals and their inversions. Deutsch and Roll concluded that intervals and their inversions were treated as equivalent.

Other Dimensions While some of the dimensions discussed above may be related, at least to some extent, to corresponding acoustical or perceptual phenomena (such as size-frequency ratio, consonance-fusion, inversions-octave equivalence) others are specific to the musical syntax discussed in this thesis (such as same scale steps). There are more syntax-specific ways in which some intervals may be perceived as related to each other and thus expected to be confused with each other.

One such dimension is intervals that imply similar tonal meaning (TT-m7, M3-

³⁴D. Deutsch and P. L. Roll, “Separate ‘What’ and ‘Where’ Decision Mechanisms in Processing a Dichotic Tonal Sequence,” *Journal of Experimental Psychology: Human Perception and Performance*, 2 (1976): 23-29.

P5). The effect of this dimension is not expected to be as marked as other dimensions discussed above, and it may be possible that the amount of data is insufficient to draw such specific conclusions about two pairs of intervals. Although the number of errors reported by Plomp *et al.* for P5-M3 is remarkably high, there is no such significant effect in their study for the pair TT-m7.

Another syntax-specific dimension is that of the major-minor. It is possible that this dimension may effect the numbers of errors for the pairs m3-m6 and M3-M6. Probably a special study needs to be designed to determine whether such effect exists. Such an experiment would have to be able to link the consonant major intervals to each other (M3-M6) rather than to their inversions (M3-m6, M6-m3), and the same with the consonant minor intervals (m3, m6). In the studies reviewed here there are some clues of a major-minor effect, but again insufficient to draw strong conclusions. In Ortmann's data, the m3, m6 and M6 are confused more often with the other major or minor interval than with their inversions, and the M3 is the only interval that is confused more often with its inversion than with the M6. In the data presented by Plomp *et al.*, such effect exists in the m3 and M6, but not in the M3 and m6.

Summary

The effect of size on the identification of intervals cannot be ignored, but it is definitely not the only dimension effecting the identification. It is most marked in one semitone

differences, and as the differences get larger, other dimensions take over. To determine and quantify the effect of all these dimensions and their interactions, more extensive and specified research is needed.

4.1.3 Chord Perception

Technically, chords can be defined as two or more simultaneous intervals, but since chords are highly idiomatic in tonal Western music, this is a case in which the whole is more than the sum of its parts. Their representation in the memory is on a cognitive, rather than a perceptual level.³⁵ It is not surprising then, that it takes longer to create a representation of a chord than that of a single tone,³⁶ that chords are retained in memory better than single tones,³⁷ and that there are marked differences in the perception of chords between musicians and nonmusicians.

Thomas Pechmann has extended Deutsch's experiments with interference tones (see Chapter 2) to chords with interference tones, with the assumption that being a more complex stimuli, chords would be less vulnerable to interference than single tones.³⁸ The decrease in percentage of errors from tones to chords found in

³⁵J. Bharucha, "Music Cognition and Perceptual Facilitation: A Connectionist Framework," *Music Perception*, 5 (1987) 1-30; J. Bharucha and K. Stoeckig, "Priming of Chords: Spreading Activation or Overlapping Frequency Spectra?" *Journal of Experimental Psychology: Human Perception and Performance*, 14, no. 6 (1987): 519-524.

³⁶T. Hubbard and K. Stoeckig, "Musical Imagery: Generation of Tones and Chords," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, no. 4 (1988): 656-667.

³⁷Thomas Pechmann, "Memory for Chords: The Retention of Pitch and Mode," *Music Perception*, 16 (1998): 43-54.

³⁸Pechmann, "Memory for Chords."

Pechmann's results was not dramatic, but statistically significant (30% versus 25% errors).

As with harmonic intervals, chords may be identified along various dimensions: their intervallic structure, their degree of dissonance or consonance, their position (root position or inversion), the pitch of their root (relative to the root pitch of the previous chord), the degree of ambiguity of their root, their type (major, minor, diminished or augmented for triads, and the different types of the 7th chords), and their mode of presentation (melodic or harmonic). The chord type may be regarded as a result of other dimensions such as intervallic structure, degree of consonance, or degree of root ambiguity, but it may also be regarded as a dimension with some degree of independence from the other dimensions.

Dimensions of Chord Perception

The assumption that a chord is represented and perceived by its intervallic structure was investigated by Roberts and Shaw, by comparing results of a same/different test to two different models; the first is based on the size of the component intervals of the four triads, and the second is based on the consonance/dissonance of the component intervals.³⁹

³⁹L. A. Roberts and M. L. Shaw, "Perceived Structure of Triads," *Music Perception*, 2 (1984): 95-124.

Intervallic Structure Based on the similarity of intervals size, the most similar pair of triad is the major/minor pair. Next are the major/augmented and minor/diminished pairs, followed by the major/diminished and minor/augmented pairs and last is the diminished/augmented pair.

Table 4.7: Discrimination test results for highly trained subjects, compared with the two models

Triads Pair	Discrimination Test Results		Rank Predicted by Interval-Size Model	Rank Predicted by Consonance-Dissonance Model
	Rank	% Correct		
M/dim	1	84.2	2-3	1-2
M/Aug	2	83.3	4-5	1-2
M/m	3	72.5	6	5-6
m/Aug	4	70.0	2-3	3-4
m/dim	5	67.3	4-5	3-4
dim/Aug	6	55.9	1	6

Source: Roberts and Shaw (1984), p. 107.

In the first four columns of Table 4.7 the results of the “highly trained” subjects group in that experiment are compared with the interval size models. The results are presented as the percentage of correct responses averaged for the nine subjects. Pairs that are ranked high in their dissimilarity (e.g. dim/Aug) are expected to have the greatest percentage of correct responses. The data show clearly that this model cannot account for the results. Roberts and Shaw’s second experiment required the subjects to rate the similarity of pairs of triads, assigning the value 9 to a “very similar” pair, and 1 to a “very dissimilar” pair. The results are presented in Table 4.8 with the

Table 4.8: Similarity test results for moderately trained subjects compared with the interval size model

Triads Pair	Similarity Test Results		Rank Predicted by Interval- Size Model
	Rank	Rating	
M/m	1	5.76	1
M/Aug	2	5.52	2-3
m/dim	3	5.33	2-3
M/dim	4	5.12	4-5
m/Aug	5	5.08	4-5
dim/Aug	6	4.94	6

Source: Roberts and Shaw (1984), p. 118.

similarity rating averaged for the nine “moderately trained” subjects. This group did show some influence of interval size, however, the differences between the pairs were less marked.

Consonance–Dissonance In the same study, results were compared to a consonance–dissonance model based on Helmholtz⁴⁰ and on empirical evidence found by Roberts.⁴¹ According to this model, the major triad is the most consonant, followed by the minor triad, and the diminished and augmented triads are rated as very dissonant. This model results in three different types of triad pairs: C,C (consonant–consonant); C,D (consonant–dissonant); and D,D (dissonant–dissonant). Looking at

⁴⁰Helmholtz, *Sensations of Tone*.

⁴¹L. A. Roberts, “Consonance and Dissonance Judgments of Musical Chords” (Paper presented at the 105th meeting of the Acoustical Society of America, 1983), cited in: L. A. Roberts and M. L. Shaw, “Perceived Structure of Triads,” *Music Perception*, 2 (1984): 95-124.

the “different” condition (pairs of two different chord types) in the discrimination task, results were best for pairs containing a major triad and a dissonant triad, and worst for pairs containing the two different dissonant triads (diminished and augmented). As can be seen in the right most column of Table 4.7 (page 106), the major/minor pair is the only one whose ranking does not match the consonance–dissonance model. Roberts and Shaw noted that performance was better for all pairs containing a major triad, and they suggested that this may be attributed to the major triad’s superior ability to define tonality. It may be enough to note that the major triad is the most important triad in Common Practice Western music to explain its superior perceptual status, and that the minor triad is the second most important triad.

A different chords consonance–dissonance model, developed by Terhardt,⁴² Terhardt *et al.*,⁴³ and later by Parncutt,⁴⁴ regards the degree of dissonance of a chord as a function of the degree of its root ambiguity. The model (which will be explained in more detail below), ranks the consonance of triads with the major triad first, followed by minor, augmented and diminished. Both Helmholtz and Roberts and Shaw did not specify which of the dissonant triads is more dissonant, though it seems (based on anecdotal experience only) that musicians tend to perceive the augmented triad

⁴²Terhardt, “Impact of Computers on Music.”

⁴³E. Terhardt, G. Stoll and M. Seewann, “Algorithm for Extraction of Pitch and Pitch Salience from Complex Tonal Signals,” *Journal of the Acoustical Society of America*, 71 (1982): 679-688.

⁴⁴Parncutt, “Revision of Terhardt’s Psychoacoustical Model.”

as more dissonant than the diminished one. On the dimension of degree of consonance, then, the major and augmented triads (or major and diminished, according to Parncutt's model) are, perceptually, the most distant from each other. Comparing both models (Parncutt's model and the anecdotal-experience based model) with their specific ranking of the dissonant triads to the discrimination test results (Table 4.9) shows one difference in each model: Parncutt's model places the m/dim pair before the m/Aug pair, while the anecdotal-experience based model places the M/Aug before the M/dim pair. In this sense, it seems that there is a better match with the anecdotal-experience based model, since, as can be seen in Table 4.7, the difference in percentage of errors between the M/dim and M/Aug is smaller than that between the m/dim and the m/Aug. There may be another reason, though, why the M/dim pair precedes the M/Aug pair. Although it seems that the augmented triad tends to be perceived as more dissonant than the diminished triad on the "pleasantness" scale, the diminished triad may be perceived as the more direct opposite of the major triad, on the "stability" scale (a tonic-implying versus dominant-implying chord), and thus regarded as perceptually more distant from it.

As was seen in Section 4.1.2, the greatest number of errors between harmonic intervals occurred between two dissonances, followed by two consonances, and the smallest number of errors occurred between a dissonance and a consonance. In the case of triads, there is only one D,D and one C,C type of pairs possible. Averaging

Table 4.9: Ranking comparison of the discrimination test results, Parncutt's model, and the anecdotal-experience based model

Triads Pair	Rank		
	Test Results	Parncutt	Anecdotal
M/dim	1	1	<i>2^a</i>
M/Aug	2	2	<i>1</i>
M/m	3	5	5
m/Aug	4	<i>4</i>	3
m/dim	5	<i>3</i>	4
dim/Aug	6	6	6

^aItalic numbers indicate deviations of the models from the test results.

the percentage of errors presented in Table 4.7 for the four possible pairs of the C,D type shows that this order is maintained, with 55.9% correct responses for the D,D type pair, 72.5% correct responses for the C,C type pair, and an average of 76.2% correct responses for the C,D type pairs.

Position In Roberts and Shaw's study, discrimination was better for root position than for first inversion triads. This effect was found in all levels of musical training, but was most marked for the highly trained musicians group. This group also maintained the same confusion pattern that was found for the root position condition. In the similarity rating test, Roberts and Shaw found that triads were rated as more similar with same-type triads (major, minor, diminished or augmented) than with same-position triads.

Roberts and Shaw interpreted these results as suggesting that triads, even in

first inversion, are represented in the memory in root position. Their alternative explanation for the discrimination test results was that the poor performance for the first inversion might have been a result of a reduced salience of the dissonance or consonance of the triads. This explanation seems more plausible since it is common for listeners to be able to identify the type of a chord without being aware of its position, or being able to sing its root.

Root Pitch Since the above cited experiments were conducted with pairs of triads, the relationship between the pitches of the roots of the two triads may have impacted the results. Experiments with interference tones showed clearly that the pitches of the roots were retained better than the modes of the triads.⁴⁵ This effect was present in Roberts and Shaw's similarity rating experiment, for the moderately trained subjects group that rated triads with same root pitch as more similar than triads of the same type (the effect on the untrained subjects group was not discussed).

Roberts and Shaw also studied effects of the relations of the suggested keys of the roots using the data from the discrimination test. Following the Schenkerian view that the root of a major or a minor triad is assumed to be the tonic, Roberts and Shaw experimented with different key relations. Thus, triads with C as a root were compared with triads with the roots of C', F, G, A and F#. Significant results were obtained for the highly trained subjects group. For major and minor triads,

⁴⁵Pechmann, "Memory for Chords."

Figure 4.2: The two diminished triads, enharmonically interpreted as parts of one diminished 7th chord.

performance for the “same” condition (where both triads were of the same type) followed the pattern of the circle of fifths; best for the octave condition, followed by the relations C/F, C/G, C/D, C/A and worst performance for C/F#. In all types of relationships, performance was better for the major triads than for the minor triads.

The effect of the circle of fifths was maintained when results were averaged for all the four chord types, but observing Roberts and Shaw’s graph of each chord type separately reveals that in the relationships of C/A and C/F# the performance was better for the diminished triad than for the minor triad (and in the case of C/F# it was even better than that of the major triad).⁴⁶ This effect was not discussed or further experimented by Roberts and Shaw, but these pairs of diminished triads may be perceived as part of the same diminished 7th chord (see Figure 4.2) and thus their similarity may be enhanced.

Studying the data of the “different” condition, Roberts and Shaw concluded that there was no effect of root relationships, but, they did find some effect of “‘good’

⁴⁶Roberts and Shaw, “Perceived Structure of Triads,” 110.

versus 'poor' relationships."⁴⁷ There were more erroneous "same" responses when the second triad was possible in the key of the first triad. It seems that this effect is not much different from the root relationships effect of the "same" condition for the major and minor triads, since the order of the circle of fifths also reflect, to some extent, the possibility of these triads in the key established by the first triad of each pair. Since the dissonant triads do not establish their roots as a tonic, it is not surprising that in the case of augmented triads pairs, this effect is greatly reduced (though still apparent), while in the case of the diminished triads, the C/C' pair is the only one that reflects the effect of root relationships.

Roberts and Shaw also pointed out asymmetries, such as more erroneous "same" answers for C-Dm than for Dm-C, since the first is a more common chords succession. Such asymmetry may also effect pairs such as C/F (or F/C) and C/G (or G/C), since both these pairs may be interpreted as I-V or IV-I in one temporal order (F-C and C-G), or as the more common and complete succession V-I in the other order (C-F and G-C). Indeed, the graph presented by Roberts and Shaw shows that in the "same" condition, the major and minor pairs C/F got more correct responses than the pairs of C/G.⁴⁸ It is not clear, though, whether the graph represents specifically the order C-G or whether it represents the average for the pairs C/G and G/C.

⁴⁷Roberts and Shaw, "Perceived Structure of Triads," 111.

⁴⁸Roberts and Shaw, "Perceived Structure of Triads," 110.

Root Ambiguity As was mentioned above, Terhardt,⁴⁹ Terhardt *et al.*,⁵⁰ and later Parncutt⁵¹ linked the degree of dissonance of chords to the degree of their root ambiguity. Parncutt maintains that since the minor triad is not a “distorted form of the major triad,”⁵² a theory explaining the nature of the root of the triad should be able to account for the root of the minor triad as well as that of the major triad. This is not the case with the theory that links the root to the fundamental, since with C as a fundamental, the only partials that construct a minor triad are partials number 6,7,9 (which construct an “out of tune” G minor triad) and 10,12,15 (which construct an E minor triad), none of which have the C as their root.

Terhardt holds that though we are unaware of the partials of a tone, our auditory system is trained to recognize the spectral patterns of complex tones, and thus is able to decide, when presented with a simultaneity, which of the components are part of which fundamental.⁵³ His model is based on the first ten subharmonics (the 7th one is categorized to the nearest tempered semitone),⁵⁴ which, as shown in Figure 4.3a and 4.3b, when octave-generalized, produces five interval classes above the fundamental: P8 (subharmonics 1,2,4,8), P5 (subharmonics 3,6), M3 (subharmonics 5,10), m7 (categorized subharmonic 7) and M2 (subharmonic 9). Figure 4.3c demon-

⁴⁹Terhardt, “Impact of Computers on Music.”

⁵⁰Terhardt, Stoll and Seewan, “Aural Key Identification.”

⁵¹Parncutt, “Revision of Terhardt’s Psychoacoustical Model.”

⁵²Parncutt, “Revision of Terhardt’s Psychoacoustical Model,” 66.

⁵³Terhardt, “Impact of Computers on Music.”

⁵⁴See Figure 4.1 on page 82

a. The first 10 subharmonics

b. The octave generalized subharmonics

c. The 5 subharmonics of C

Figure 4.3: Terhardt's model of subharmonics

strates how the subharmonics of a specific tone are derived by calculating these five intervals below the tone.

To find a root of a chord, the subharmonics of all its component tones are searched for matches, and the subharmonic that is supported by the greatest number of the chord tones is the most likely root. In Figure 4.4, for the C major chord, the tone C is supported by all three tones (C,E,G), while F, D and A, which are also diatonic to C major, are supported only by two of the chord tones. B, the least stable tone in C major, is not supported by any of the tones.

As Parncutt points out, this model fails to account for the root of the minor triad.⁵⁵ Figure 4.5 demonstrates that following this model, F is the best supported

⁵⁵Parncutt, "Revision of Terhardt's Psychoacoustical Model."

Figure 4.4: Possible roots for C major

Figure 4.5: Possible roots for C minor

root for Cm, while C is supported by only two chord tones, as are $E\flat$ and $A\flat$. Another deficiency of the model, according to Parncutt, is its failure to reflect the root ambiguity of the half-diminished 7th chord (the only tone supported by the four chord-tones in a C_7 half-diminished is $A\flat$, which turns the chord into a V_7^9).

Parncutt's criticism of the model is that it identifies possible roots but does not weight them appropriately. Terhardt *et al.* developed a model in which higher subharmonics (which are less audible) are rendered less weight than lower ones.⁵⁶ Parncutt used the method of weighting the subharmonics a little differently, and rendered these intervals the following weight values: $P1=1$, $P5=1/2$, $M3=1/3$, $m7=1/4$, $M2=1/5$. These intervals were named by Stuckey "root-support intervals,"⁵⁷ and they

⁵⁶Terhardt, Stoll and Seewan, "Aural Key Identification."

⁵⁷R. Stuckey, "Music in Cyclic Twelfths" (Unpublished manuscript, London, 1983), cited in: R. Parncutt, "Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord," *Music Perception*, 6 (1988): 65-94.

Figure 4.6: A m3 as an indirect root support

Table 4.10: Ranking of root ambiguity for triads and 7th chords.

Triads		7th Chords	
Chord	Weight	Chord	Weight
Major	2.0	V ₇	2.1
Minor	2.1	Minor	2.3
Aug.	2.3	Major	2.4
Dim.	2.5	Half Dim.	2.4
		Dim.	2.9

Source: Parncutt (1988), p. 79.

also included the m3 which Parncutt regarded as an indirect root support through its harmonics (see Figure 4.6) with the value of 1/10. The remaining six intervals, which contributed to the root ambiguity of the chord were named “detractors.”

Parncutt regards this model as able to estimate the root ambiguity of chords as well as their degree of dissonance, and the salience of their component pitch classes. A chord’s root ambiguity is low when one of its pitch class’ weight is significantly higher than that of all the other possible roots. It is lower for harmonic intervals and increases as the number of pitch classes in the chord increases. Table 4.10 shows the ranking of triads and 7th chords according to this model.

Although Parncutt maintains that this ranking reflects “expectations based on music theory,”⁵⁸ it is very likely that musicians will debate his conclusions, both as a basis for degree of dissonance and as a basis for root ambiguity.

As was discussed in Section 4.1.1, there is no objective measure of consonance and dissonance. The question of the ranking of the dissonant triads was discussed earlier; some musicians may regard the major 7th chord as more dissonant than the half-diminished and diminished 7th chords. As for root ambiguity, the main deviation is that of the augmented triad with its root ambiguity rated as less than that of the diminished triad.

Chord Type The two dimensions that were shown to interact with chord type were the degree of consonance and the degree of root ambiguity. It does not necessarily imply, though, that these are the only two determinants of the chord type. It was shown above that the major triad, and maybe also the minor triad, have a special perceptual status that could not be accounted for by the consonance–dissonance, or by the interval size model. There is no agreement on the ranking by degree of dissonance for the diminished and augmented triad. Yet, the highly trained musicians group in Roberts and Shaw’s similarity experiment rated them as less similar than the major and minor triads (discrimination test results, though, were better for the major/minor than for the diminished/augmented).

⁵⁸Parncutt, “Revision of Terhardt’s Psychoacoustical Model,” 80.

Pechmann's study showed that the retention of chord type was poor.⁵⁹ This study was conducted with major and minor triads only, and as Pechmann noted, the stimuli were acoustically simple. This effect of timbre is known to be significant in judging similarity of chords. Beal has shown that comparison of chords played on different instruments was difficult even for trained musicians.⁶⁰

It seems that the validity of chord type as an independent dimension needs to be studied further. It may be that interference type experiments (see Deutsch⁶¹ and Pechmann) with chords and interference chords may reveal some information about retention, consolidation and interference of chord types.

Mode of Presentation A common strategy for music students is the arpeggiation of chords in order to facilitate their identification. It is questionable, though, to what extent can a simultaneous and an arpeggiated chord be regarded as an equivalent stimuli. Roberts and Shaw's study showed that for both highly trained and moderately trained musicians, melodic presentation improved discrimination for the first inversion chords, but had the opposite effect on the root position chords. It may demonstrate that for a familiar form, arpeggiating impairs the performance.

Hubbard and Stoeckig showed that it takes longer to form a mental image of a

⁵⁹Pechmann, "Memory for Chords."

⁶⁰A. L. Beal, "The Skill of Recognizing Musical Structures," *Memory and Cognition*, 13 (1985): 405-412.

⁶¹Deutsch, "Tones and Numbers."

chord than of a tone, but that this time period is not long enough to form images for all the component tones and then combine them into a chord.⁶² They concluded that “images could have resulted from the activation of previously formed, stored complex representation of chords.”⁶³

Summary and conclusions

The cognitive level of chord perception results in many differences between musicians and nonmusicians. Howard *et al.* has shown that while musicians performed somewhat better than nonmusicians in a major/minor discrimination task (in which differences of less than a semitone in the 3rd of the triad were presented on a continuum), their performance was much better in a labeling task, and it was only the trained musicians subjects group that showed evidence of categorical perception.⁶⁴

It seems that chord type, whether influenced by consonance–dissonance, root ambiguity, or as an independent dimension, is not a prominent perceptual entity for nonmusicians. In Roberts and Shaw’s similarity test, moderately trained musicians made their similarity judgments by the pitch of the chords’ roots rather than the chord type.

Last, as seen by the different conclusions drawn from Roberts and Shaw’s study,

⁶²Hubbard and Stoeckig, “Musical Imagery.”

⁶³Hubbard and Stoeckig, “Musical Imagery,” 660.

⁶⁴David Howard, Stuart Rosen and Victoria Broad, “Major/minor Triad Identification and Discrimination by Musically Trained and Untrained Listeners,” *Music Perception*, 10 (1992): 205-220.

the two tasks, discrimination and similarity judgment yielded different results. It does not follow from chords that are confused with each other in a discrimination task, to be judged as highly similar in a similarity task.

4.2 Perception of “Horizontal” Simultaneities

As was discussed in Section 4.1, any musical simultaneity has a vertical aspect to it, which varies depending on the texture (harmony, polyphony, and an endless variety of textures in between the two). For a figure to be formed in our perception, some selection needs to take place, and this selection may even take place when no real simultaneities are present.

4.2.1 Channelling Attention

It is common knowledge that there are limits to the capacity of our attention, but psychologists hold different views as to what this limit is and how selection and rejection of stimuli is done. The most extreme view, presented by D. E. Broadbent, is that there is a single all-purpose channel that is capable of attending to one task at a time, and that any illusion of multi-tasking is in fact due to rapid shifts between the tasks.⁶⁵ The strongest supporting evidence of this view is the human incapacity

⁶⁵D. E. Broadbent, *Perception and Communication* (London: Pergamon, 1958); D. E. Broadbent, *Decision and Stress* (London: Academic Press, 1971).

of attending to two verbal messages simultaneously.⁶⁶ As to the selection or rejection of the messages, Broadbent held that this was done by filtering the messages by their characteristics.

This theory was challenged by studies that showed capacity of attending to two different types of stimuli, such as speech and pictures, or speech and music.⁶⁷ These experiments led to an alternative theory of task-specific attention channels by which it is possible to attend to more than one task provided the different tasks use different attention channels. As Sloboda has pointed out, the incapacity to attend to two different verbal messages is tended to by the social conventions of conversation, but there are no such conventions in the case of music: two or more simultaneous melodies are not only permissible but desirable.⁶⁸

The theory of the task-specific attention channels does not exclude Broadbent's concept of filtering, but it places the filter in the specific channel rather than in the assumed single all-purpose attention channel. For a "figure,"⁶⁹ or a "channel of attention"⁷⁰ to form, the series of stimuli has to have some common features. A figure may be disrupted by "attention grabbing" features (on the attraction of attention, see

⁶⁶E. C. Cherry, "Some Experiments on the Recognition of Speech, with One and Two Ears," *Journal of the Acoustical Society of America*, 25 (1953): 975-979.

⁶⁷D. A. Allport, "Attention and Performance," in *Cognitive Psychology: New Directions*, ed. Guy Claxton (London: Routledge and Kegan Paul, 1980).

⁶⁸Sloboda, *Musical Mind*.

⁶⁹Francès, *The Perception of Music*.

⁷⁰D. Deutsch, "Memory and Attention in Music," in *Music and the Brain*, eds. M. Critchley and R. Henson (London: Heinemann, 1977).

also Bharucha's model in Section 6.3). *Gestalt* principles of perceptual organization such as the law of similarity, law of proximity and law of good continuation are readily applicable to musical streaming, together with other perceptual principles, and music specific features.

Dowling *et al.* developed an auditory model of directing attention.⁷¹ In their model of "expectancy windows," attention is voluntarily directed to a specific pitch range or time frame. Since knowledgeable expectations were shown to improve perception (see also Chapter 3),⁷² it is expected that tones outside the "expectancy windows" will not be perceived as well as those that are within the pitch range or time frame.

In a series of experiments listeners were presented with various tasks involving familiar melodies interleaved with distractor tones.⁷³ Although the melodies and the distractors shared pitch range and timbre, listeners were able to follow the stream of the familiar melodies. Dowling *et al.* compared this to the visual phenomenon demonstrated by Neisser in which two interleaved basketball games were shown on a video, with two different players wearing different outfits.⁷⁴ Observers were able to stream the stimuli into the two separate games based only on the kinetic information.

⁷¹Dowling, Lung and Herrbold, "Aiming Attention in Pitch and Time."

⁷²D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (New York: Wiley, 1966).

⁷³Dowling, Lung and Herrbold, "Aiming Attention in Pitch and Time."

⁷⁴Ulric Neisser, *Cognition and Reality; Principles and Implications of Cognitive Psychology* (San Francisco: W. H. Freeman, 1976).

In the same way, listeners were able to follow the stream of the familiar melody, based only on melodic and rhythmic information.

When not directed one way or another, listeners seemed to choose to narrow their focus on time more than on pitch.⁷⁵ This resulted in more correct responses about altered pitches in an expected time, than about altered times in an expected pitch. When directed to focus on one dimension, though, both dimensions had the same effect: detection of pitch alterations was better in expected (versus unexpected) times, and detection of time alterations was better in expected pitches.⁷⁶

Attention may be directed to any feature that distinguishes one melodic line from the others, or that is common to a group of tones, such as pitch, timbre, location, loudness and familiarity. It may be directed to changes as well as tend to continue with a melodic line that has already established itself as a figure.

Pitch Pitch range was shown to articulate figures even when the articulation was in contrast with other perceptual principles. Deutsch has presented listeners with a simultaneous ascending and descending major scales through ear phones, alternating the high and low tones between the right and left ear (Figure 4.7a).⁷⁷ The most commonly perceived figures were of two simultaneous melodies, one of which was constructed by the low tones and the other by the high tones (Figure 4.7b). In a

⁷⁵Dowling, Lung and Herrbold, "Aiming Attention in Pitch and Time," experiment 5.

⁷⁶Dowling, Lung and Herrbold, "Aiming Attention in Pitch and Time," experiment 6.

⁷⁷Deutsch, "Memory and Attention in Music."

Figure 4.7: Perception of simultaneous ascending and descending scale

study of perception of polyphony, Gregory showed that the melody of the higher pitch range was recognized more often than that of the lower pitch range.⁷⁸

The special status of the high pitch tones can be explained on the basis of masking.⁷⁹ Although both the highest and lowest parts are outer voices, the highest is more prominent because the high tones mask each other less than the low tones, since their critical bandwidth is narrower (see Table 4.11).

Figure 4.8 demonstrates two alternative perceptions of melodic lines using Deutsch's stimulus. In Figure 4.8a the spatial location is assumed to dominate the perception, while in Figure 4.8b the most familiar form, which also follows best the *Gestalt* law of good continuation, is assumed to dominate the perception.

⁷⁸ A. H. Gregory, "Listening to Polyphonic Music," *Psychology of Music*, 18 (1990): 163-170.

⁷⁹ Sloboda, *Musical Mind*, 172-173.

Table 4.11: Average masking ranges of some pitches.

Pitch	Masking range
A2	octave
A3	perfect fifth
A4	major third
A5	major second
A6	minor second

Source: Scharf (1970), p. 161.

a. Spatial location

b. Familiar form, good continuation

○ = Right ear ● = Left ear

Figure 4.8: Alternative melodic lines

Dowling's experiments with interleaved melodies demonstrated how pitch range can articulate figures and create polyphony even in the absence of true simultaneities.⁸⁰ Creating polyphony by compound melodies was a common practice in the Baroque era. Examining pieces such as Bach's partitas for violin solo or suites for cello solo reveals that the two parts do not overlap in pitch, and are always at an interval of no less than three semitones.

Dowling created interleaved melodies by combining pairs of familiar melodies and alternating their tones. These pairs of melodies were then presented to the listeners with various degrees of pitch overlapping (Figure 4.9). Dowling found that the melodies were separated from the distractor tones and identified when their ranges were separated by an average of a little over one semitone.

Dowling *et al.* showed how the focus of attention can alter the effect of the pitch range.⁸¹ A change in contour or pitch of one target tone resulted in that tone being out of the expected range. When asked to determine whether there was a change in direction (contour) or whether the pitch was higher or lower than that in the original melody, the listeners reported that the pitch has seemed to "disappear," and the percentage of correct responses was low. On the other hand, when the focus of attention was narrowed by eliminating the distractor tones, the out-of-range pitches were very salient, and the percentage of correct responses was high.

⁸⁰W. J. Dowling, "Perception of Interleaved Melodies," *Cognitive Psychology*, 5 (1973): 322-337.

⁸¹Dowling, Lung and Herrbold, "Aiming Attention in Pitch and Time."

a. Overlapping pitch ranges

Frere Jacques

Twinkle Twinkle

Range

b. Non-overlapping pitch ranges

Frere Jacques

Twinkle Twinkle

Source: Dowling (1973), p. 325.

Figure 4.9: Interleaved melodies

Rate of Presentation A simpler form of the separation into distinct streams (also called “fission”), is the trill, which may be regarded as two interleaved pitches. Experiments with this relatively simple stimulus show that the pitch range needed for creating fission depends on the rate at which the tones are presented. Miller and Heise have found that at a rate of 10 tones per second, a trill as wide as almost three semitones is no longer perceived as a trill, but rather as a melodic fission (or two interleaved, separated tones).⁸²

Bregman experimented with interleaved melodies with no overlapping of their ranges and found that for a compound melody to be perceived, the rate of presentation

⁸²G. A. Miller and G. A. Heise, “The Trill Threshold,” *Journal of the Acoustical Society of America*, 22 (1950): 637-638.

may be from 1.4 tones per second to 20 tones per second.⁸³ At more than 20 tones per second, individual tones are not perceived, while at a rate of less than 1.4 tones per second, attention may shift from one tone to another as they are presented, and the two interleaved melody will form a single melodic line.

Bregman's melodies were constructed with a small pitch range. Schouten has found that it was possible to direct the attention to perceive either temporal coherence or fission within a relatively wide range of presentation rates.⁸⁴ When attempting to perceive temporal coherence, the maximum pitch interval was a M3 at a rate of twelve tones per second, but it increased to a m9 as the presentation rate decreased to four tones per second. Decreasing the tempo had no effect on the perception of a compound melody, though.

Timbre Deutsch maintains that sequences with different timbres are difficult to perceive as one figure.⁸⁵ This assertion is based on experiments with temporal coherence held by Warren and Obusek,⁸⁶ and Van Noorden.⁸⁷ Warren and Obusek ex-

⁸³A. S. Bregman, "Primary Auditory Stream Segregation and the Perception of Tunes," Unpublished manuscript, Department of Psychology, McGill University, Montreal, 1971, cited in: W. J. Dowling, "Perception of Interleaved Melodies," *Cognitive Psychology*, 5 (1973): 322-337.

⁸⁴J. F. Schouten, "On the Perception of Sound and Speech; Subjective Time Analysis," 4th International Congress on Acoustics (Copenhagen Congress Report II, (1962): 201-203), cited in: D. Deutsch, "Memory and Attention in Music," in *Music and the Brain*, eds. M. Critchley and R. Henson (London: Heinemann, 1977).

⁸⁵Deutsch, "Memory and Attention in Music."

⁸⁶R. M. Warren and C. J. Obusek, "Identification of Temporal Order within Auditory Sequences," *Perception and Psychophysics*, 12 (1972): 86-90.

⁸⁷L. A. P. S. Van Noorden, "Temporal Coherence in the Perception of Tone Sequences" (Unpublished doctoral dissertation, 1975), cited in: D. Deutsch, "Memory and Attention in Music," in

perimented with repetitive sequences of sounds such as hisses and buzzes, and found that it was difficult to identify the order in which the sounds were presented. Van Noorden's experiments involved the spectral components of the tones. Using tones with missing fundamentals, he concluded that temporal coherence is achieved with tones that shared spectral components, but not necessarily the fundamental pitch.

The role of timbre in streaming of melodic lines will be discussed in more details in Chapter 5.

Spatial Location Deutsch held a series of experiments with spatial locations and concluded that when presented with two stimuli simultaneously, emanating from two different locations, subjects attend to one stimulus and reject the others.⁸⁸ She also found a tendency of right handed subjects to localize the high notes at their dominant ear. Thus, the stimulus presented in Figure 4.10a was perceived, most of the times, as illustrated in Figure 4.10b. Deutsch's hypothesized that this phenomenon may be due to the existence of two separate mechanisms, one of which determines the pitch, and the other one determines the location.

Loudness Dowling and Hollombe found that when a familiar melody and an interleaved random background were at different levels of loudness, it was easier to

Music and the Brain, eds. M. Critchley and R. Henson (London: Heinemann, 1977).

⁸⁸D. Deutsch, "An Auditory Illusion," *Journal of the Acoustical Society of America*, 55 (1974): 518-519; D. Deutsch, "Musical Illusions," *Scientific American*, 233, no. 4 (1975): 92-104; Deutsch and Roll, "Separate 'What' and 'Where' Decision Mechanisms."

Figure 4.10: Spatial location experiment

attend to the familiar melody, although its pitch range overlapped with that of the background.⁸⁹ It was very difficult to attend to the unfamiliar background.

Familiarity While familiarity with the melodic line may facilitate attending to it, it seems to have a distracting effect when the intent is to direct the attention to the unfamiliar melody. Dowling experimented with short melodies interleaved with a background with which some of the subjects became familiar with prior to the testing.⁹⁰ He found that not only familiarity with the background did not facilitate the recognition task, the group that was not familiar with the background performed best on this task.

⁸⁹W. J. Dowling and Ava W. Hollombe, "The Perception of Melodies Distorted by Splitting into Several Octaves: Effects of Increasing Proximity and Melodic Contour," *Perception and Psychophysics*, 21, no. 1 (1977): 60-64.

⁹⁰Dowling, "Perception of Interleaved Melodies," experiment 4.

4.2.2 Attending Musical Simultaneities

The various ways of channelling attention discussed above, have demonstrated either attending to one figure while rejecting the other, or creating a figure by selecting parts of the stimuli, following various perceptual principles. This does not explain, though, the perception of musical styles such as polyphony and melody/harmony. The melody/harmony style is simpler in the sense that one figure emerges, and the rest becomes the unattended ground. In the case of polyphony, there is at least an illusion of attending to more than one melodic line simultaneously. As opposed to simultaneous verbal messages, simultaneous melodies normally enhance each other and create harmony. In this aspect, these are not two simultaneous messages, but one.

Gregory holds that a true simultaneous perception of two melodies is possible, in the particular case of polyphony: "The two parts probably need to have sufficient factors in common to be perceived as musically related, but at least one musically distinguishing factor, such as pitch level or timbre."⁹¹ Indeed, the results of Gregory's experiment showed that simultaneous melodies were better recognized when they were related in key, shared pitch range and tempo.

It is still possible that these results were achieved without attending to both melodies simultaneously, but by frequent shifts between the lines. Familiarity with

⁹¹A. H. Gregory, "Listening to Polyphonic Music," 169.

one of the lines, or both, make it possible to recognize them with a quick “auditory glance.” Familiarity with the style may be enough, in this sense, since this short “glance” is actually a mere confirmation that expectations are met.

Chapter 5

The Influence of Timbre on Pitch Perception

So far, this thesis has discussed components of music which are parts of its inner syntax; pitch combinations are transformed and manipulated over time in a way that enables us to derive syntactical rules and to perceive the unfolding of a musical piece as a well organized whole, or a “well formed formula.”

Other components of music to be discussed are timbre and dynamics. There is not much to say about dynamics except for the fact that it sometimes interacts with the perception of pitch (a higher tone may be perceived as a louder tone and vice versa). As for timbre, although it serves an important role in organizing and articulating music, it works mostly on the level of complete phrases or larger sections,

and it is not subject to such systematic transformations as are pitch and rhythm.

The reason for that might be the same reason given by Singh and Hirsh, referring to the difficulties in conducting controlled experiments with timbre:

The multidimensionality of timbre does not lend itself easily to quantifiable comparisons. While pitch can be differentiated along ordinal scales (from low to high) and acoustic ratio scales (defining musical intervals like the octave, fifth, etc.), timbre differences are more usually described by nominal, semantic labels characterizing sound sources or qualities.¹

Indeed, the correlation of acoustical aspects to the perception of timbre is not as direct as the pitch–frequency correlation. The differences in the roles of the music components started changing in the Romantic era together with the expanding of the musical syntax in other directions (such as high chromaticism). In the twentieth century, attempts have been made to treat timbre in ways similar to the treatment of pitch, i.e., finding an hierarchical organization,² or using transpositions and inversions.³ Such transformations became possible due to the recent advanced research of timbre, its acoustical components and their perceptual correlates.

The relevance of timbre to this thesis is in the way it interacts, influences, and even interferes with the perception of pitch. The problem that timbre presents when dealing with ear training is not how to identify the timbre, but rather, how to identify familiar musical structures performed with different timbres. Most of the research

¹Punita G. Singh and Ira J. Hirsh, "Influence of Spectral Locus and Fo Changes on the Pitch and Timbre of Complex Tones," *Journal of the Acoustical Society of America*, 92 (1992): 2651.

²F. Lerdahl, "Timbral Hierarchies," *Contemporary Music Review*, 2 (1987): 135-160.

³W. Slawson, *Sound Color* (Berkeley, California: University of California Press, 1985).

concerning timbre deals with the way we identify different timbres. This discussion will be limited to experiments investigating interactions of pitch and timbre. To my knowledge, there is very little research on the influence of timbre on the perception of musical structures, probably because this is a very complex issue (both on the part of timbre and on the part of musical structures), and therefore it is almost impossible to conduct a controlled experiment. On the other hand, it is quite common to see students in ear training classes who are experienced with harmonic dictations performed on a piano, experiencing disorientation when presented with a simple harmonic phrase played by a string quartet.

5.1 Physical Attributes of Sound Influencing Timbre

The most important factor in creating a specific timbre — the spectrum of a musical sound, has a more or less fixed scheme of partials, appearing at various amplitudes. When the partials are a whole-number multiplication of the fundamental, they are harmonics, and then we perceive a specific pitch, even when the fundamental is very weak or missing. When all the partials are not harmonics the sound will be perceived as having no pitch; but we are capable of perceiving a pitch from nearly harmonic partials, or from a complex sound in which only some of its partials are harmonics. Timbre is influenced by changes in amplitudes (loudness) of each partial (some of which maybe very weak or missing), the synchronization of their rise, and the rate of

their decay.

All these factors are relatively easily measured and controlled, which means they can be analyzed and synthesized for creating electronic sound and for experiments. Older synthesizers used to produce sound based on these factors, but every musician presented with a sound synthesized with this information alone would immediately identify it as an electronic sound, mostly because transient components (components that change over time) were missing. These components, which by now are also analyzed and synthesized, create the characteristic attack, decay and release, but also the relatively steady part of a musical sound is not really steady when played by a human being on a musical instrument. The sound will change in intensity, frequency (i.e. vibrato) and in many other ways.

Other sound attributes are characteristic to specific instruments. The trumpet, for example, has a characteristic peak of the amplitude of the partials which is at a fixed frequency region, regardless of the frequency of the fundamental. An experiment showed that keeping this formant region (the peak in amplitude) at the same frequency resulted in better identification of the trumpet's timbre than keeping a fixed spectrum.⁴ On the other hand, the spectrum of the same instrument may

⁴R. Plomp, *Aspects of Tone Sensation: A Psychophysical Study*, Academic Press Series in Cognition and Perception (New York: Academic Press, 1976).

change in different frequencies⁵ or loudness levels.⁶

5.2 Interaction of Timbre and Pitch

As Grau and Kemler Nelson noted, many of the studies of the interaction of timbre, pitch and also loudness were held as part of an ongoing psychological debate of whether we perceive multidimensional stimuli *holistically* (as being overall similar, as *Gestalt* psychologists claim) or *analytically* (by their different properties).⁷ The answer probably depends on the nature of the stimulus and its dimensions. Strategies for the research are largely based on some definitions developed by Garner's experiments of visual perception.⁸ Garner suggested two types of dimensional interaction: one is of stimuli composed of separable dimensions, and the other one is of stimuli composed of integral dimensions. The separable dimensions stimuli have dimensional structure which is directly perceived, and selective attention may be paid to individual dimensions. On the other hand, the integral dimensions stimuli are perceived more in terms of their overall similarity, and their analysis is a secondary process. These

⁵J. C. Risset and M. V. Mathews, "Analysis of Musical Instrument Tones," *Physics Today*, 22, no. 2 (1969): 23-30.

⁶M. Clark and P. Milner, "Dependence of Timbre on the Tonal Loudness Produced by Musical Instruments," *Journal of the Audio Engineering Society*, 12 (1964): 28-31.

⁷James W. Grau and Deborah G. Kemler-Nelson, "The Distinction Between Integral and Separable Dimensions: Evidence for the Integrality of Pitch and Loudness," *Journal of Experimental Psychology: General*, 117 (1988): 347-370.

⁸W. R. Garner, *The Processing of Information and Structure* (Potomac, Maryland: Erlbaum, 1974).

distinctions have implications on perceptual processes such as selective attention, appreciation of similarity, and holistic versus analytic processing. Later studies tried to apply Garner's findings to auditory stimuli and to investigate the validity of the term "integral dimensions." One of the points of criticism of this term was that "all examples of integral dimensions may simply reflect cases in which the experimenter's description of the stimulus does not correspond to the dimensions that the subject perceives."⁹

The interrelations of pitch and timbre go both ways, meaning that timbre can distort the perception of pitch, and changes in pitch may be perceived as changes in timbre. Many experiments explored various aspects of timbre and came to different conclusions about the independence of these two attributes.

Some experiments were held with absolute pitch possessors to investigate whether and to what extent timbre carries pitch information. The simplicity of both spectrum and transient aspects of sine tones were shown to create difficulties with pitch identification. Lockhead and Byrd held an experiment with four subjects,¹⁰ Balzano incorporated three subjects¹¹ and Miyazaki incorporated 28 subjects.¹² Lockhead and Byrd's results had the most dramatic drop in accuracy — from 99% for piano

⁹Grau and Kemler-Nelson, "Distinction Between Integral and Separable Dimensions," 348.

¹⁰G. R. Lockhead and R. Byrd, "Practically Perfect Pitch," *Journal of the Acoustical Society of America*, 70 (1981), 387-389.

¹¹G. J. Balzano, "Absolute Pitch and Pure Tone Identification," *Journal of the Acoustical Society of America*, 75 (1984): 623-625.

¹²Miyazaki, "Absolute Pitch Identification."

tones to 58% for sine tones. Ignoring octave errors raised the results for sine tones to 69%. Balzano's results for sine tones for the three subjects were 91.2%, 84.3%, 87.5%. Balzano also tried to show a frequency dependence; most of the errors occurred in the lowest octave experimented, and were one half-step too sharp. Miyazaki also showed that absolute pitch possessors performed poorly when presented with sine tones as opposed to piano tones, and that the best performances were in the middle of the musical pitch range. Her conclusion was that timbre carries pitch information.

A different approach to research was the investigation of pitch and brightness. Pitch is said to be determined by the frequency of the fundamental, while brightness is said to be determined by the relative amplitude of the higher partials. Some scholars claimed that these two attributes are independent of each other and have different perceptual correlates.¹³ Other studies, though, were able to demonstrate a clear dependence of pitch and timbre. Crowder measured the response time to a "same/different" pitch question and found that it was faster for sounds of the same timbre than to sounds of different timbres.¹⁴ Melara and Marks concluded that pitch and timbre, and also pitch and loudness, interact.¹⁵ Krumhansl and Iverson showed that the subjects' responses to pitch were influenced by changes in timbre, and their

¹³Plomp, *Aspects of Tone Sensation*; Laurent Demany and Catherine Semal, "Pitch Versus Brightness of Timbre," *Music Perception*, 11 (1993): 1-14.

¹⁴Crowder, "Role of One's Own Voice in Immediate Memory."

¹⁵R. D. Melara and L. E. Marks, "Perceptual Primacy of Dimensions: Support for a Model of Dimensional Interaction," *Journal of Experimental Psychology: Human Perception and Performance*, 16 (1990): 398-414.

responses to timbre were influenced by changes in pitch.¹⁶ They also investigated whether timbre is perceived in relation to other timbres the way pitch is. One of their experiments showed that while other pitches interfered with the recognition of a target pitch, other timbres did not interfere with the recognition of a target timbre. Singh and Hirsh showed that changes in brightness were perceived as changes in pitch when the fundamental was fixed or slightly altered.¹⁷ They also referred to the differences between individual listeners — some tend to follow changes in pitch, while others pay more attention to changes in timbre, or as referred to by Singh and Hirsh — changes in “spectral pitch.” Another interesting study, supporting the dependence of pitch and timbre, was conducted by A. Lynne Beal.¹⁸ Beal presented musicians and nonmusicians with chords played on different instruments. The subjects were presented with a “same/different” question. Nonmusicians did very poorly in identifying chords as the same when played on different instruments, but even the musicians’ performance was less accurate when different instruments were used. The instruments used in this experiment were piano, harpsichord and guitar, all having a fast decay. Incorporating more sustaining instruments is likely to impact these results.

¹⁶C. L. Krumhansl and Paul Iverson, “Perceptual Interactions Between Musical Pitch and Timbre,” *Journal of Experimental Psychology: Human Perception and Performance*, 18, no. 3 (1992): 739-751.

¹⁷Singh and Hirsh, “Influence of Spectral Locus.”

¹⁸Beal, “Skill of Recognizing Musical Structures.”

5.3 Timbre as an Organizing Factor

As was noted above, timbre takes a secondary role in Common Practice musical syntax, but it does have an organizing role. Gregory conducted one of the most recent studies of the interaction between pitch and timbre and the way they compete as organizing factors.¹⁹ The way this study, and previous studies²⁰ addressed this issue was by studying the way auditory streaming is formed. Gregory's experiments might serve as a good example to the research method of this issue. He presented subjects with pairs of ascending and descending scales played either simultaneously or with interleaved tones, on different instruments. Some of the different stimuli and the hypotheses for the possible segregation that they may produce are presented in Figure 5.1.

In the first stimulus (5.1a), segregation by pitch may produce two half-scales (5.1b); segregation by timbre may produce two crossing scales (5.1c); and segregation by continuity may also produce the same segregation as the one by timbre. In order to distinguish between the last two cases a second kind of stimulus is presented (5.1d). Now only segregation by continuity may produce crossing scales (5.1f), while segregation by timbre is expected to produce two "jumpy" melodies (5.1e).

¹⁹A. H. Gregory, "Timbre and Auditory Streaming," *Music Perception*, 12, no. 2 (winter 1994): 161-174.

²⁰W. M. Hartmann and D. Johnson, "Stream Segregation and Peripheral Channeling," *Music Perception*, 9 (1991): 155-184; Bregman, *Auditory Scene Analysis* (Cambridge, Massachusetts: MIT Press, 1990).

a. Stimulus 1: different instruments, simultaneous crossing scales

b. Segregation by pitch

c. Segregation by timbre and continuity

d. Stimulus 2: different instruments, simultaneous scales, with the scale notes alternating between the instruments

e. Segregation by timbre

f. Segregation by continuity

♩, ♪ denote different perceptual streams

Source: Gregory (1994), p. 164.

Figure 5.1: Possible segregations of interleaved melodic lines

Bregman suggested that streaming is formed by identification of different sound sources (by clues such as pitch, timbre simultaneity and continuity),²¹ but Hartmann and Johnson found that some dimensions of timbre are more important for the segregation and streaming than others.²² On the other hand, they found that other identifiers of sound source, such as envelope differences, did not produce stream segregation.

Gregory's study compares segregation by timbre to segregation by pitch using tonal stimuli (major scales) and thus, the segregation by pitch is very likely to rely on the existence of a syntax which the subjects are familiar with. One reason often mentioned for timbre taking a secondary role in organizing the music material is a lack of timbral syntax. Some twentieth century composers and music theorists took this challenge to create timbral syntax. The composers tried to manipulate timbre, hoping that with the use of the new language, an inner coherent syntax would be revealed. In 1909 Schoenberg composed his orchestral piece Op. 16 no. 3, known as "Summer Morning by a Lake," but also titled "Changing Chord" and "Colors."²³ This piece explores the possibilities of "timbral melodies" with its opening three measures which are composed of one sustained chord with changes of timbre. In 1911, Schoenberg stated:

²¹Bregman, *Auditory Scene Analysis*.

²²Hartmann and Johnson, "Stream Segregation."

²³Arnold Schoenberg, *Summer Morning by a Lake* (Five Orchestral Pieces Op. 16 No. 1, New York: C. F. Peters, 1973).

Now, if it is possible to create patterns out of tone colors that are differentiated according to pitch . . . then it must also be possible to make such progressions out of the tone colors of the other dimension, out of that which we call simple "tone color," progressions whose relations with one another work with a kind of logic entirely equivalent to that logic which satisfies us in the melody of pitches.²⁴

Later, in the 1950's, other composers of total serialism incorporated timbre as one of the musical dimensions to be ordered and manipulated by rows. Lerdahl criticized this attempt as being arbitrary, not based on the properties of timbre and therefore not readily apprehensible by the listener.²⁵

Theorists went the other way, trying to create speculative theories of timbre and hoping that composers would use these theories to create compositions that make sense. Lerdahl introduced such a theory, based on his former work with Ray Jackendoff.²⁶ He regards a hierarchical organization as superior to other types of organizations in facilitating learning and memory; timbre, though, is not hierarchical in nature. Lerdahl created a "stability scale" of consonance and dissonance for some dimensions of timbre; the next step was looking for possible operations on timbre, such as transposition, inversion, prolongation and resolution. "Intervals" along each of the "stability scales" were determined on perceptual rather than acoustical basis, and Lerdahl suggested ways of operating on one or more dimension at the same time. His musical explorations of operating on timbre use a computer program and started

²⁴Arnold Schoenberg, *Harmonielehre* (Vienna: Universal, 1911), 421.

²⁵Lerdahl, "Timbral Hierarchies."

²⁶Lerdahl and Jackendoff, *Generative Theory of Tonal Music*.

with operating on vowels. In the conclusion to his work, he admitted: “the sphere of exploration has been limited in order to isolate the phenomenon . . . there is no pretension to aesthetic merit . . .”²⁷ and later: “There is a conflict between the fluidity needed for sonority and the fixity needed for syntax.”²⁸ He suggested that timbral hierarchies would not be used independently, but for the reinforcement (or blurring) of pitch hierarchies. Another important work in the realm of timbral syntax is that of W. Slawson.²⁹ This theory is based on vowels and restricted to non-temporal dimensions (eliminating factors such as attack and decay).

Whether in theory or in composition, timbre is being explored in new directions in the twentieth century. Much of this exploration is made possible by the electronic synthesis of sound with the possibilities to control specific dimensions of sound. For the purpose of this thesis, though, it remains a fact that timbre has a secondary role in Common Practice musical syntax. Yet, since it interacts with pitch and with our perception of pitch, changes of timbre should be incorporated in the process of ear training.

²⁷Lerdahl, “Timbral Hierarchies,” 155.

²⁸Lerdahl, “Timbral Hierarchies,” 156.

²⁹Slawson, *Sound Color*.

Chapter 6

Perception of Tonality

The perceptual meaning of tonality is that the tonic, as perceptually superior, is easily and directly perceived, while other tones are perceived in relation to the tonic. Unlike other perceptual prototypes, the tonic is context-dependent; no specific pitch is perceptually superior, but a specific context may create a hierarchy in which the tonic is superior.

There are two main approaches to the study of the perception of tonality; the theory of tonal hierarchy, developed by Carol Krumhansl,¹ and the theory of intervallic rivalry, developed by David Butler, in collaboration with Helen Brown and others.²

¹Krumhansl, *Cognitive Foundations of Musical Pitch*.

²Butler, "Describing the Perception of Tonality;" Butler and Brown, "Describing the Mental Representation of Tonality in Music."

6.1 Tonal–Hierarchy Theory

The theory of tonal hierarchy attempts to rate the stability of different tones within a context of a specific key. This hierarchy of stability is demonstrated in the results of quite a few experiments, such as those held by Krumhansl and Shepard,³ Jordan,⁴ and Speer and Adams,⁵ in which subjects were asked to rate the twelve tones (and sometimes also microtones) within a specific tonal context, as to the degree of their stability. In all the experiments, the overall results with some variations between the different studies, showed that for subjects with at least some musical training, the tonic got the highest rating, followed by other tones of the tonic triad, the other diatonic tones, and last by the nondiatonic tones. Jordan was able to show that microtones were rated lower than nondiatonic tones, while in Krumhansl and Shepard's experiment these tones were assimilated to the neighboring chromatic tones.

Krumhansl and Shepard's experiments showed differences between three groups of subjects according to the degree of their musical expertise. The highly trained musicians showed better octave generalization (meaning low and high tonics got the

³C. L. Krumhansl and R. N. Shepard, "Quantification of the Hierarchy of Tonal Functions within a Diatonic Context," *Journal of Experimental Psychology: Human Perception and Performance*, 5 (1979): 579-594.

⁴D. S. Jordan, "Influence of the Diatonic Tonal Hierarchy at Microtonal Intervals," *Perception and Psychophysics*, 41 (1987): 482-488.

⁵J. R. Speer and W. E. Adams, "The Effects of Musical Training Upon the Development of the Perception of Musical Pitch" (Paper presented at the meetings of the Society for Research in Child Development, 1985), cited in: C. L. Krumhansl, *Cognitive Foundations of Musical Pitch* (New York: Oxford University Press, 1990).

same rating); subjects with less musical training preferred the low tonic, and their preference for diatonic tones was lower. The pattern for subjects with no musical training showed that proximity to the low tonic had an overall stronger effect than scale membership.

Speer and Adam's study showed a developmental sequence for groups of children with no musical training; the preference of diatonic over nondiatonic tones appeared earlier than the hierarchy of the diatonic tones. Musically trained children in all grades followed the hierarchy mentioned above (tonic, tonic triad, other diatonic tones, nondiatonic tones).

A later experiment by Krumhansl and Kessler was designed to expand the variety of tonal contexts, both in major and minor modes, and to minimize the effect of pitch height.⁶ In a C major mode context, pitch classes were rated for stability in the following hierarchy: C, G, E, F, A, D, B, F#, Ab, Bb, Eb, C#. In the minor mode context, the third was rated higher than the fifth, probably due to its being the tonic of the relative major.

6.1.1 Psychological Distances

The perceptual significance of the rate of stability is presented by Krumhansl in the form of three perceptual principles that apply to tones that are rated higher in the

⁶C. L. Krumhansl and E. J. Kessler, "Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys," *Psychological Review*, 89 (1982): 334-368.

tonal stability hierarchy.⁷ These principles, supported by empirical evidence, are presented using a, b to denote tones, and d to denote psychological distance between the tones.

The first principle is *Contextual identity*: “ $d(a, a)$ decreases as the stability of a increases,”⁸ which means that the sense of identity is stronger for two instances of the same stable tone than for instances of less stable tones. In other words, the psychological distance between two instances of a tone gets smaller as its stability increases.

The second principle is *Contextual distance*: “ $\frac{1}{2}[d(a, b) + d(b, a)]$ decreases as the stability of a increases and as the stability of b increases.”⁹ (The average distance is taken, since the psychological distance a, b does not necessarily equal the psychological distance b, a .) This principle relates to the psychological distance between two distinct tones, which decreases as the stability of the tones increases.

The third principle is *Contextual asymmetry*: “ $[d(a, b) - d(b, a)]$ will increase as the tonal stability of tone a increases and the tonal stability of tone b decreases.”¹⁰ This principle governs the effect of the temporal order of the tones, a being the first tone, and b the second. It means that as a becomes more stable and b becomes less stable, their temporal placement will have a greater effect on the distance between

⁷Krumhansl, *Cognitive Foundations of Musical Pitch*, 143–150.

⁸Krumhansl, *Cognitive Foundations of Musical Pitch*, 143.

⁹Krumhansl, *Cognitive Foundations of Musical Pitch*, 148.

¹⁰Krumhansl, *Cognitive Foundations of Musical Pitch*, 150.

them. A sequence of an unstable tone followed by a stable tone spans over a smaller psychological distance than a sequence of a stable tone followed by an unstable tone.

The empirical evidence for the first principle may be derived from Krumhansl's study in 1979.¹¹ In the third experiment of this study, probe tones were heard at the beginning and end of an intervening melody: some melodies were in C major and some were atonal. The probe tones best and most confidently identified as "same" were the diatonic C major tones with the C major intervening melody. This confirms the assumption that tonally stable tones are more stable in our memory.

The evidence for the second principle comes from experiments that show that the same interval is judged differently according to its tonal position. Two stable tones will be judged as more closely related to each other than two less stable notes, in the same interval. Such an experiment is described in Krumhansl's *Cognitive Foundations of Musical Pitch*.¹² Subjects were presented with sequences of two tones within an established tonal context; these sequences were all possible combinations of the twelve tones (total of 132 combinations). A group of trained musicians was asked to rate the combinations as to how well the second tone followed the first. The high rating of pairs of stable tones implies that stable tones are perceived as more related than unstable tones. In a different experiment, Cuddy, Cohen and Miller

¹¹C. L. Krumhansl, "The Psychological Representation of Musical Pitch in a Tonal Context," *Cognitive Psychology*, 11 (1979): 346-374.

¹²Krumhansl, *Cognitive Foundations of Musical Pitch*, 123-4.

presented subjects with a melody and two transpositions; one exact transposition, and one with an alteration of one tone, to either diatonic or chromatic tone.¹³ The chromatic changes were detected more than the diatonic ones, again, implying that diatonic tones, or highly stable tones, are at a smaller psychological distance.

Krumhansl's experiment with pairs of tones, described above, also provides evidence for the third principle. In that experiment, the order of the two tones affected the subjects' judgment, with preference to ending on a stable tone. Moreover, the more distant the two tones were in the tonal hierarchy, the more effect the temporal order had. Krumhansl's 1979 experiment, also described above, provided additional evidence. The identification of nondiatonic tones was rated lower in confidence than that of the diatonic tones, which implies that the less stable tones are more easily confused with the stable tones than vice versa.

Krumhansl's conclusions were:

the findings . . . fit nicely with the *Gestalt* notion of singularity of elements within a whole. Individual tones within tonal contexts are perceived as more or less stable, normative, or consistent with the tonality. This has the consequence that certain elements are better remembered, are less often confused with other — less stable — elements, and are preferred in final positions.¹⁴

¹³Cuddy, Cohen and Miller, "Melody Recognition."

¹⁴Krumhansl, *Cognitive Foundations of Musical Pitch*, 150-51.

Table 6.1: Correlations of C major to other major keys (The distance is shown by the number of positions from C in the circle of fifths, or number of differences in the key signature from C.)

Key	Correlation	Distance
G/F	0.591	1
D/B $\flat$	0.040	2
A/E $\flat$	-0.105	3
E/A $\flat$	-0.185	4
B/D $\flat$	-0.500	5
F $\sharp$	-0.683	6

Source: Krumhansl (1990) p. 39.

6.1.2 Key Relations

The next step in the tonal hierarchy theory is the investigation of interkey distance. Keys are considered related in as much as their hierarchies are similar. Krumhansl compared the hierarchies of different keys by computing the correlations between the profiles of hierarchies of each pair of major and minor keys, with the number “1” denoting full correlation, and the number “-1” denoting negative correlation.¹⁵ While the distance of C major from other major keys showed the relations created by the circle of fifths (Table 6.1), the correlation of C minor to other minor keys showed some interference of parallel and relative relations (Table 6.2).

While the effect of the circle of fifths is evident by the equal values given to keys that are equally distant from C minor, keys that are at a distance of three and four

¹⁵Krumhansl, *Cognitive Foundations of Musical Pitch*, 39.

Table 6.2: Correlations of C minor to other minor keys

Key	Correlation	Distance
F/G	0.339	1
E $\flat$ /A	0.055	3
A $\flat$ /E	-0.003	4
B $\flat$ /D	-0.160	2
C $\sharp$ /B	-0.394	5
F $\sharp$	-0.673	6

Source: Krumhansl (1990) p. 39.

positions precede keys that are at a distance of two positions. This can be explained by relations of relative and parallel minors. Thus, the correlation between C minor and E $\flat$ minor is produced by the relations of C minor to E $\flat$ major, and E $\flat$ major to E $\flat$ minor; the correlation between C minor and A minor goes through C major; the correlation between C minor and A $\flat$ minor goes through F minor and F major, and so forth.

Correlating major and minor keys shows a stronger effect of relative and parallel relations, and keys of the same distance on the circle of fifths do not have equal values any more (Table 6.3).

These distances between keys are asymmetrical. Thompson found that modulations up the circle of fifths are perceived as less distant than modulations down the circle.¹⁶

¹⁶W. F. Thompson, "Judgments of Key Change in Bach Chorale Excerpts: An Investigation of the Sensitivity to Keys, Chords, and Voicing" (Unpublished doctoral dissertation, Department of

Table 6.3: Correlations of C major to other minor keys

Key	Correlation	Distance
A	0.651	0
E	0.536	1
C	0.511	3
G	0.241	2
D	0.237	1
F	0.215	4
B	-0.158	2
C#	-0.298	4
F#	-0.369	3
Bb	-0.402	5
G#	-0.508	5
Eb	-0.654	6

Source: Krumhansl (1990) p. 39.

6.1.3 Criticism

Butler criticizes both the theory of tonal hierarchies and the procedures applied to obtain the empirical data supporting it.¹⁷ Krumhansl's main method, the probe tone method, involves creating a tonal context mostly by playing a scale, and then sounding a probe tone to be rated for its stability. Butler compares the validity of the use of a scale to investigate tonality to the use of the alphabet to investigate deep lingual structures. Furthermore, with the diatonic scale heard just before the probe tone, short-term memory may interfere with the judgment and thus diatonic

Psychology, Queen's University. Kingston, Ontario, Canada, 1986), cited in: C. L. Krumhansl, *Cognitive Foundations of Musical Pitch* (New York: Oxford University Press, 1990).

¹⁷Butler, "Describing the Perception of Tonality;" David Butler, "Response to Carol Krumhansl," *Music Perception*, 7 (1990): 325-338.

tones would be preferred because they would sound familiar, and the tonic would be preferred because it was the first and last tone of the scale. Butler and Brown present a wider view of tonality as an idiom, dialect or even “an entire grammatical system.”¹⁸

As for the theory itself, Butler claims that it is static, it fails to account for the effect of the temporal order of the tones, and it does not explain how the tonality of a specific musical piece is perceived. The first point, the theory being static, refers to the fact that all tones, diatonic and chromatic, are given a specific rate of stability, which does not allow for a tone to have different meanings within the same key. While this possibility is quite obvious regarding chromatic tones, it cannot be ignored even regarding diatonic tones which may carry different meanings on different harmonic levels.

The only temporal consideration present in Krumhansl’s theory seems to be the influence of the order of pairs of stable and unstable tones (see Krumhansl’s third principle, *Contextual asymmetry*). Butler and Brown regard temporal order as much more significant to tonality. Brown’s experiment of 1988 shows that by changing the temporal order in pitch sets, they may suggest a variety of tonalities.¹⁹ Figure 6.1 demonstrates how temporal re-ordering of a sequence may change the tonality of

¹⁸Butler and Brown, “Describing the Mental Representation of Tonality in Music,” 197.

¹⁹Helen Brown and David Butler, “Diatonic Trichords as Minimal Tonal Cue-cells,” *In Theory Only*, 5 (1981): 39-55.

Figure 6.1: Effects of temporal order on the perception of tonality

a sequence to suggest a diatonic or a chromatic set. Extracting the pitches of an excerpt from Schubert, *Sonata D. 664*, third movement, Brown and Butler have demonstrated how an unambiguous diatonic set can be temporally manipulated to suggest chromaticism. While in the first temporal setting (Figure 6.1b) 95.2% of the listeners identified D as a tonic, in the second setting (Figure 6.1c) only 40.5% identified D as a tonic, while 45.2% identified G as a tonic (probably interpreting C# as a leading tone to the dominant, D, which immediately follows it).

Regarding meeting the goal of key-identification, Brown Butler and Jones claim that the tonal hierarchy theory operates *within* a specific key:

It emphasizes the stability of certain prominent pitches — tonic, mediant, dom-

inant — that happen to be outlined by those interval classes more common in the diatonic set. The strength of the tonal hierarchy model lies in descriptions of ways these common intervals reinforce key. By reinforcing key, we mean: Given a fixed cognitive anchor pitch, the presentation of more-common intervals of the diatonic set (major seconds, minor and major thirds, and perfect fourths) supports tonal functions within the tonal framework of the anchor pitch.²⁰

6.2 Intervallic–Rivalry Theory

The intervallic rivalry theory, on the other hand, sets key discovery as its main goal. In a more recent study held by Brown, Butler and Jones, the theory of intervallic rivalry was presented as based on three hypotheses: (1) *Primacy*: the first tone heard is assumed to be the tonic, until a better candidate is heard; (2) *Rare intervals*: intervals that occur only once or twice in the diatonic set are better clues for key discovery than other intervals; (3) *Temporal order*: rare intervals are most efficient as clues for key discovery when they are arranged in a temporal order implying common harmonic motions.²¹ Brown Butler and Jones regard the interaction between the second and the third hypotheses as the core of the theory of intervallic rivalry.

The rare intervals hypotheses is based on Richmond Browne's presumption that since the tritone and the minor second are the rarest intervals of the major scale, they

²⁰Helen Brown, David Butler and Mari Riess Jones, "Musical and Temporal Influences on Key Discovery," *Music Perception*, 11 (1994), 377.

²¹Brown, Butler and Jones, "Musical and Temporal Influences on Key Discovery," 372.

Source: Brown, Butler and Jones (1994), p. 375.

Figure 6.2: Position finding facilitated by the Tritone

may serve in identifying a tonic.²² These intervals facilitate what Browne calls *position finding* and the more unique they are, the less ambiguous their position is.

Brown, Butler and Jones demonstrate how the tritone facilitates *position finding* by looking at the combined set of diatonic tones in C major and in F# major (Figure 6.2): the tritone is represented by the gray keys; the C major set by the gray and white keys; and the F# major set by the gray and black keys. Any combination of the tritone with a third tone (black or white key) will determine whether the sequence is in C major or in F# major. This theory is supported by empirical findings such as Brown and Butler's 1981 experiment in which three diatonic tone sequences containing a tritone were shown to yield significantly stronger agreement on the tonic than diatonic sequences with no tritone.²³

Since in the minor mode there are two tritones, Butler regards the augmented second (Figure 6.3a) and the diminished fourth (Figure 6.3b) as the tonic-identifying

²²R. Browne, "Tonal Implications of the Diatonic Set," *In Theory Only*, 5 (1981): 3-21.

²³Brown and Butler, "Diatonic trichords."

Source: Butler (1992), p. 166.

Figure 6.3: Tonic-identifying rare intervals

rare intervals, and his 1992 experiment supports this assumption.²⁴ The short melodic patterns containing an augmented second were sufficient for a high level of agreement on a tonic.

Brown and Butler's 1981 experiment also showed that the sequence F-B, implying subdominant to dominant, yielded more agreement on the tonic than the sequence B-F, implying dominant to subdominant. This effect of the temporal order of the tones on the judgment led Brown and Butler to the conclusion that the sequences were carrying harmonic meaning, and that when the implied harmony represent common harmonic motion, the sequences facilitate key discovery.

In 1982 Cuddy and Badertscher conducted an experiment, using Krumhansl's probe tone technique, which yielded some surprising results.²⁵ The context of C major triad produced a profile resembling Krumhansl's profiles for major keys, only with F being rated higher than E, which implies that the triad was interpreted as a dominant.

²⁴David Butler, "On Pitch-set Properties and Perceptual Attributes of the Minor Mode," in *Cognitive bases of musical communication*, eds. R. Jones and S. Holleran (Washington, DC: American Psychological Association, 1992).

²⁵L. L. Cuddy and B. Badertscher, "Recovery of the Tonal Hierarchy: Some Comparisons Across Age and Levels of Musical Experience," *Perception and Psychophysics*, 41 (1987): 609-620.

The context of B diminished triad, on the other hand, did not produce a profile of any key, which led Cuddy and Badertscher to the conclusion that the tritone was not effective in key discovery.

Brown, Butler and Jones replicated this experiment using the same stimuli, and then altered the temporal order of the tones. They noted that the difference between the two theories implied different predictions as to the results of such experiments. Since the tonal hierarchy theory rated the stability of tones within a given key, the highest ratings would be given to tones 1, 3 and 5 of the scale, and the tone profile would be jagged (altering diatonic and chromatic tones). The intervallic rivalry theory, on the other hand, rated the tones that are the best possible tonics and thus the prediction is that the highest ratings would be given to tones 1, 4 and 5 of the scale, and the jagged key profile would not necessarily appear.²⁶ Obtaining such different results may also reflect differences in the instructions given to the subjects, or in the subjects' interpretation of the instructions.

The replication of the experiment yielded similar results to Cuddy and Badertscher's results for the C major triad context. Significant differences, though, were found for the diminished triad context. While Cuddy and Badertscher's profile was quite flat, with no significant peak at any pitch class, Brown Butler and Jones' profile had a significant peak at C, indicating that C got the highest rating as a possible

²⁶Brown, Butler and Jones, "Musical and Temporal Influences on Key Discovery," 390.

tonic (Figure 6.4).

The diminished triad profile kept the same directions as that of their C major triad profile, but was flatter, with a local peak at B (Figure 6.5).

Figure 6.4: Profiles of the diminished triad: Brown, Butler and Jones, and Cuddy and Badertscher, superimposed

When Brown Butler and Jones altered the order of the tones, the results changed. Most striking is the fact that while F was judged as tonic in the original order (Figure 6.6a), C was judged as tonic in the altered pattern (Figure 6.6b).²⁷ Altering the order of the tones in the diminished triad (Figure 6.7) resulted in a much higher rating for C as a tonic, which implies that the temporal order in Cuddy and Badertscher's experiment (Figure 6.7a) made the tonality of the sequence more am-

²⁷Brown, Butler and Jones, "Musical and Temporal Influences on Key Discovery," 382, 392.

Figure 6.5: Profiles of Brown, Butler and Jones' diminished triad and major triad. superimposed

biguous. The relative minor (A) also received high rating (lower than the C, though) but F and G did not.

Temporal aspects of tonality were also investigated by Butler's student, C. R. Strader, in her Ph.D. dissertation.²⁸ She concluded that while temporal order affected the subjects' choice of tonic, temporal duration did not. While meter is not mentioned in Butler's work, it is explicitly excluded in Strader's research. However, it seems that metric position may affect the choice of tonic in the case of the major triad. Changing the order of the tones "pushes" the C to a different metric position (referred to by

²⁸C. R. Strader, *The Effects of Variations of Duration on Musicians' Perceptions of Tonic in Tonally Ambiguous 3 Notes Set* (Ph.D. diss., Ohio State University, 1996).

Figure 6.6: Effect of the temporal order of the major triad tones on the judgment of the preferred tonic

Figure 6.7: Effect of the temporal order of the diminished triad tones on tonal ambiguity

Brown Butler and Jones as a “prominent serial location”²⁹). While in Figure 6.6b the second C is a downbeat on some metric level, in Figure 6.6a the second C is easily interpreted as an anacrusis which is typically a dominant. To further demonstrate how the role of the temporal order may become secondary to the role of meter, one may alter the rhythm of Figure 6.6a to push the second C to a down beat metric position (Figure 6.8).

Another possibility worth experimenting with is to eliminate the temporal order aspect by sounding all the tones simultaneously, in two different rhythms which imply two different metric positions of the second sound (Figure 6.9).

²⁹Brown, Butler and Jones, “Musical and Temporal Influences on Key Discovery,” 393.

Figure 6.8: Effect of metric position on tonality

Figure 6.9: Effect of metric position with the Elimination of temporal order

6.3 Combining the Two Theories

J. Bharucha suggested a model by which unstable tones served to attract attention to their pitch neighborhood, and created a “directional expectation” toward the nearest stable tone.³⁰ This model seems to incorporate both the theory of tonal hierarchy and the theory of intervallic rivalry. The stable tones, or chord tones, act as anchors, or “cognitive reference points.”³¹ Bharucha made an interesting distinction between an event and its mental representation, or a mental object, by which an unexpected event “attracts attention to the extent that its mental representation is inactive before the event occurs,” while the mental representation attracts attention “to the extent that they are activated.”³² For our purpose, the unexpected event is a nonchord (or unsta-

³⁰J. Bharucha, “Melodic Anchoring,” *Music Perception*, 13 (1996), 387.

³¹Bharucha, “Melodic Anchoring,” 383.

³²Bharucha, “Melodic Anchoring,” 386.

ble) tone; the mental object is one of the stable tones of a tonal template. Although Bharucha is not explicit about what it is that activates the mental representation, it is probably the establishment of a key as described by the intervallic rivalry model. The resolution of the nonchord tone, on the other hand, is highly predictable and thus it serves as a pointer to one of the tones on the template — the closest one to the pointer's pitch neighborhood: "The resulting attentional focus on the representation of the highly activated chord tone neighbor while the nonchord tone is still sounding gives rise to the subjective demand or yearning for resolution to the chord tone."³³

Although Butler and Krumhansl conduct an ongoing debate through various publications,³⁴ the two theories are not mutually exclusive, as demonstrated by Bharucha's model, and as expressed in the article by Brown Butler and Jones: "The two models address different aspects of tacit knowledge about tonality."³⁵ While the theory of intervallic rivalry suggests a model for key identification by rare intervals whose temporal order carries a harmonic meaning of dominant to tonic, the theory of tonal hierarchy describes the template of stable tones engaged by the established tonality.

³³Bharucha, "Melodic Anchoring," 386.

³⁴C. L. Krumhansl, "Tonal Hierarchies and Rare Intervals in Music Cognition," *Music Perception*, 7 (1990): 309-324; Butler, "Response to Carol Krumhansl."

³⁵Brown, Butler and Jones, "Musical and Temporal Influences on Key Discovery," 377.

Chapter 7

Perception of Musical Time

7.1 Introduction

Like the previous chapter, this chapter concentrates on the perception of one aspect of music, while bearing in mind that all musical aspects interact in our perception, and any investigation of the perception of one aspect can happen in a controlled experiment but not in real music. But is time only an *aspect* of music? It is involved in every aspect of music; it affects not only meter and rhythm but also aspects such as pitch, timbre, melody and form: “The processes of growth and decline, antecedence and consequence, preparation and resolution, and motion toward and recession from climax, regardless of how they are described and in whatever terms, all depend on

orderings of musical elements and relationships in time.”¹

Time is the space in which music operates. When looking at a work of visual art, we are able to alter our perspective by changing our distance from the piece. We can look at a picture from a distance, get an overall view of the composition, then get closer, examine details in the order of our choice and finally go back for an overview with a greater awareness of the details. As Lester points out, there is a big difference between our perception of visual space and that of the temporal space: “The very notion of a melody, a harmony, and even a tonality would disappear if the pitches or sections of a piece could be heard in any order that the listener chose.”² Furthermore, to simulate the equivalent of a visual overview of the whole piece, we would have had to hear all the pitches simultaneously. A musical “overview” can only happen in retrospect.

While time is the space of all performing arts, perhaps none of these arts structure time to the degree that music does. It was Susanne Langer who had said that music “makes time audible.”³

Some writers go as far as to regard time as *the* meaning of music. Marvin Minsky compared human fascination with the way music structures time to that of a child’s fascination with building blocks: “He plays with blocks and boxes, stacking

¹J. Lester, *The Rhythms of Tonal Music* (Carbondale: Southern Illinois University Press, 1986), 1.

²Lester, *Rhythms of Tonal Music*, 2.

³Susanne Langer, *Feeling and Form* (New York: Scribner’s, 1953), 110.

them and packing them; he lines them up and knocks them down. What is that all about? Clearly, he is learning Space! But how, on earth, does one learn Time? Can one time fit inside another, can two of Them go side by side? In Music, we find out!"⁴

Susanne Langer distinguished two kinds of time: "one-dimensional, infinite succession of moments,"⁵ which will be called here *objective time* or *clock time*, and "the sequence of actual happenings,"⁶ which will be called here *subjective time*. Accordingly, there are two main theories of time perception: (1) Perception of time by an internal clock, or the *pacemaker theory*;⁷ (2) Perception of time as a function of the amount of information processed.⁸ Clarke noted that these two theories are applicable to different kinds of human behavior, and to different structural levels of music.⁹ The internal clock theory is applicable to behaviors such as conducting and performing, both having a motoric component, and to a surface rhythmic level that has relatively short durations. The information-depended theory is applicable to the perception of longer durations.

In terms of musical skills, the internal clock allows us to "keep the tempo."

⁴Marvin Minsky, "Music, Mind and Meaning," in *Music, Mind and Brain: The Neuropsychology of Music*, ed. Manfred Clynes (New York: Plenum Press, 1982), 5.

⁵Langer, *Feeling and Form*, 111.

⁶Langer, *Feeling and Form*, 109.

⁷A. B. Kristofferson, "A Quantal Step Function in Duration Discrimination," *Perception and Psychophysics*, 27 (1980): 300-306.

⁸Paul Fraisse, "Rhythm and Tempo," in *The psychology of music*, ed. D. Deutsch, Academic Press Series in Cognition and Perception (New York: Academic Press, 1982).

⁹E. F. Clarke and C. L. Krumhansl, "Perceiving Musical Time," *Music Perception*, 7 (1990): 213-251.

and is related to our ability to perceive durations. How accurate is our memory for durations? To what extent can we perceive, remember and reproduce accurately a silent time interval? Kramer commented on these types of questions: "In the process of experiencing and then remembering lengths of time, we alter them. It is misleading, however, to think of subjective time alterations as distortions. Subjective duration is *more* relevant to the understanding of music than is duration measured by a clock."¹⁰

Clock time, or objective time, is an important aspect of music, but it is a stylistic, rather than a structural aspect. While a piece played in a "wrong" tempo can still be easily recognized, a piece played with wrong rhythm is very hard to recognize.¹¹

7.2 Definitions

Most of the research concerning the perception of rhythm and meter is directed to isolated aspects of music and is conducted by psychologists, psychoneurologists and psychomusicologists. Since even music theorists do not agree on the terminology of the temporal aspects of music, it is not surprising to find a complete chaos in the use of those terms in research. While in *The New Harvard Dictionary of Music* there is

¹⁰J. D. Kramer, *The Time of Music: New Meanings, New Temporalities, New Listening Strategies* (New York: Schirmer, 1988), 327.

¹¹E. C. Carterette *et al.*, "Rhythmic and Melodic Structures in Perceptual Space," *Journal of the Acoustical Society of America* 72, Supplement 1 (1982): S11.

no definition of *Pulse*, “Beat” is defined as “A metrical pulse.”¹² Such a definition calls for checking the *meter* entry where it reads: “The pattern in which a steady succession of rhythmic pulses is organized.”¹³ Unfortunately, this interchangeable use of pulse and beat, rhythm and meter is common in writings of psychomusicologists as well as music theorists. In the *rhythm* entry, Harold S. Powers explains some of the confusion: “In modern English musical usage, the word rhythm appears on two semantic levels. In the widest sense, it is set beside the terms melody and harmony ... covers all aspects of musical movement as ordered in time ... In the narrower and more specific sense, rhythm shares a lexical field with meter and tempo.”¹⁴

Before continuing this discussion it is necessary to define these terms together with some other related terms. Some definitions from both disciplines (music theory and psychomusicology) will be quoted, and the way each term will be used here will be stated.

Pulse and Beat While the terms *beat* and *pulse* are very often used interchangeably, this discussion may benefit from making a distinction.

Cooper and Meyer’s initial definition of pulse is “one of a series of regularly

¹²David Fuller, “Beat,” *The New Harvard Dictionary of Music* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1986), 85.

¹³Don Michael Randel, “Meter,” *The New Harvard Dictionary of Music* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1986), 489.

¹⁴Harold Powers, “Rhythm,” *The New Harvard Dictionary of Music* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1986), 700.

recurring, precisely equivalent stimuli.”¹⁵ There are two problems with this definition. First, a pulse is not a stimuli or a physical entity, but a perceptual entity. As Cooper and Meyer themselves note later, once a pulse is established, it may continue in the listener’s mind, and they add a better definition: “pulses mark off equal units in the temporal continuum.”¹⁶ This leaves us with the second problem: while pulses are primarily equal, they may be maintained through changes such as rubato, accelerando, etc. Lester formulated the following definition of beats: “*Beats mark off functionally equivalent spans of time.*”¹⁷ Now, a pulse is a musical point: it has no duration (or, it is not a “span of time”) just as a point has no dimensions. It only has a point of attack. The duration of the beat, on the other hand, is equivalent in length to the time interval between one pulse to the next pulse. Thus, Lester’s definition may be adjusted for beats and pulses: *Pulses mark off functionally equivalent intervals of time; beats mark off functionally equivalent spans of time.*

Another feature of the pulse is that it exists on different levels, and that these levels are usually nested, with one level of pulse considered primary (for the considerations by which the primary level is chosen see Section 7.3.2).

¹⁵Grosvenor W. Cooper and Leonard B. Meyer, *The Rhythmic Structure of Music* (Chicago: University of Chicago Press, 1960), 3.

¹⁶Cooper and Meyer, *Rhythmic Structure of Music*, 3.

¹⁷Lester, *Rhythms of Tonal Music*, 46.

Accent David Fuller defined accent as an “Emphasis on one pitch or chord. An accent is dynamic if the pitch or chord is louder than its surroundings, tonic if it is higher in pitch, and agogic if it is of longer duration.”¹⁸ An accent though, may occur on an unpitched sound as well, and indeed, Cooper and Meyer hold a different view about these three different ways in which an accent is supposed to be created:

none of them appears to be an invariable and necessary concomitant of accent. Accents may occur on short notes as well as long, on soft notes as well as loud, on lower notes as well as higher ones ... Accents have to relate to the other sounds yet be different ... An accent, then, is a stimulus (in a series of stimuli) which is *marked for consciousness*.¹⁹

This definition of accent may be supported by a series of experiments held by Fraisse *et al.* (in 1954 and 1956) and summarized by Fraisse.²⁰ Subjects were presented with sequences of isochronous sounds that differed in intensity or pitch. One of the findings of these experiments was that the least frequent element was perceived as accented, whether it was loud or soft.

Meter Of all the terms discussed and defined so far, the definitions of meter are the most blurred ones. This may be accounted for by the fact that meter, like a pulse, is a perceptual entity, but while a pulse may be stored “as is” in the sensory memory (that is, without any further cognitive processes), the perception and storage of meter involves much higher-level cognitive processes.

¹⁸David Fuller, “Accent,” *The New Harvard Dictionary of Music* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1986), 3.

¹⁹Cooper and Meyer, *Rhythmic Structure of Music*, 7.

²⁰Fraisse, “Rhythm and Tempo,” 162-164.

Cooper and Meyer define meter as “the measurement of the number of pulses between more or less regularly recurring accents.”²¹ While it is plausible that the *number* of pulses (or beats) may be perceived by experienced musicians (such as a feeling of “twoness” or “threeness”²²), meter can be perceived regardless of this number, by perceiving the time interval between two accents.

Cook explains meter as a higher level pulse: “In common practice a pulse level is organized by the pulse level above it,” and he defines meter as “an organizational scheme of at least one pulse level constitutes a meter and is often created by an accented pattern.”²³

While the explanation of the nested pulses does not take into account the perceptual difference mentioned above, the definition fails to give a full account of what it is that creates the organization: it is “often created by an accented pattern,” but how is it created when there is no accented pattern to create it?

Using Meyer’s notion of expectations,²⁴ Common Practice meter may be defined as *an expectation for a regularly accented pulse, one level above the primary pulse level.*²⁵

²¹Cooper and Meyer, *Rhythmic Structure of Music*, 4.

²²Brower, “Memory and the Perception of Rhythm,” 28.

²³K. E. Cook, *A New Approach to Musical Meter: Contributions Towards a Theory of Metric Perception in Music* (Ph.D. diss., Michigan State University, 1991), 11.

²⁴L. B. Meyer, *Emotions and Meaning in Music*.

²⁵Non-Common Practice meters would raise questions such as irregular meters or meters having beats of different lengths.

Figure 7.1: Rhythms of equal spaces but different durations

Rhythm Definitions of meter are unclear mostly due to the lack of understanding of its perception. But it seems that definitions of rhythm are unclear because scholars seem to differ widely on what is the phenomenon that they believe the word “rhythm” refers to.

Fraisse quotes Plato, stating that rhythm is “the order in the movement.” and Fraisse adds that “rhythm is the perception of an order.”²⁶ This is a very wide definition that may work also for visual rhythm.

Harold S. Powers also starts from Plato’s notion of movement: “The pattern of movement in time ... A perceivable pattern of temporal space between attacks constitutes a rhythm.”²⁷ It should be noted that since Powers does not include duration in his definition, it should follow that rhythms A and B in Figure 7.1 are identical.

Amos’ definition is not far from that line of thought: “rhythmic patterns are successions of stimuli in time which are perceived as temporal Gestalts.”²⁸

²⁶Fraisse, “Rhythm and Tempo,” 150.

²⁷Powers, “Rhythm,” 700.

²⁸T. E. Amos, *Perceiving Musical Rhythms in Different Sense Modalities*, (Ph.D. diss., University

In both Powers' and Amos' definitions, the element of relationships is missing; Amos refers to "successions," but relationships between the stimuli are necessary for a rhythmic pattern to be perceived.

Cook includes relationships in his definition, maintaining that rhythm is "the duration and temporal relationships of sound and silence in music."²⁹ This definition does not relate to the formation of conceptual units, or patterns, by those relationships.

Another aspect, the accentuation, is attributed to rhythm by Dowling and Harwood: "*Rhythm* refers to a temporally extended pattern of durational and accentual relationships."³⁰

Cooper and Meyer also incorporate accents into their definition, which represents a completely different approach: "The way in which one or more unaccented beats are grouped in relation to an accented one."³¹ This definition relates to the prosodic feet that Cooper and Meyer used in their rhythmic analysis. Lester criticized this approach, claiming that while in prosody syllables are either accented or unaccented, in music there is a full range of accentuations. Further criticism on the use of prosodic feet in music will be discussed in Section 7.3.5.

of California, San Diego, 1987), 7.

²⁹Cook, *New Approach to Musical Meter*, 16.

³⁰W. J. Dowling and Dane L. Harwood, *Music Cognition*, Academic Press Series in Cognition and Perception (New York: Academic Press, 1986), 185.

³¹Cooper and Meyer, *Rhythmic Structure of Music*, 6.

Figure 7.2: Rhythmic patterns created by various means of accentuations

It seems natural to include accents in the definition of rhythm when the accent is agogic; but should other kinds of accents be included? If rhythm is to be regarded as a strictly temporal phenomenon, then tonic and dynamic accents should be excluded. On the other hand, as Figure 7.2 may demonstrate, any means of accentuation — intensity, pitch or even timbre, may create a rhythmic pattern. It is this kind of rhythmic pattern, using notes of equal durations, that is most frequently investigated in the realm of psychomusicology.

It is true that rhythm interacts with many other musical aspects, but since it is necessary for this discussion to isolate the temporal aspects, accents should be left out of the definition of rhythm. Yet, experiments incorporating isochronous rhythms will be cited, since they are still valuable for the insights they provide on the perception of rhythm.

As for the question of durations versus time intervals, since the literature of psychomusicological experiments examined the perception of rhythmic groups as time intervals (although the term “duration” was often used, misleadingly), it seems sensible to regard here durations as aspects of articulations and simultaneity rather than rhythm.

A rhythmic group for our purposes then, is *a pattern of relationships of time intervals*.

7.3 Perception of Temporal Aspects of Music

7.3.1 Perception of Durations

Though duration usually refers to the actual duration of a specific pitch (or silence), the time interval between two points of attack may also be perceived as duration. Cook used the example in Figure 7.3 to stress the importance of duration in its first sense.³² Yet, this example demonstrates that although the durations of the tones of the two melodies are different, there is more similarity than difference between the two. The changes in durations maybe perceived as changes in articulation rather than changes in the tonal meaning of the melodies. It is only when simultaneities are used that this sense of duration may assume structural importance and may impact the

³²Cook, *New Approach to Musical Meter*, 15.

Source: Cook (1991), p. 15.

Figure 7.3: Durations versus points of attack

tonal meaning of the piece.

Research shows that human perception of clock time is very poor. Tasks such as reproducing silent time intervals, or judging the similarity of two silent time intervals produced poor results. It should be noted, though, that the performance was best for the durations around 0.6 seconds (M.M.=100).³³ A much easier task, with more successful results, was synchronizing with a steady pulse: again, the best results were achieved at around M.M.=100.³⁴

Such poor discrimination for clock time implies that although it is a shaping force in music (since the choice of tempo *is* meaningful), it cannot be a component of the musical syntax. “Pace-maker” type musical tasks such as conducting or keeping tempo, are more similar to the synchronizing in the experiment mentioned above.

More relevant and complex is the theory of time perception as a function of

³³H. Woodrow, “Time Perception,” in *Handbook of Experimental Psychology*, ed. S. S. Stevens (New York: Wiley, 1951).

³⁴Dirk-Jan Povel, “Internal Representations of Simple Temporal Patterns,” *Journal of Experimental Psychology: Human Perception and Performance*, 7 (1981): 3-18.

the processed information, only one must keep in mind that this theory investigates our experience of subjective time, and not our perception of clock time. There are several aspects that influence our experience of duration, and in 1890 William James found that the influence of the same aspect may be reversed when we switch from the experience of time in passing, to the experience of retrospective time.³⁵

The first aspect is the number of stimuli in a specific time interval. A time interval filled with more stimuli seems to be longer.³⁶ A second aspect is the complexity (or unpredictability) of the stimuli which was also found to make the time interval seem longer.³⁷ These findings confirm James' earlier finding, but James also noted that in passing, a time interval filled with varied stimuli seems shorter than an empty time interval.³⁸ A third aspect is the amount of time it takes to process the stimuli. Ornstein found that an automatic response to the stimuli resulted in the experiencing of a shorter duration.³⁹ Idson and Massaro investigated along this line and concluded that since the experience of duration depends on the processing time of the information in that duration, it follows that judgments of durations are reflexive.⁴⁰

³⁵William James, *The Principles of Psychology* (New York: Holt, 1890).

³⁶R. E. Ornstein, *On the Experience of Time* (Harmondsworth, Middlesex: Penguin Books, 1969); L. Buffardi, "Factors Affecting the Filled-duration Illusion in the Auditory, Tactual, and Visual Modalities," *Perceptions and Psychophysics*, 7 (1971): 36-45.

³⁷Ornstein, *On the Experience of Time*.

³⁸James, *Principles of Psychology*, 624.

³⁹Ornstein, *On the Experience of Time*, 81-2.

⁴⁰W. L. Idson and D. W. Massaro, "Perceptual Processing and Experience of Auditory Duration," *Sensory Processes*, 1 (1977): 316-337.

7.3.2 Perception of Pulses

Although the term “beat” is more commonly used by musicians, the term “pulse” will be used here, since this discussion refers to the time interval rather than the duration.

Fraisse has found that children as young as three years old were able to tap in synchronization with a pulse.⁴¹ Synchronization occurs from the third heard pulse, and is actually easier to perform than tapping *after* the pulse. Fraisse explains that this is a unique behavior; while normally a response to a stimulus comes *after* the stimulus, in the case of pulses, “the signal for the response is not the sound stimulus but the temporal interval between successive signals.”⁴² It seems that it is the regularity that allows synchronization, but since other stimuli, such as rhythmic patterns, may be synchronized, Fraisse concludes that anticipation is the more crucial factor. Although synchronization is unlikely to occur within tempo changes such as *rubato*, anticipation allows the perception of the pulse to persist. While Fraisse found synchronization to be possible in time intervals that ranged between 0.2 and 1.8 seconds (both beyond the M.M. markings; but they would have been M.M.=33 and 300, respectively), the performance was best between M.M.=75 and M.M.=150.⁴³

Pulses are regarded by many researchers as a grid against which rhythm is perceived. Povel explains the perception of the pulse (using the term “beat”) as

⁴¹Fraisse, “Rhythm and Tempo,” 154.

⁴²Fraisse, “Rhythm and Tempo,” 154.

⁴³Fraisse, “Rhythm and Tempo,” 155.

a "segmentation of the sequence [that is, the rhythmic pattern] into parts of equal length. The segmentation is based on the detection of accented events occurring at equal intervals in the sequence."⁴⁴ (See page 197 for more details on Povel's research.)

Monahan followed this line and suggested that the metric structure of pulses and meter as a temporal grid is analogous to the scale schemata as a melodic grid.⁴⁵

Clarke observed that it is only the beat which is directly timed by the listeners and that any subdivisions of the beat are perceived as parts of the beat.⁴⁶ Moreover, the perception of tones shorter than the beat is less precise than the perception of the beat.

How do the listeners determine the level of pulses they refer to as beats? Following Povel's model, there should be no difference between metrically accented pulses (or "downbeats") and the regular pulses. However, there are many experiments that show that there is a certain rate that makes a pulse level primary. A primary level, for our discussion is the level on which the temporal events are processed most efficiently. Researchers were looking for a natural pace for psychological events (also referred to by Fraisse as "spontaneous tempo").⁴⁷ In addition to being the optimal tempo for perception, it is believed to be the preferred tempo for tapping, and stimuli at faster

⁴⁴Povel, "Internal Representations of Simple Temporal Patterns," 4.

⁴⁵C. B. Monahan, "Parallels Between Pitch and Time: The Determinants of Musical Space" (Unpublished Ph.D. Diss., University of California, Los Angeles, 1984).

⁴⁶E. F. Clarke, "Structure and Expression in Rhythmic Performance," in *Musical Structure and Cognition*, eds. P. Howell, I. Cross and R. West (New York: Academic Press, 1985).

⁴⁷Fraisse, "Rhythm and Tempo," 151.

or slower paces tend to be assimilated to this preferred tempo. Both Woodrow⁴⁸ and Fraisse maintained that this tempo is M.M.=80-100. As was mentioned above, M.M.=100 was the tempo in which performance was best.

Fraisse maintains that all living creatures have characteristic motions that are regulated physiologically at a certain rate (walking, heart beat, sucking, flying).⁴⁹ The tempi of these spontaneous, reflex-like activities are highly individual. Research (conducted mostly with tapping on a table) shows that while there are big differences between individuals, the same person has repeatedly almost the same tempo.⁵⁰

Since the pace of both walking and heart-beating is M.M.=60-120, there were some attempts to investigate the existence of a physiological "pacemaker." Mishima found correlation between spontaneous tempo and the pace of walking,⁵¹ but no such correlation was found with the heart-beat.⁵²

⁴⁸Woodrow, "Time Perception."

⁴⁹Fraisse, "Rhythm and Tempo," 151.

⁵⁰Fraisse, "Rhythm and Tempo," 153.

⁵¹J. Mishima, *Introduction to the Morphology of Human Behavior. The Experimental Study of Mental Tempo* (Tokyo: Tokyo Publishing, 1965), cited in: Paul Fraisse, "Rhythm and Tempo," in *The psychology of music*, ed. D. Deutsch, Academic Press Series in Cognition and Perception (New York: Academic Press, 1982).

⁵²M. Tisserand and J. Guiolt, "Etude du Tempo de 335 Sujets Masculins dans la Région Parisienne," *Biotypologie*, 10, no. 11 (1949-1950): 89-94, cited in: Paul Fraisse, "Rhythm and Tempo," in *The psychology of music*, ed. D. Deutsch, Academic Press Series in Cognition and Perception (New York: Academic Press, 1982).

7.3.3 Perception of Meter

As was pointed out by Cook, rhythm exists in any piece of music, but meter is not necessarily a part of any music. For meter to be created, the music must have clearly defined durations and “some sense of steady motion.”⁵³ Clarke regards pulses as temporal markers around which rhythms are perceived.⁵⁴ Cook, using the same approach, regards meter as a set of higher level temporal markers.

Povel’s model of pulse perception by segmentation was mentioned in Section 7.3.2. Models of meter perception are based mostly on grouping. The most basic example of grouping is the spontaneous grouping of equal sounds in groups of two or three. Fraisse, in his review of a series of such experiments starting as early as 1894,⁵⁵ holds that the phenomenon revealed in these experiments “underscores the perceptual and spontaneous character of rhythmic grouping.”⁵⁶ These groupings, called by Fraisse *subjective rhythmization*, happened within the same range that synchronization occurred (0.2 to 1.8 seconds), the range within which it is possible to link two separate stimuli; at a faster rate than 0.2 seconds it is impossible to discriminate two separate stimuli, while at a slower rate than 1.8 seconds the two stimuli are no longer linked to

⁵³Cook, *New Approach to Musical Meter*, 18.

⁵⁴Clarke, “Structure and Expression in Rhythmic Performance,” 224.

⁵⁵E. Meumann, “Untersuchungen zur Psychologie und Aesthetic des Rhythmus,” *Philosophische Studien*, 10 (1894): 249-322, 393-430, cited in: Paul Fraisse, “Rhythm and Tempo,” in *The psychology of music*, ed. D. Deutsch, Academic Press Series in Cognition and Perception (New York: Academic Press, 1982).

⁵⁶Fraisse, “Rhythm and Tempo,” 156.

a perceptual unit. In musical terms, this range implies several pulse levels: the upper limit, 1.8 sec, corresponds to the length of a $\frac{3}{4}$ measure with a beat of M.M.=100, while the lower limit would correspond to the second subdivision of the beat (16th notes in a quarter note beat) in a tempo of M.M.=75. It follows that when sounds are equal in all aspects, we tend to group them in groups of two or three on most of the musically relevant pulse levels; a third subdivision of the beat is quite uncommon and usually very local or ornamental, while the perceptual validity of larger units than the measure (or hyper measures) is debatable (see Section 7.3.5). One example of how music can stretch the limits of human time perception is the Adagio of Bach's *Brandenburg Concerto No. 5* (Figure 7.4). Written in $\frac{4}{4}$, it is often conducted in a pattern of four divided beats, the division of the beats being MM=60 or slower. This results in a measure which is four seconds long or more, yet this time unit is perceivable. The perception is retrospective, though; it gets clearer as the piece gets more familiar and more anticipation takes place.

If meter may be perceived without any clue in the stimuli, then to prevent the listener from creating his/her own meter the composer's clues should be strong enough to overcome the *subjective rhythmization*. Furthermore, once a meter is established, it perpetuates itself even in the face of contradicting information.⁵⁷ This persistence of the meter allows us to perceive changes such as syncopations and hemiolas, without

⁵⁷H. A. Simon and R. K. Summer, "Pattern in Music," in *Formal Representation of Human Judgment*, ed. B. Kleinmuntz (New York: John Wiley, 1968).

Figure 7.4: Bach, *Brandenburg Concerto No. 5*, Adagio

switching to the new metric grouping suggested by it.

Vos has measured recorded performances of Bach's *Well Tempered Clavier*, and found that the duration of the measures ranged between 1.75 seconds and 4.8 sec.⁵⁸ Fraisse cites this experiment and derives from it conclusions about the duration of a "rhythmic group."⁵⁹ He suggests a duration of two to five seconds for a rhythmic group, but as shown by MacDougall, to perceive longer groups, their elements must be processed into chunks to reduce their number.⁶⁰

Brower suggests two strategies for meter perception, depending on the duration of a measure.⁶¹ The first model, for shorter measures, is called by Brower *entrain-*

⁵⁸P. G. Vos, "Identification of Metre in Music," *Report 76 ON 06* (University of Nijmegen, 1976).

⁵⁹Fraisse, "Rhythm and Tempo," 171: It seems though, that the issue discussed at that point is meter, and not rhythm, and the following is based on the assumption that Fraisse correlates "rhythmic group" with a measure, and that at either case the issue is the possible length of a perceptual unit.

⁶⁰R. MacDougall, "The Structure of Simple Rhythm Forms," *Psychological Review* (Monograph Supplements), 4 (1903): 309-416.

⁶¹Brower, "Memory and the Perception of Rhythm."

ment,⁶² and it resembles the phenomenon of *synchronization*. The *entrainment* strategy can be used as long as the measure is short enough to be stored in the sensory memory, and it involves producing internal temporal markers that will synchronize with those in the music. As mentioned in Section 7.3.2, the upper limit for synchronization is 1.8 sec. For longer measures, Brower maintains that the *counting* strategy is used. This does not mean literally counting the number of pulses, but rather perceiving each pulse as a “one” or a “two.” A simple way to experience the difference between the two strategies is to set a metronome to M.M.=80 and synchronize with the pulse (preferably using voice, to reduce the motoric component of the memory); then, to set it to M.M.=40, and to try to synchronize *without* adding a mental pulse in between each two pulses. One can go back and forth between those two tempi, trying to find out at which metronome marking the tendency to add a subpulse begins. Adding a subpulse is a cognitive process, the occurrence of which indicates that the capacity of the sensory memory was exhausted. This is the limit of synchronization, or *entrainment*. Note that M.M.=40 is only a 1.5 sec pulse; synchronization may go as far as 1.8 sec, but it becomes imprecise. Again, this does not mean that we have no “feel” for measures longer than 1.8 sec. The counting strategy means that there is less duration-based anticipation and more stimulus-based retrospective perception of the measure. As Brower explains, when using the *counting* strategy “we are more

⁶²Brower, “Memory and the Perception of Rhythm,” 26.

likely to pay attention to cues found in the music itself, such as similarity, change, pattern repetition, and accent.”⁶³

We have seen so far that the duration of an average beat is 0.60-0.75 seconds, and that it is the beat which is directly timed by our perception; that the duration of a measure is in the range of 2-5 seconds; and that we tend to group elements in twos and threes. Pulses and measures though, were treated so far as perceptual entities, regardless of their representation in music notation. We may find meter signatures indicating more than two or three beats in a measure, and other meter signatures implying beats longer or shorter than these durations. This means that meter signatures cannot be read literally.

There is disagreement between music theorists about compound meters. Joel Lester regards compound meters as multiples of duple or triple meters, implying that these meters contain more than three beats in a measure.⁶⁴ This view contradicts all the above conclusions drawn from research of time perception. Allen Forte, on the other hand, regards compound meters as duple or triple meters in which the beat is subdivided into three.⁶⁵ As for quadruple meters, perception theories do not accept their existence other than as two groups of two. It seems that the difference between a duple and a quadruple meter lies in two musical aspects: (1) the degree

⁶³Brower, “Memory and the Perception of Rhythm,” 26-27.

⁶⁴Lester, *Rhythms of Tonal Music*, 46, and note 1 p. 267.

⁶⁵Allen Forte, *Tonal harmony in Concept and Practice*, 3rd ed. (New York: Holt, Rinehart & Winston, 1979), 25.

of accentuation of the third beat and (2) the motivic structure. If the third beat is strongly accented it may articulate a new measure, while if the accent is relatively weak, it may be perceived as a secondary accent of the same measure. Of course, this is a continuum and there is no specific degree of accentuation that is enough to articulate a new measure. As for the motivic structure, a repeated motive which is two beats long may articulate duple meter, while a continuous four beat long motive may guide our perception to a four beats measure.

Even in simple meters which usually indicate the number and value of the beats quite clearly, the performer needs to use some judgment as to whether the indicated value really represents a beat. A famous example is the second movement of Beethoven's *Symphony No. 9*. Sections of this movement were often used as examples for hyper-measures, as Beethoven has indicated "*Ritmo di tre battute*" (Rhythm of three measures). Since the metronome marking is $\text{♩} = 116$, it is obvious that the beat is $\text{♩} = 116$, and not $\text{♩} = 348$ as may be implied by the $\frac{3}{4}$ meter signature (Figure 7.5). It follows that the "hyper-measure," referred to by theorists such as Cooper and Meyer,⁶⁶ is perceptually a regular measure.

⁶⁶Cooper and Meyer, *Rhythmic Structure of Music*, 5.

Figure 7.5: Beethoven, *Symphony No. 9*, second movement

7.3.4 Perception of Rhythmic Groups

A rhythmic group is a small perceptual unit, within which smaller durations are not perceived individually, but as a part of the larger unit, the *Gestalt*, as described by Brower: “Instead of measuring individual durations against the background of an internally generated metric grid, the listener recognizes subpulse rhythms by their qualities of evenness or unevenness, twoness or threeness, accentedness or unaccentedness, and so on.”⁶⁷

As a perceptual unit, the duration of a rhythmic group is limited. While Fraisse was quoted earlier as regarding the duration of a rhythmic pattern from two to five seconds,⁶⁸ we must keep in mind that Fraisse also equated a rhythmic pattern with a measure. It is understandable why it is difficult to perceive a rhythm longer than five seconds as a unit, but two seconds may definitely contain more than one unit. Amos gives an example of an extremely short rhythmic group in Beethoven’s *Symphony No.*

⁶⁷Brower, “Memory and the Perception of Rhythm,” 24.

⁶⁸Fraisse, “Rhythm and Tempo,” 171.

Figure 7.6: Beethoven, *Symphony No. 7*, Scherzo

7. Scherzo (Figure 7.6), which lasts, in Karajan's recording, 0.455 sec.⁶⁹

Within this time span (from about half a second to five seconds) there are many other factors that articulate a group. Following the *Gestalt* concept of a perceptual unit, elements cannot form a group when they are too close together to be differentiated, or too far apart to be associated. Fraisse regards two to three elements per second as an optimal frequency.⁷⁰

In 1983, Handel and Lawson conducted an experiment that showed the effect of tempo on the perception of rhythmic patterns.⁷¹ Subjects were presented with two different patterns of two against three, in two different pitches, and were asked to tap the rhythm. As shown in Figure 7.7, at the slow tempo of three seconds per measure, subjects tapped the combination of both rhythms; at 1.6 sec per measure

⁶⁹Amos, *Perceiving Musical Rhythms*, 7.

⁷⁰Fraisse, "Rhythm and Tempo," 164.

⁷¹S. Handel and G. R. Lawson, "The Contextual Nature of Rhythmic Interpretation," *Perception and Psychophysics*, 34 (1983): 103-120.

Source: Handel and Lawson (1983), experiment 3.

Figure 7.7: Effect of tempo on the perception of rhythmic patterns

they followed the triple subdivision (0.53 sec apart); at 1.2 sec per measure, they followed the duple subdivision (0.6 sec apart), and at 0.6 per measure, some of the subjects shifted to one tap per measure.

Grouping Principles There is a wide agreement about two articulating factors, which correspond to two *Gestalt* laws. The first is that a long duration, or a long time interval between elements, tends to produce an end of a group. In *Gestalt* terminology, this is a manifestation of the law of proximity. The second is that a louder sound tends to articulate a beginning of a new group, which corresponds to

the *Gestalt* law of good continuation.⁷² But as Fraisse points out, duration and intensity interact in our perception: a louder sound is perceived as longer, and a longer sound as louder.⁷³ Fraisse concludes that a slight prolongation is perceived as an increase in loudness and thus articulates a beginning of a group, while a much longer sound creates a pause and articulates an end of a group. In this sense, a string of similar elements corresponds to a long duration. A series of experiments, conducted by Garner and his colleagues and reviewed by Fraisse, used isochronous sounds with rhythmic patterns created by differences in intensity or pitch.⁷⁴ These experiments showed that while in a sequence with a three elements string the string most frequently began the group (XXXOXXOO), in a sequence with four elements string, the string most frequently ended the group (OOIOXXXX). Fraisse added that a higher pitch tends to begin a group, and indeed higher pitches are often perceived as louder. These general tendencies should not be taken as rules, since, as shown by Vos, it is possible to perceive a beginning or an end of a group without any differences in intensity.⁷⁵

⁷²Fraisse, "Rhythm and Tempo," 159; Cooper and Meyer, *Rhythmic Structure of Music*, 10.

⁷³Fraisse, "Rhythm and Tempo," 160.

⁷⁴F. L. Royer and W. R. Garner, "Response Uncertainty and Perceptual Difficulty of Auditory Temporal Patterns," *Perception and Psychophysics*, 1 (1966): 41-47; F. L. Royer and W. R. Garner, "Perceptual Organization of Nine Element Auditory Temporal Patterns," *Perception and Psychophysics*, 7 (1970): 115-120; D. Preusser, W. R. Garner and R. L. Gottwald, "The Effect of Starting Pattern on Descriptions of Perceived Temporal Patterns," *Psychonomic Science*, 21 (1970): 219-220.

⁷⁵P. G. Vos, "Temporal Duration Factors in the Perception of Auditory Rhythmic Patterns," *Scientific Aesthetics*, 1 (1977): 183-199.

Cooper and Meyer explain that “though *similarity* tends to create cohesion, repetition usually makes for the separation of groups.”⁷⁶ This phenomenon was shown in an earlier series of experiments using isochronous sounds (1954 and 1956, summarized by Fraisse⁷⁷). In these experiments, the patterns were perceived with similar sounds grouped together, and thus, in a sequence of L (loud) and S (soft) sounds such as LSSLLSSLLSS, the pattern perceived was LLSS, and not LSSL.

In a research on Morse code in 1943, Plotkin showed that the difficulty to learn a signal increased with the number of elements, but more so with the number of strings.⁷⁸ Garner's experiments also confirmed the findings regarding the tendency to find the simplest pattern possible.⁷⁹ In the first experiment in the series, the tendency to keep the pattern simple proved to be so strong that in the case of the three elements sequence, the pattern XOXOXOXX, (which has four changes as opposed to the previously mentioned pattern XXXOXOXX which has three) was perceived only once out of 128 times.⁸⁰

Within-Group Organization We have seen a few principles by which the beginning and the end of rhythmic groups are perceived. The next step is to examine

⁷⁶Cooper and Meyer, *Rhythmic Structure of Music*, 9.

⁷⁷Fraisse, “Rhythm and Tempo,” 162-164.

⁷⁸L. Plotkin, “Stimulus Generalization in Morse Code Learning,” *Archives of Psychology*, no. 287 (1943): n.p.

⁷⁹W. R. Garner, “Good Patterns Have Few Alternatives.”

⁸⁰Royer and Garner, “Response Uncertainty.”

the organizational processes within the rhythmic group.

It seems that most rhythmic patterns are products of only two durations. While research measuring channel capacity showed that we can distinguish stimuli in about five different levels, the capacity for distinguishing absolute durations was only two levels, and rarely three.⁸¹ Fraisse supports this assertion in three ways: (1) When he asked subjects to produce both regular and irregular rhythmic groups, they used two kinds of durations: short and long; (2) When subjects were asked to reproduce a pattern with three different durations, in the reproduction the third duration was assimilated to one of the other two; (3) Fraisse points to the human tendency to incorporate two durations such as in the Greco-Latino prosody, the Morse code, and music: "In music, there is, at any given moment, a play of two notes that are in a ratio of one to two or one to three . . . These two notes represent 85% to 95% of the movement."⁸²

The perceptual superiority of the 1:2 ratio is affirmed in the first of a series of experiments held by Povel. Subjects were asked to repeat rhythmic patterns of repeated two time-intervals. Figure 7.8 is a musical representation of Povel's chart of durations and ratios.⁸³ The subjects tended to assimilate all ratios to 1:2, and the most imprecise patterns were the ratios 3:4 and 4:5 in which the tendencies to

⁸¹L. E. Murphy, "Absolute Judgments of Duration," *Journal of Experimental Psychology*, 71 (1966): 260-263.

⁸²Fraisse, "Rhythm and Tempo," 168.

⁸³Povel, "Internal Representations of Simple Temporal Patterns," 7.

Figure 7.8: Povel's durations ratios

assimilate 1:2 and 1:1 conflicted.

A common case of 1:2 assimilation as a musical behavior is the tendency to alter the rhythm $\text{♩} \text{♩}$ to $\text{♩} \text{♩}$. Fraisse found that higher ratios, or smaller duration-differences are more typical to irregular groups, or *arrhythmia* than to rhythmic patterns.⁸⁴

Fraisse has sampled music literature from Beethoven to Stravinsky for the frequency of the different durations and found that in each piece, 80% to 90% of the notes were of two durations in a 1:2 ratio. This “durations count” should be taken with some reservation, though. First, the music sampled does not fully correspond to the Common Practice style; rather, it represents a time in which instrumental style was the prevailing style. Second, Fraisse did not measure how much time is spent on each duration, but how many times each duration appeared. Thus, if, for example, two $\frac{1}{4}$ measures were to be filled with 16th notes and five measures with whole notes, the count would still be thirty-two 16th notes and only five whole notes, although

⁸⁴Fraisse, “Rhythm and Tempo,” 165.

more time was actually spent on whole notes.

A more musically-relevant interpretation of the prevalence of the 1:2 ratio is given by Povel, who found it hard to believe that professional musicians could not repeat a simple ratio such as 1:3 or 1:4, as was suggested by the results of his first experiment. Assuming the problem was the context, Povel constructed another experiment in which each short interval was repeated until it equaled the longer interval.⁸⁵ Thus, the subjects were able to perceive the short intervals as subdivisions of the long intervals. The improved performance in this experiment supported Povel's model of perceiving rhythms within the framework of beats, and led him to the following conclusion:

The most significant aspect of this model seems to be that the description of the sequence does *not* start with the smallest duration and try to define the total sequence in terms of that duration; rather, it starts at a higher level, the beat interval, and describes the sequence in terms of this interval. Any duration shorter than the beat interval is described as a subdivision of the beat interval.⁸⁶

The difficulty of having two types of subdivisions (duple and triple) at the same time is demonstrated by the poor results of stimulus no. 6 (Figure 7.8).

Many theorists agree with this concept that rhythm is not perceived as a sequence of durations, but holistically. Royer and Garner found that subjects did not begin to respond until a pattern was perceived.⁸⁷ It was also found that it takes twice

⁸⁵Povel, "Internal Representations of Simple Temporal Patterns," 13.

⁸⁶Povel, "Internal Representations of Simple Temporal Patterns," 13.

⁸⁷Royer and Garner, "Response Uncertainty."

as much time to describe a pattern than it takes to reproduce it, which implies that while analysis is necessary for the description (or the notation) of a rhythm, it is not necessary for its reproduction.⁸⁸

In discussing the relative nature of rhythmic perception, Monahan showed that a contour of relative durations takes part in the encoding of rhythm. the way a melodic contour does in the encoding of melodies.⁸⁹ Her first experiment showed that a change in tempo of similar rhythms improved the subject's judgment of similarity. Her interpretation of these results was that the changed tempo shifted the subjects' focus from the absolute to the relative durations.

7.3.5 Perception of Hyper-Measures and Large-Scale

Rhythmic Groups

The attempt to discuss the perception of larger groups raises some problems. First, larger rhythmic groups are "real" musical events, and as such, they do not fit well into the laboratory experiments. Second, larger groups exceed the capacity of the sensory memory and the short-term memory. Brower maintains that one or two levels above the pulse is the highest level possible before the capacity of the short-term memory is exceeded.⁹⁰ As groups processed by the long-term memory, they are subjected to

⁸⁸Preusser, "Effect of Structure and Rate on Recognition of Temporal Patterns."

⁸⁹Monahan, "Parallels Between Pitch and Time."

⁹⁰Brower, "Memory and the Perception of Rhythm," 26.

elaborate cognitive processes that differ from one person to another. It follows that music theorists do not agree as to what these larger groups are, and the discussion deviates to the realm of music analysis. Lester commented on the concept of hyper-measures: "Even among those who accept the notion of metric and other accents on higher levels of structure, there is disagreement concerning the nature of such accents, the factors that give rise to them, and their placement."⁹¹ Brower questions theorists who use metric reduction to the highest structural levels (such as Arthur Komar⁹² and Jonathan Kramer⁹³) or even those who limit its use to the foreground and middleground (such as Carl Schachter⁹⁴ and William Rothstein⁹⁵):

The appeal of metric reduction is that it allows us to do in theory what we are unable to do in fact . . . bring large-scale rhythmic patterns within the bounds of our perceptual horizon . . . Theorists who apply the principles of reduction to rhythm tend to make two assumptions: 1) that the principles of rhythmic organization are the same at all levels; and 2) that we perceive large-and small-scale rhythmic patterns within the bounds of our perceptual horizon.⁹⁶

Cooper and Meyer developed a comprehensive rhythmic theory that leads to large-scale rhythmic groups, based on prosodic feet, and some criticism of their theory was mentioned in Section 7.2. Lester maintains that the use of prosodic feet cannot work for large-scale musical accents since in prosody, a syllable is considered accented

⁹¹Lester, *Rhythms of Tonal Music*, 12.

⁹²Arthur Komar, *Theory of Suspensions* (Princeton: Princeton University Press, 1971).

⁹³Kramer, *The Time of Music*.

⁹⁴Carl Schachter, "Rhythm and Linear Analysis: Aspects of Meter," *The Music Forum*, 6, no. 1 (1987) 1-59.

⁹⁵W. Rothstein, *Phrase Rhythm in Tonal Music* (New York: Schirmer Books, 1989).

⁹⁶Brower, "Memory and the Perception of Rhythm," 20.

throughout the span of time until the next unaccented syllable. A musical accent, on the other hand, happens at a point of time and not over a span of time.

Theorists who accept the existence of higher level meter agree that at a higher level the meter is not rigid anymore, but it becomes a more flexible and organic structure, with the sections articulated by aspects such as orchestration and tonality. While measure-level meter tends to persist and deviations are perceived as syncopations, higher level meter changes with the musical cues. As for large-scale rhythmic groups, Cooper and Meyer maintain that on this level, temporal cues become more important (as opposed to a dynamic accent that has an effect over only one or two measures). Thus, larger groups tend to be end-accented since, as discussed in Section 7.3.4, temporal accents tend to articulate the end of a group.

Chapter 8

Summary and Conclusions

8.1 Summary of Research

So far, different aspects of music have been discussed in isolation. None of these aspects, though, exist in “real music” independently of other aspects. Time, for one, is an aspect that exists in any musical event, as isolated as it may be, and it impacts the perception of melody, harmony and tonality. Other aspects, such as timbre and loudness, which are less intrinsic to the musical syntax have received little or no attention here, but they, too, interact with each other and with the perception of pitch and time. These interactions may facilitate perception, as in the case of pitch perception which is facilitated by temporal organization. It can also hinder perception, as in the case of interaction between timbre and pitch.

The musical excerpts discussed were no longer than a simple short melody, and in most cases they were shorter. Larger scale musical excerpts were only mentioned with the discussion of hyper-measures, and thus only from the angle of time perception. The perception of large-scale pitch relations, such as the tonalities of large sections of a movement is not readily susceptible to a laboratory experiment.

It is no wonder, then, that ear training teachers do not find research of music perception relevant for developing an effective and meaningful approach to music education. Instead, they bring to their classrooms the approaches that they have encountered in the course of their own music education experience, and, hopefully, alter, modify and improve them as they gain more and more teaching experience. This teaching experience, the answers it provides, and more so, the questions it raises, may be an ample source for experiments in music perception and cognition, as was put by David Butler: "Who is best equipped to do the sort of research that will begin to give us knowledge about aural training, based on evidence, rather than beliefs about aural training, based on faith? Who better than music educators?"¹

It is clear, then, that researchers of music perception and cognition can benefit from the experience of ear training teachers. However, the impetus for this thesis is the belief that ear training teachers can benefit from music perception research as

¹David Butler, "Why the Gulf Between Music Perception Research and Aural Training?" by R. E. Radocy (A D. Butler presentation), *Bulletin of the Council for Research in Music Education*, 132 (spring 1997), 47.

well.

It was shown that perception is facilitated by mental representations and that these representations operate, at least for some of the musical aspects, as categories. These categories are learned, and their number, the different aspects that they exist for and the stability of their mental representations in memory vary mostly according to level of expertise. Musicians, for example, have categories for intervals, while non-musicians do not. In the case of single-pitch perception, though, the correlation with expertise is not direct; although all musicians have some degree of memory for pitch, there is no evidence that nonpossessors of absolute pitch have mental representations of specific pitches. If they do, these representations must be so unstable that the use of relative pitch is preferred whenever possible. While a high degree of musical expertise does not imply possessing absolute pitch, absolute pitch possessors are not necessarily trained musicians.

Trained musicians, then, are able to perform categorizations such as categorizing out-of-tune intervals to the nearest interval category, triads with an out-of-tune third to a major or a minor triad, or complex duration ratios to simple ratios of 1:2 or 1:3. absolute pitch possessors are also able to categorize out-of-tune pitches to standard musical pitches. While these categorizations happen involuntarily, it is possible to suppress them when finer discriminations are required. When presented with a smaller choice of categories, trained listeners are likely to make finer judgments.

For nonmusicians, on the other hand, the pitch dimension was shown to be perceived as a continuum; intervals are judged by their size, octave generalization is weaker, a tonic is perceived as more stable when lower, and the hierarchy of stability correlates with proximity to the tonic. Interval size was also shown to impact musically untrained listeners' judgments of chords similarity. The same difference in perception also exists between possessors and nonpossessors of absolute pitch; possessors perceive pitch classes by their categories, while nonpossessors perceive pitch height.

Some categories are rendered a higher perceptual status: they are perceived directly and therefore faster and with greater accuracy. Such superior perceptual units are the major triad, the tonic (within the context of its key), intervals such as an octave and a fifth, the beat and the down beat. Within a major tonal context, there is a perceptual scale-degree hierarchy of tonic, 5th, 3rd, 4th, 6th, 2nd, 7th, #4th, b7th, b3rd, #1st. Fixing the tonic on C, this hierarchy is very similar to the hierarchy of errors made by absolute pitch possessors; C is the most easily perceived pitch class, followed by G, E and so forth.

Perception of music was shown to be both holistic and analytical, depending, again, on expertise, but also on the type of stimulus, and the type of musical activity. While most people perceive single pitches holistically (i.e., as their fundamental, rather than a series of partials), musicians tend to perceive more complex stimuli

analytically, using their left brain hemisphere. The level of the analysis depends on the type of activity; chords, for example, tend to be perceived holistically,² but may be analyzed and arpeggiated at will. Rhythmic groups may be recalled without analysis, but analysis may be necessary for their notation. While listening to a familiar harmonic progression, though, it is likely that the chords will not be analyzed, while the chord progression will. Put in a wider context of a complete movement, that same chord progression may be perceived holistically as one phrase which, in the context of a larger section, is analyzed into phrases.

Although analysis is possible for some musical entities, the analyzed stimulus is not perceptually equivalent to the whole. In accordance with the *Gestalt* principle that the whole is more than the sum of its parts, chords are more than the sum of their component intervals; melodies are more than the sum of their pitches, intervals, or rhythmic groups; intervals are more than the sum of their component semitones. The more meaningful aspect of all these musical phenomena are the relations, rather than the individual components; pitch relations, rather than absolute pitch, and duration ratios rather than specific, clock-time durations. Moreover, our perception for these components (absolute pitch and specific durations) was shown to be poor. Without disregarding the value of analytical perception, it should be noted that holistic perception is faster than analytical perception.

²Hubbard and Stoeckig, "Musical Imagery."

Figure 8.1: Chunking into memory bits

The higher cognitive-level perception performed by trained musicians also implies better retention of musical aspects in both short-term and long-term memory. With pitch being processed in a specialized, single-channel system, short-term memory is impaired by interference of the same type of stimulus, and improved by consolidation (repetition of the stimulus). Correlation between speed and accuracy was found in the case of absolute pitch possessors, who were using a “one step” strategy; extra steps are not only time consuming, but may also interfere with the stimulus.

The capacity of short-term memory is believed to be five to nine “bits” of information.³ What these bits are may vary by the type of stimulus, and by the experience of the listener. In Figure 8.1 there are fifteen notes, and they may occupy fifteen memory bits of a nonmusician. It takes some musical expertise to be able to create larger chunks of arpeggiated triads, which will reduce the number of bits to four. For an experienced musician, though, there is only one bit of information in this passage: an arpeggiated authentic cadence.

³Miller, “Magical Number Seven.”

Not all types of groupings, though, are based on musical expertise. Grouping processes equivalent to the visual perception principle described by *Gestalt* theory are spontaneous and universal. Such principles that are transferable to music are the laws of *common fate* (percepts that move in the same direction are grouped together), *good continuation*, *proximity* and *similarity*. Various aspects such as pitch, timbre, spatial location, loudness and familiarity determine the relative force of the perception principles when more than one force is in effect. Pitch was shown to be a dominant aspect that tends to overcome changes in timbre, location and loudness, and it operates even with no musical expertise. Experience does have an impact on streaming by pitch, as was shown by Smith, Hausfield, Power and Gorta, who found that musicians were influenced by musical expectations, and when these were met, stimuli were streamed by pitch even in the presence of conflicting influences of other aspects.⁴

When the capacity of short-term memory is exceeded, perception can only be retrospective; one can perceive groups containing more bits of information, once the stimulus is stored in long-term memory where music is stored in an abstracted form. The Western tonal system is one of the most long lasting schemata. Against this template of relationships and forces, expectations are fulfilled or denied; music stored in this abstracted form can be recalled by paraphrasing without distorting

⁴J. Smith *et al.*, "Ambiguous Musical Figures and Auditory Streaming," *Perception and Psychophysics*, 32 (1982): 454-464.

its meaning. Music that does not invoke such specific expectations (such as atonal music) is harder to memorize.

8.2 Goals of Ear Training

It was seen that research of music perception and cognition provides information regarding perceptual and cognitive processes. It sheds light on weaknesses and demonstrates cognitive abilities that develop with training. In order to examine how this information can be put to use in college level aural training programs, it is necessary to examine the goals of such programs.

If such goals are regarded as a list of skills that a program aspires to equip its student with, then they may differ widely, reflecting the variety of entry and exit levels in colleges. These skills may be certain levels of sight singing, dictation, identification of isolated intervals and chords, and so forth. These goals are easily managed; a program is planned, using an approach chosen or developed by its head, and at the end of the average two years allocated for aural training, the students are tested on their proficiency in the specific skills and within the specific approach.

If, however, these skills are regarded as means rather than goals; if a program aspires to create professional musicians rather than successful students, and if it is agreed that creating literate musicians with understanding of small and large scale musical structures takes more than two years, then it may be that any college level

music program would have the same goal, which will be to instill in its students cognitive skills, concepts and work habits that will enable them to embark on a life long process, most of which will take place after their graduation. This goal, though, is not as readily testable, not easily graded, and does not fit into the formal demands of college education.

Paula Telesco has stated that her goal in teaching ear training is “to produce musicians who can hear and think about music equally well — musicians who can understand and see what they hear and hear and understand what they see.”⁵ It should be noted that for the purpose of this thesis which focuses on aural skills, literacy, or mastering music notation, is regarded as a means rather than a goal. Being literate, in music as well as in a spoken language, is not only a handy tool, but a means of understanding structures and relations in ways that are not accessible to the illiterate.

Large scale structures that are beyond the cognitive abilities of students, find their way into classes such as harmony, counterpoint and analysis. This is done with the intent and hope that the theoretical understanding will later facilitate the aural perception of these large structures. As was mentioned above, this goal is often beyond the scope of a two years ear training class, but it may serve as the far-end marker for the purpose of ear training, as was put by David Butler: “what is aural

⁵Paula Telesco, “Contextual Ear Training,” *Journal of Music Theory Pedagogy*, 5, no. 2 (fall 1991), 180.

training, after all, if it is not the acquisition of cognitive skills in music?"⁶

Mental Representations Mental representations are products of continuous exposure to stimuli, followed by recognition of repeated patterns and sorting them into categories. It is safe to assume that labeling these categories improves their stability in long-term memory in the same way that language facilitates thinking, though thinking is possible without verbal activity. Audiation (a term coined by Gordon⁷) is a way of thinking about sounds, or creating a mental image of music. It is a necessary basic skill that enables the activation of mental representations. These activated mental representations facilitate the perception of the represented stimulus.⁸

The process of creating mental representations and categories incorporates such musical aspects as intervals, chords and rhythmic and melodic patterns. A clear and stable mental representation of these categories is likely to overcome interferences such as changes in timbre, context or other interactions with other musical aspects. Expanding aural training to the use of other instruments than the piano, may assist in creating categories that include other timbres.

The process of labeling includes labeling common patterns (such as intervals and chords), the use of syllables for individual pitches and for rhythmic patterns,

⁶Butler, "Why the Gulf . . .," 47.

⁷Edwin E. Gordon, "Research Studies in Audiation," *Bulletin of the Council for Research in Music Education*, (1984): 34.

⁸Hubbard and Stoeckig, "Musical Imagery."

and labeling phenomena pertaining to the tonal system such as modes, functions and cadences. L. Campbell has noted that in the English language, the order of adjectives and nouns requires labeling the type of chord or interval (major, minor etc.) before labeling the interval or chord itself (a third, a triad etc.).

Expansion of Short-Term Memory Short-term memory is the working buffer of perception. Its capacity can be expanded not by stretching the ability to store more information, but by reducing the storage load: “memory-storage needs influence our estimates of perceptual time: a percept containing a large amount of information will require more storage capacity.”⁹ Understanding of relationships facilitates creating larger chunks and thus reducing the number of chunks to be stored. Experience in tonal relationships builds expectations which enables us to extract meaning and store abstractions of the musical excerpts.

Understanding relationships means understanding the forces which operate in the tonal system and create tension and release. Steve Larson has identified three main forces; *gravity*, *magnetism* and *inertia*.¹⁰ Gravity and magnetism are the tendencies of unstable tones to resolve to a more stable tone; while gravity is the tendency to resolve by descending, magnetism is the tendency to resolve to the nearest, more

⁹David Butler, *The Musician's Guide to Perception and Cognition* (New York: Schirmer Books, 1992), 154.

¹⁰Steve Larson, “Scale-degree Function: A theory of Expressive Meaning and its Application to Aural-skill Pedagogy,” *Journal of Music Theory Pedagogy*, 7 (1993): 69-84.

stable tone, in any direction. Inertia, a manifestation of the *Gestalt* law of good continuation, is the tendency of a pattern to perpetuate itself in the same manner it has started. There may be more than one force operating at the same time and with different magnitudes. Larson regards the combination of these forces and their magnitude as characterizing the different tonal functions. Understanding these forces, together with identifying those categories that are of a special status (such as the stable tones) will result in understanding of the tonal system.

Although it is true that the tonal system is based on relativity, there is still a point in improving the retention of single pitches, which was shown to be poor, but susceptible to improvement by training. Better retention of pitch in short-term memory may assisted in retaining a tonic through ambiguous passages and tonicizations, and perceiving larger scale tonal relationships, as well as in a variety of practical situations such as getting cues for sight-singing from a non-adjacent tone, being able to start over from the same pitch after stopping, discussing and maybe hearing other pitches, and so forth.

Encouraging a holistic, rather than analytical perception may also improve short-term memory, since a holistic response is faster. Having mental representations to more complex patterns may assist in a faster perception of these patterns. Thus, a listener who has a mental representation of a major triad is likely to recognize this triad faster than a listener who only has a mental representation of a major third and

a minor third, or a major third and a perfect fifth.

Directed Attention Since attention was shown to be both voluntary and involuntarily, it is a goal of ear training to train listeners to direct their attention at will. This goal has two components: the first is the ability to control the direction of attention to various features at will, and the second is to make knowledgeable choices about where to direct attention.

8.3 Research Based Ear Training Approach

A suggested approach, rather than a suggested sequence of concepts and topics may be more suitable for a variety of programs, regardless of their starting point, scope and exit requirements. The only sequence that is suggested here is the general learning sequence. A learning process should be based on the learner's prior knowledge.¹¹ In the case of music of the Western world, familiarity with the syntax of tonal music exists, at least to some extent, by exposure, and it is quite unlikely to find a person living within this culture and not being exposed to some form of tonal music. Therefore, Common Practice music, being the strictest form of tonal music is a good starting point. Later musical developments within the tonal system are expansions of tonality and can be based on that knowledge. Even atonal music is not meant for

¹¹Jerome S. Bruner, "The Act of Discovery," *Harvard Educational Review*, 31 (1961): 21-32.

the ears of a hypothetical listener who has erased his/her prior knowledge of tonality; such a listener would be hard to find among the population of the Western world above six years old, including the composers of atonal music whose music education happened within the system of tonal music.

Another aspect of the learning sequence is the type of activities involved in the learning process. Teachers often introduce a concept in theory first, assuming that what is not theoretically understood cannot be aurally perceived. Bruner names three stages of learning that correspond to three different types of activities: action stage, iconic stage and symbolic stage.¹² The very first stage, then, is *doing* or *experiencing*; it can be listening, recalling, playing back etc. The iconic stage, creating an image, or a mental representation to a generalization of percepts that were experienced, cannot precede the action stage. The symbolic stage, assigning symbols such as notation or Roman numerals must be last. There is no meaning or need for a symbol that does not stand for a meaningful phenomenon.

8.3.1 Professor Rubinstein's Aural Drill

The aural drill presented in this section is a part of an ear training method developed by my former ear training teacher, Professor Bat-Sheva Rubinstein, and was presented to me while working on my own aural skills under her instruction.¹³ During

¹²Bruner, "The Act of Discovery."

¹³Rubin Academy of Music, Tel Aviv University, Israel.

a course of one year of training I acquired abilities resembling those of absolute pitch possessors, and also expanded my ability of recognizing complex melodic and harmonic tonal patterns. I was able to retain pitch during interferences of other pitches, and identify instantly pitches within the context of keys such as F# major in any octave of the piano. The context of the key was a starting point, but during the drill it has expanded to the point that all twelve tones were used in a nearly random order, and each tone was presented in a different octave. It should be clarified that acquiring absolute pitch is not regarded here as one of the goals of ear training, but acquiring the ability to identify pitches and patterns instantly and accurately even in the most complex tonal contexts certainly is. Professor Rubinstein reports that the average time it takes the students in her classes to acquire such skills is two years. The following is not a full description of Professor Rubinstein's method, but some examples of strategies utilizing research findings cited and ear training goals stated in this thesis. Some of the strategies presented later, in the discussions of sight singing and dictation, are also drawn from Professor Rubinstein's ear training classes.

The aural drill is performed in class, mostly by the teacher playing and the students responding by singing instantly. Since this drill has no specific rhythm or meter, it enables the students to respond with full letter names, including flats, sharps and naturals. The students are encouraged to repeat this drill daily by playing and singing the patterns presented in class. All octaves of the piano are used for the

playing, and one octave, within the students' comfortable vocal range, is used for the recall. Single pitches are presented first, followed by pairs of pitches, longer sequences and simultaneities. Other class activities include melodic and harmonic dictations and sight singing, the strategies of which will be discussed later.

An important aspect of this aural drill is an instant, one-step response; if no such response is possible, it means that the students are not ready for that level of drills, do not yet have the mental representations needed, or the understanding of the relationships presented. Although single pitches, intervals and chords are presented, they are all presented within a tonal context, or a variety of contexts, created by the teacher, and matching the students' ability.

The drill is performed in one key, starting with C major, and as this key is mastered on a basic level, the next key is introduced, alternating both sides of the circle of fifths (C, G, F, D, B $\flat$ etc.). Using a spiral model of learning, once the circle of fifths is completed, C major is revisited with more complex patterns.

Single Pitch Recognition Single pitches are introduced in the order of their perceptual prominence within a given tonal context, which is, in C major, C, G, E, F, A, B, F $\sharp$, B $\flat$, C $\sharp$ and D $\flat$, E $\flat$, A $\flat$. The drill starts with one pitch class (the tonic), repeated over all octaves, and then the second pitch class (the dominant) is added and alternated with the tonic. As more pitch classes are added, they are introduced

in patterns that are typical to the key, and when these patterns are mastered, more complex patterns are introduced. Minor keys are introduced both as parallel and relative. This drill resembles Brady's training¹⁴ and Waters' Experiment¹⁵ (see Section 2.2.6), only it repeats in every key. Thus, tonal relationships are learned and mastered with the pitch names for every key in the circle of fifths. This type of drill can be expanded to responding to two different keys on two different registers of the piano, e.g., drilling C major in the register of C4 and above, alternating with G major in the register of C4 and below. The rate of alternating can increase, and different key relations can be drilled.

Pitch-Combination Recognition Pitch combinations, both sequential and simultaneous, are presented in the same manner that single pitches are presented, but the recall is delayed to after the pattern is completed. In this manner, melodic and harmonic intervals as well as chords can be introduced in the variety of contexts in which they typically occur in real music. Since higher pitches are more conspicuous, simultaneous combinations are initially sung from the top down. In addition to sequential and simultaneous patterns, a sequence of simultaneous events may be presented (Figure 8.2).

¹⁴Brady, "Teach Yourself Absolute Pitch."

¹⁵Waters, "Absolute Pitch."

Figure 8.2: Modes of presentation of pitch combinations

Pitch Retention All the above drills may improve long term memory for single pitches, but it is also possible to improve short-term pitch retention with interferences of other pitches. One such example was the drilling of two different keys simultaneously. In that case, the tonic is either retained in memory, or inferred from the tonal context; but in the moment of actual switch from one tonality to another, it has to be retained.

A more specific pitch retention drill is also done within a tonal context, starting with the easiest pitch to retain, the tonic, and progressing in the same order of perceptual prominence mentioned above. The first stage, the action stage, is done by the teacher playing various pitches or pitch combinations, and the students constantly responding with the pitch to be retained. Initially, the teacher's playing provides tonal support to the retained pitch, but gradually it becomes more distractive (Figure 8.3a).

Figure 8.3: Pitch retention in aural drill

In the second stage, the iconic stage, students are asked to maintain an activated mental representation of the pitch, and while the teacher plays various pitches, students respond only when the retained pitch is played (Figure 8.3b). This drill can be expanded to simultaneities, with the teacher playing harmonic intervals or chords and students respond only when the retained pitch is one of the components of the simultaneity played.

Melodic Tendencies Although patterns of tension and resolution are presented in Professor Rubinstein's aural drill, and students are encouraged to recognize typical and less typical patterns of musical motion, it is possible to further direct students' attention to these patterns by singing the basic melodic relationships of the major (and later minor) key, as presented for C major in Figure 8.4. As each dynamic tone resolves according to its "natural" tonal tendency, this melody is very easy to memorize. After memorizing, those relationships can be drilled by both identification and recall tasks of single pitches. In this context, the dynamic tones are very easy

Figure 8.4: Tendencies in C major

to recall and identify, as their “tendency” makes them more conspicuous. Recalling pitch classes within a context of a key, with attention directed to their tendencies, gives a sense of “relative absolute pitch.” Responses tend to be very fast, especially to the dynamic tones which seem to be accessed directly, without the need to recall the tonic.

8.3.2 Isolated Intervals and Chords

Identifying isolated intervals and chords is one of the most common activities in ear training classes. It is easy to administer, test and grade, but it seems that this type of training gets more than its share of time compared to its importance and urgency for ear training in tonal music. The need to use intervals strategy in sight singing, for example, usually indicates that the student is not familiar with a particular tonal relations in a particular passage. While it is helpful to have the intervals strategy available when the more relevant scale-degrees strategy fails, promoting the study of intervals for this purpose places sight singing as a goal, rather than a means. If,

on the other hand, understanding tonal relationships is a goal and sight singing is a means, then intensive study of isolated intervals can be postponed to a later stage. Even when atonal music is kept in mind as a future goal, the ability of singing and identifying isolated intervals is likely to develop in the process of tonal ear training.

As with single pitch, the goal is a fast response based on a mental representation of the percept as a whole, rather than on analysis into its components. In the case of harmonic intervals and chords, it means a one-step response without arpeggiating. This does not mean that there is no benefit in arpeggiating harmonic intervals and chords, or in analytical perception in general: such an activity was described above (see Figure 8.2 page 218). Yet, this type of activity, perception of isolated simultaneities, provides an opportunity to create mental representations for more complex stimuli, based on their musically-meaningful dimensions.

The dimensions for harmonic intervals were discussed in Section 4.1.2. Intervals may be identified by their degree of consonance, mode (major/minor), size and tonal implications. Chords can be identified by their degree of consonance, type, position and tonal implications. Each of these dimensions may be practiced in isolation (asking students to sort the stimuli by one particular dimension), creating larger categories, or in combination, creating finer categories such as “a consonant small imperfect interval.”

As was discussed in Section 4.1.1 the concepts of consonance and dissonance

may cause confusion. Most students entering college level music programs already have some concept of dissonance, usually based on acoustic considerations. Graybill commented on this issue:

It seems unwise to deny any status whatsoever to acoustical factors . . . At some point the instructor needs to demonstrate systematically the distinction between acoustical and contextual criteria and to show that they may be considered independently. The fact that they so often reinforce each other in tonal music is precisely why this distinction is necessary.¹⁶

As for the perceptual reality of perfect versus imperfect consonances, the musical distinction may be based on whether or not they imply major/minor quality.

In the case of isolated simultaneities, tonal implications refer mostly to simultaneities that imply dominant function, such as the tritone, the minor seventh, the major-minor seventh chord and the diminished seventh chord. It is true that in some contexts these simultaneities may not imply dominant function, but yet, this is their most common implication, and the one that is most likely to be assumed when heard in isolation. Chord position may also be perceived without analysis, i.e., without singing or audiating the bass, since inverted chords are characteristically less stable than root position chords.

Some scholars suggested an order of intervals practice based on their difficulty, i.e., starting with the easiest intervals and gradually adding the more difficult ones.¹⁷

¹⁶Roger Graybill, "Pedagogically Speaking: Consonance and Dissonance in the Freshman Theory Class," *In Theory Only*, 9, no. 5-6 (1987), 53.

¹⁷Ann K. Blombach and Regena T. Parrish, "Acquiring Aural Interval Identification Skills: Random vs. Ordered Grouping," *Journal of Music Theory Pedagogy*, 2, no. 1 (spring 1988): 113-132; Ortmann, "Problems in the Elements of Ear-dictation."

Dissonant intervals were shown to be more conspicuous and therefore easier to identify,¹⁸ but in general, “difficult” intervals are those that are most often confused with other intervals, and thus their degree of difficulty depends on the degree of their similarity to other intervals presented at the same identification task. As was discussed above, presenting listeners with a smaller choice of categories enables them to make finer distinctions. Thus, when all intervals are presented, the choice of categories needs to be limited (e.g., dissonance versus consonance, major versus minor etc.); when the task is identifying a specific interval, a small choice of intervals should be presented, based on similarity (or high degree of confusion) along one dimension.

8.3.3 Sight Singing

Sight singing is one of the activities that tends to be regarded both as a means and as a goal. No doubt an important skill, real-life musical situations that require sight singing are limited to some choir rehearsals and less often, rehearsals of solo vocal pieces. As a means, though, it is a valuable way of incorporating almost all musical aspects and demonstrating pattern recognition, understanding of the tonal system and the variety of relationships implied by it.

¹⁸Cohen *et al.*, “Response to Intervals.”

Modification of Sight Singing Texts When using a specific sight singing text, the tasks can be modified and specialized to strengthen various cognitive skills such as retention of specific scale degrees, pattern recognition, major/minor discrimination, metric alterations and so forth. Figure 8.5 illustrates some strategies of modifying one melody; it should be noted, though, that different melodies may lend themselves to different modifications.

Adding the tonic in each rest (Figure 8.5b) has two main effects: the first, and obvious, is retaining the tonic in the various contexts of the melody. This can be done with other scale degrees, as well as with nondiatonic tones, wherever appropriate. In a modulatory melody, it is possible to retain the tonic relevant for each section, and also to retain the original tonic throughout the modulation. A secondary effect of inserting an extra tone is directing attention to the second beat, a non-prominent metric position which is usually not attended to.

Re-constructing a major melody in a minor mode (Figure 8.5c) directs attention to the mode more intensely than when the melody is initially familiar in that mode. Alternating between major and minor modes (Figure 8.5c) is done first by comparing repeated phrases (Figure 8.5d1), but proceeds to more difficult tasks such as starting a phrase with the new mode (Figure 8.5d2-d3) and altering the mode within a phrase (Figure 8.5d3).

a. Sight singing text

b. Adding low tonic in each rest

c. Minor mode

d. Alternating major/minor modes

e1. Upper auxiliary, triple vs. duple subdivision

e2. Lower auxiliaries and passing tones

e3. Chromatic lower auxiliary

e4. Alternating chromatic/diatonic lower auxiliaries

Source of melody: *Solfège des Solfèges*, eds. H. Lemoine and Danhauser, vol. 1A no. 24.

Figure 8.5: Modifications of a sight singing text

Numerous patterns of variations are possible, some of which are demonstrated in Figure 8.5e1-e4, providing ample opportunities for aural training. Especially important are the three variations incorporating lower auxiliary tones (Figure 8.5e2-e4) which allow students to experience the difference between a diatonic lower auxiliary and its chromatic counterpart. Such distinctions are common pit-falls in dictations, when a chromatic lower auxiliary sounds so natural that it is perceived as a diatonic tone.

Intervals versus Scale Degrees The role of intervals versus scale degrees was discussed above (Section 8.3.2). It cannot be stressed enough that if sight singing is a means of understanding tonal relationships, then these relationships should be used as a main strategy, and not intervals. A student that is “stuck” in a modulatory passage and resorts to the use of intervals strategy, may be able to complete his/her sight singing task better, and probably get a higher grade; but a student that will use this opportunity to get familiarized with that specific type of modulation and understand its tonal meaning, will probably become a better musician.

Movable and Fixed Methods

While there is an agreement that tonal sight singing is an exercise in tonal relationships, teachers strongly disagree whether the appropriate method is the use of a fixed or a movable set of syllables. The main methods currently used in the U.S. and

Europe are:¹⁹

- Fixed *do*: solfège syllables assigned to pitch classes
- Fixed A: letter names assigned to pitch classes
- Movable “one”: numbers assigned to scale degrees
- Movable *do/la minor*: solfège syllables assigned to scale degrees with *do* as the tonic of the major scale, and *la* as the tonic of its relative minor
- Movable *do/do tonic*: solfège syllables assigned to scale degrees with *do* as the tonic of both the major scale and its parallel minor scale

The methods differ also by the way flats and sharps are treated; they may be omitted during sight singing, for the sake of metric and rhythmic continuity; the vowels of solfège syllables may be altered to reflect changes (e.g. *do*, *di*, *de*); or sharps and flats may be sung, creating a two-syllables name. It would have been helpful if we could encompass in one syllable all relevant features of a tone within a tonal context: its pitch class, its scale degree, its melodic and harmonic function, and the mode of its context. Since this is not possible, the debate is about which information is more important, which features must be regarded as “same,” and which features

¹⁹For a more comprehensive list of methods and their critique, see Timothy A. Smith, “A Comparison of Pedagogical Resources in Solmization Systems,” *Journal of Music Theory Pedagogy*, 5, no. 1 (1991): 1-23.

must be differentiated. It is my belief that understanding tonal music is predominantly understanding relationships, but also that pitch classes cannot be ignored. For professional musicians, there is no way around mastering the pitch-class names of all scale degrees in every major and minor key. For relationships to be clearly understood, it is preferable not to assign any relation-implying syllables, rather than to assign same syllables to not-quite-the-same relations. In language, this would be a case of double-meaning which tends to obscure the perception of meaning.

Some surveys administered in the U.S. showed that statistically, movable methods incorporating numbers or solfège syllables are used more often than fixed methods.²⁰ On the other hand, letter names are commonly used for discussion about music and for naming pitch classes. Thus using a movable method for sight singing means assigning one language for literacy, including activities such as instrumental sight reading, references to the music during rehearsal, analysis and any discussion among musicians, while a different language is assigned to aural perception and cognition. David Butler, who has interviewed directors of aural-training programs in a few dozens universities, observed that "sight-singing with fixed *do* ... has nothing to do with developing absolute pitch ... fixed *do* is viewed as a performing musician's

²⁰R. G. Pembroke and H. Lee Riggins, "'Send Help!': Aural Skills Instruction in U. S. Colleges and Universities," *Journal of Music Theory Pedagogy*, 4, no. 2 (fall 1990): 231-241; Irma Collins, *Current Attitudes and Trends in the Teaching of Sight Singing in Higher Education* (D.M.A. in Mus. Ed. diss., Temple University, 1979), cited in: M. Houlahan and P. Tacka, "The Americanization of Solmization: A Response to the Article by Timothy A. Smith, 'A Comparison of Pedagogical Resources in Solmization Systems'," *Journal of Music Theory Pedagogy*, 8 (1994): 137-151.

Esperanto.”²¹ Although Butler refers particularly to performing musicians who will have the opportunity to perform with European conductors, it seems that even musicians within the U.S. are in need of Esperanto to be able to use one language for all musical activities.

The main criticism against the fixed methods is that the syllables do not reflect tonal relationships. That this does not prevent sight singers from perceiving tonal relationships is demonstrated by all those sight singers, mostly in Europe, but also in the U.S., who use a fixed method, do not possess absolute pitch, and probably do not constantly calculate intervals at an amazing speed. It is probably true that movable methods are faster and easier for beginners, and are especially efficient if they do not aspire to become professional musicians, in which case they need to be fluently literate. In this sense, these methods may be regarded as short cuts suitable for the short run, but for the long run they may lead to a dead end.

Movable methods are also called “relative methods” because they aspire to teach relationships while sight singing by constantly naming these relationships. But is it possible to assign syllables that will consistently describe specific tonal relationships? Scholars who promote these methods believe that it is possible, but they strongly disagree on which types of relationships are to be assigned the same syllables. An on-going discussion of this matter was carried between Houlahan and Tacka, who promote

²¹Butler, “Why the Gulf . . .,” p. 40

the *la minor* approach,²² and Smith, who promotes the *do tonic* approach.²³ Smith maintains that a tonic should always be assigned the same syllable, *do*, reflecting its stability, and the same goes for other scale degrees: a dominant is a dominant, named *so* in major or minor mode, a mediant is a mediant, and is named *mi* and so forth. Houlahan and Tacka, on the other hand, maintain that using one set of tonal hierarchy does not reflect the differences in dynamic tensions between scale degrees in major and minor modes.

Indeed, the difference between these two methods points to a major problem of the movable approach; one set of syllables is just not enough to reflect all possible, or even all common types of relationships in tonal music. Melodic dynamic tensions between scale steps are different than harmonic relationships. Thus, the A $\flat$ (marked “a”) in the soprano part of Figure 8.6 should be assigned one syllable, reflecting its melodic tendency to resolve down to G, while the A $\flat$ in the bass line (marked “b”), which definitely follows any experienced listener’s expectations by ascending to B $\flat$, should be assigned a different syllable, reflecting its harmonic function. Both, would be named *fa* in any movable solfège-based method.

The situation gets worse when interkey relations are considered, starting with the major/minor relations; one set of syllables can reflect either relative or parallel

²²M. Houlahan and P. Tacka, “The Americanization of Solmization: A Response to the Article by Timothy A. Smith, ‘A Comparison of Pedagogical Resources in Solmization Systems’,” *Journal of Music Theory Pedagogy*, 8 (1994): 137-151.

²³Smith, “A Comparison of Pedagogical Resources.”

Figure 8.6: Movable syllables assigned to melodic tendencies and harmonic functions

major/minor relations, but not both. In the same manner, any modulatory process results in the new tonic being accepted as “The” tonic, with no special syllable reflecting its being “the tonic of the dominant key,” or any other interkey relationships. This should present a problem to any musician, but especially to those accepting the Schenkerian view that there are no “true” modulations, just tonicizations.

Movable methods are also regarded as methods which promote analysis during sight singing. The fact of the matter is that analysis, at least in the sense of organization of the material, takes place with perception, though it may not be fully conscious. The highly cognitive analysis needed for assigning the correct movable syllables may be, in some cases, contrary to the way the music unfolds. A passage that starts as tonally ambiguous, for example, cannot maintain its ambiguity if scale degree syllables are to be assigned. Modulatory pivot tones have to be assigned a scale degree syllable either in the initial key or in the new key, and thus lose their pivotal character.

It seems, then, that there is a need for numerous different syllables to name the many common relationships that interplay in tonal music. Since it is not possible to manage assigning that many syllables in real time, a compromise must be made, and the compromise made by movable methods is assigning similar syllables to nonsimilar musical phenomena. Such a compromise is more detrimental to the understanding of these phenomena than using fixed syllables that are not designated to indicate relationships, and using other strategies to direct students' attention to these relationships.

In addition to avoiding a semantic chaos, using a fixed method, in the U.S. particularly the fixed A, has the benefit of promoting literacy, allowing an easier integration of cognitive skills acquired in an ear training class with performance and analysis, and allowing pitch classes to be relevant.

8.3.4 Dictation

Dictation is the other main activity that is taken both as a means and as a goal, but in fact, this particular activity is quite rare in a musician's life. Assuming composers are most likely to need to notate what they hear, they do not need to rely on their short-term memory, since hopefully, their composition is organically evolving in their minds as they notate it. As a means, though, this is an activity of great value, since it requires organization and understanding of what is heard, and its integration with

what is seen.

Regarded primarily as a means, rather than a goal, it is not necessary to complete every dictation that is started in the classroom, especially since this is a time consuming activity in which students vary widely in the amount of time they need to complete the task. If completing the dictation is a goal, then the effective use of short-term memory is the first task, but this does not necessarily have to be the case. Not underrating the value of the expansion of short-term memory, a dictation activity may also be used for temporal organization of the material, and for a controlled direction of attention. By the time these tasks, which do not require notating the complete dictation, are completed, the dictation has been played several times and the students are likely to have it memorized.

Short-Term Memory If dictation taking is expected to expand the students' short-term memory, its current capacity needs to be assessed and exceeded. It should be pointed out that some activities related to dictation taking are likely to interfere with memory. Since pitch memory was shown to work on a single channel (see Section 2.1) and deteriorate with the interference of other pitch information, any pitch-related activity may interfere with short-term memory of the dictated material. Such interferences may be playing the dictation while students focus on a specific section of it, other than the one played at that moment; recalls of the dictation by singing,

in cases that this strategy is recommended by the teacher; and even mental activities such as audiating sections of the dictation.

Pembrook has observed that while correct recall (by singing) of a dictation improves its memory (resembling the phenomenon of consolidation), incorrect recalls worsen it.²⁴ This is also supported by the results of Sloboda and Parker's recall experiment in which incorrect recalls were maintained through repeated hearings of the original stimulus.²⁵ For this matter, transposed recalls are "incorrect recalls," since they may act as pitch interferences. This is one of those cases in which pitch retention is a helpful skill: students whose recalls are not in the original key, are faced with the task of establishing tonality with each recall, while students who retain the pitch of the original tonic, have a more stable sense of the tonality.

Temporal Organization Since time is music's space, notation is impossible without a specific location in that space. Thus, temporal organization must happen, at least at some level, before notation, and may serve as a template on which the rest of the information is added. Such a template is often provided by teachers and in work books, but such "help" allows students to skip the stage of temporal organization, and they may end up being lost in the space of time. Directing attention to features such as meter, symmetry or asymmetry in the number of phrases and in phrase lengths,

²⁴R. G. Pembrook, "The Effect of Vocalization on Melodic Memory Conservation," *Journal of Research in Music Education*, 35 (1987): 155-169.

²⁵Sloboda and Parker, "Immediate Recall of Melodies."

the actual number of phrases, repetition of phrases and number of measures in each phrase, may assist in creating such template (see Figure 8.7b). This type of activity may also incorporate a “self dictation” of familiar tunes. Eliminating the task of short-term memory may help focus on creating the temporal template.

Directed Attention

Attention may be directed to other features, in the same manner that it was directed to temporal organization. Different students will spontaneously attend to different features of the dictation at different times and in different order, but taking dictation is an opportunity to gain control over the direction of attention. Initially, attention will be directed to the most perceptually stable and structurally important features, which will facilitate memorization, understanding and notation. At a more advanced stage, to improve control of attention, it may be intentionally directed to the less stable features.

Perceptually stable features, which facilitate organization and understanding, are: (1) pitches at prominent metric positions and temporal locations (i.e., down beats and beats, beginnings and endings of phrases etc.); (2) tonally prominent features such as mode, the pitches of the tonic, the dominant and the leading tone, highest and lowest tones; (3) specific characteristics of the specific dictation. Less stable features are non-diatonic tones, unaccented beats, etc.

Figure 8.7 illustrates some of the features extraction processes mentioned above. In directing students' attention it is helpful, and musically meaningful, to see the big picture before attending to details. In the process of creating the temporal template, attention is directed first to the larger structure (e.g., number of phrases and repeated phrases). Functional direction of the phrases may be discussed (e.g., tonic to dominant and dominant to dominant). It is likely that when attending to repetitions, some students will notice the sequence in phrase B (Figure 8.7b). A special feature of this melody, that may be pointed out to the students at this stage, is that although the phrases are of equal length, their arrangement is asymmetrical. The six phrases (rather than the more common four or eight), do not adhere to the expected groups of two phrases which may be implied by the repetition of phrase A; phrase C, which is expected to be coupled with phrase B, is reminiscent of phrase A, and phrase D, which picks up the new rhythmic material introduced in phrase C, may be heard as an ending, only to be followed by another ending. Encouraging students to make these types of observations assists in organizing the time space. In discussing the meter, it is the students' task to determine whether the meter is duple or triple, simple or compound. The use of duple versus quadruple meter, as well as the choice of the beat value, can be agreed upon after the meter is perceived.

a. Dictation melody

b. Temporal template

c. Pitches at down beats

d. Pitches at 1st and 3rd beats

e. Tonic pitches, location of nondiatonic tone (x)

Source of melody: *The Oxford Folk Song Sight Singing Series*, book V, p. 16.

Figure 8.7: Features extraction in taking a dictation

Typically, notating the down beat pitches (Figure 8.7c) reveals some of the tonal structure of the melody. Less typical, in this particular melody, is the fact that taken in $\frac{4}{4}$ time, the tonic appears only on the final down beat; other occurrences of the tonic are on metrically less stable positions. Adding the pitches on the 3rd beats adds two more occurrences of the tonic (Figure 8.7d). It is likely that the task illustrated in Figure 8.7e — notating all occurrences of the tonic will be more difficult, since even the tonic, when occurs outside a prominent metrical position, tends to be perceived as unstable.

In other melodies, it may be helpful to direct attention to tonic occurrences first: even if only those on prominent metrical positions are perceived, students performing this task are likely to be audiating the tonic while listening to the various relationships occurring in the specific melody, and thus improve its stability in short-term memory.

8.4 Summary

Research of music perception and cognition focuses mostly on materials that can be controlled and experimented. It is intrinsic to the methodology of psychomusical and acoustical experiments to isolate components in order to be able to draw empirical evidence. Ear training for tonal music, as was shown, is the study of relationships and meaning, rather than components. Unfortunately, it is the large scale relations that present more questions, and these cannot be given an answer based on empirical

evidence.

However, although many of the findings cited in this theses are relevant and useful for designing effective approaches to developing musical cognitive skills, they have not yet become common knowledge for teachers of ear training. This unfortunate situation may be due, in part, to the level of accessibility of this information. Since each experiment is focused on a single aspect of a single component, its contribution is limited. Moreover, many of the relevant publications are published under various disciplines such as psychology, acoustics and even physics. Musicians are not very likely to search for information in such periodicals, and if they do, the style in which data are presented may give the impression that the information is irrelevant; musicians tend to find little interest in the number of vibrations per second of a specific tone, or in the number of milliseconds of its durations. They would rather discuss pitch and rhythm.

It is this unfortunate alienation between the two fields, the research of music perception and cognition on one hand, and teaching ear training on the other, that was the impetus for this thesis. It is my hope that this thesis has demonstrated the way these fields can benefit from each other.

Bibliography

- Abraham, O. "Das absolute Tonebewusstsein." *Sammelbde. International Musikges*, 3 (1901): 1-86. Cited in: Fraisse, Paul. "Rhythm and Tempo." In *The psychology of music*. Ed. D. Deutsch. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- Aiello, R., ed. (With John Sloboda). *Musical Perceptions*. New York: Oxford University Press, 1994.
- Aldwell, Edward and Schachter Carl. *Harmony and Voice Leading*. 2nd ed. New York: Harncourt Brace Jovanovich, Publishers, 1989.
- Allport, D. A. "Attention and Performance." In *Cognitive Psychology: New Directions*. Ed. Guy Claxton. London: Routledge and Kegan Paul, 1980.
- Amos, T. E. *Perceiving Musical Rhythms in Different Sense Modalities*. Ph.D. diss., University of California, San Diego, 1987.
- Bachem, A. "Absolute Pitch." *Journal of the Acoustical Society of America*, 27 (1955): 1180-1185.
- . "Time Factors in Relative and Absolute Pitch Determination." *Journal of the Acoustical Society of America*, 26 (1954): 751-753.
- Balzano, G. J. "Absolute Pitch and Pure Tone Identification." *Journal of the Acoustical Society of America*, 75 (1984): 623-625.
- Beal, A. L. "The Skill of Recognizing Musical Structures." *Memory and Cognition*,

13 (1985): 405-412.

Bever, T. and R. J. Chiariello. "Cerebral Dominance in Musicians and Nonmusicians." *Science*, 185 (1974): 537-539.

Bharucha, J. "Melodic Anchoring." *Music Perception*, 13 (1996): 383-400.

———. "Music Cognition and Perceptual Facilitation: A Connectionist Framework." *Music Perception*, 5 (1987) 1-30.

Bharucha, J. and K. Stoeckig. "Priming of Chords: Spreading Activation or Overlapping Frequency Spectra?" *Journal of Experimental Psychology: Human Perception and Performance*, 14, no. 6 (1987): 519-524.

Blackwell, H. R. and H. Schlosberg. "Octave Generalization, Pitch Discrimination, and Loudness Thresholds in the White Rat." *Journal of Experimental Psychology*, 33 (1943): 407-419.

Blombach, Ann K. and Regena T. Parrish. "Acquiring Aural Interval Identification Skills: Random vs. Ordered Grouping." *Journal of Music Theory Pedagogy*, 2, no. 1 (spring 1988): 113-132.

Brady, P. "Fixed-scale Mechanism of Absolute Pitch." *Journal of the Acoustical Society of America*, 48 (1970): 883-887.

———. "Teach Yourself Absolute Pitch in 365 Hard Lessons." *Journal of the Acoustical Society of America*, 48 (1970): 88.

Bregman, A. S. *Auditory Scene Analysis*. Cambridge, Massachusetts: MIT Press, 1990.

———. "Primary Auditory Stream Segregation and the Perception of Tunes." Unpublished manuscript. Department of Psychology, McGill University, Montreal, 1971. Cited in: W. J. Dowling. "Perception of Interleaved Melodies." *Cognitive Psychology*, 5 (1973): 322-337.

Broadbent, D. E. *Decision and Stress*. London: Academic Press, 1971.

———. *Perception and Communication*. London: Pergamon, 1958.

- Brower, Candace. "Memory and the Perception of Rhythm." *Music Theory Spectrum*, 15, no. 1 (1993): 19-35.
- Brown, Helen. "The Interplay of Set Content and Temporal Context in a Functional Theory of Tonality Perception." *Music Perception*, 5 (1988): 219-249.
- Brown, Helen and David Butler. "Diatonic Trichords as Minimal Tonal Cue cells." *In Theory Only*, 5 (1981): 39-55.
- Brown, Helen, David Butler and Mari Riess Jones. "Musical and Temporal Influences on Key Discovery." *Music Perception*, 11 (1994): 371-407.
- Browne, R. "Tonal Implications of the Diatonic Set." *In Theory Only*, 5 (1981): 3-21.
- Bruner, Jerome S. "The Act of Discovery." *Harvard Educational Review*, 31 (1961): 21-32.
- Buffardi, L. "Factors Affecting the Filled-duration Illusion in the Auditory, Tactual, and Visual Modalities." *Perceptions and Psychophysics*, 7 (1971): 36-45.
- Burns, E. M. and W. D. Ward. "Categorical Perception — Phenomenon or Epiphenomenon: Evidence from Experiments in the Perception of Melodic Musical Intervals." *Journal of the Acoustical Society of America*, 63 (1978): 456-468.
- Butler, David. "Describing the Perception of Tonality in Music: A Critique of the Tonal Hierarchy Theory and a Proposal for a Theory of Intervallic Rivalry." *Music Perception*, 6 (1989): 219-242.
- . *The Musician's Guide to Perception and Cognition*. New York: Schirmer Books, 1992.
- . "On Pitch-set Properties and Perceptual Attributes of the Minor Mode." In *Cognitive bases of musical communication*. Eds. R. Jones and S. Holleran. Washington, DC: American Psychological Association, 1992.
- . "Response to Carol Krumhansl." *Music Perception*, 7 (1990): 325-338.
- . "Why the Gulf Between Music Perception Research and Aural Training?" By

- R. E. Radocy (A D. Butler presentation). *Bulletin of the Council for Research in Music Education*, 132 (spring 1997): 39-50.
- Butler, David and Helen Brown. "Describing the Mental Representation of Tonality in Music." In *Musical Perceptions*. Ed. R. Aiello. New York: Oxford University Press, 1994.
- Carroll, J. B. "Speed and Accuracy of Absolute Pitch Judgments: Some Latter-day Results." Research Bulletin No. ETS-RB-75-35. Princeton, New Jersey: Educational Testing Service (ERIC Document Reproduction service No. ED 150 153). n.d.
- Carterette, E. C., C. B. Monahan, E. Holman, T. S. Bell and R. A. Fiske. "Rhythmic and Melodic Structures in Perceptual Space." *Journal of the Acoustical Society of America* 72, Supplement 1 (1982): S11.
- Cazden, N. "The Definition of Consonance and Dissonance." *International Review of Aesthetics and Sociology of Music*, 11 (1980): 123-168.
- Cherry, E. C. "Some Experiments on the Recognition of Speech, with One and Two Ears." *Journal of the Acoustical Society of America*, 25 (1953): 975-979.
- Clark, M. and P. Milner. "Dependence of Timbre on the Tonal Loudness Produced by Musical Instruments." *Journal of the Audio Engineering Society*, 12 (1964): 28-31.
- Clarke, E. F. "Structure and Expression in Rhythmic Performance." In *Musical Structure and Cognition*. Eds. P. Howell, I. Cross and R. West. New York: Academic Press, 1985.
- Clarke, E. F. and C. L. Krumhansl. "Perceiving Musical Time." *Music Perception*, 7 (1990): 213-251.
- Clynes, Manfred, ed. *Music, Mind and Brain*. New York: Plenum Press, 1982.
- Cohen, A. J. and K. Baird. "Acquisition of Absolute Pitch: The Question of Critical Periods." *Psychomusicology*, 9 (1990): 31-37.
- Cohen, D., J. Mendel, H. Prat and A. Barnea. "Response to Intervals as Revealed by

- Brainwave Measurement and Verbal Means." *Journal of New Music Research*, 23, no. 3 (1994): 265-290.
- Cohen, Morris R. *A Preface to Logic*. New York: Henry Holt & Co., 1944.
- Collins, Irma. *Current Attitudes and Trends in the Teaching of Sight Singing in Higher Education*. D.M.A. in Mus. Ed. diss. Temple University, 1979. Cited in: M. Houlahan and P. Tacka. "The Americanization of Solmization: A Response to the Article by Timothy A. Smith, 'A Comparison of Pedagogical Resources in Solmization Systems'." *Journal of Music Theory Pedagogy*, 8 (1994): 137-151.
- Cook, K. E. *A New Approach to Musical Meter: Contributions Towards a Theory of Metric Perception in Music*. Ph.D. diss. Michigan State University, 1991.
- Cooper, Grosvenor W. and Leonard B. Meyer. *The Rhythmic Structure of Music*. Chicago: University of Chicago Press, 1960.
- Corliss, E. L. "Remark on 'Fixed-scale Mechanism of Absolute Pitch'." *Journal of the Acoustical Society of America*, 53 (1973): 1737-1739.
- Crowder, R. G. "The Role of One's Own Voice in Immediate Memory." *Cognitive Psychology*, 1 (1970): 157-178.
- Crowe, Edger, Annie Lawton and W. Gillies Whittaker, eds. *The Folk Song Sight Singing Series*, Book V. London: Oxford University Press, 1933.
- Cuddy, L. L. "Practice Effects in the Absolute Judgment of Pitch." *Journal of the Acoustical Society of America*, 43 (1968): 1069-1076.
- . "Training the Absolute Identification of Pitch." *Perception and Psychophysics*, 8 (1970): 256-259.
- Cuddy, L. L. and B. Badertscher. "Recovery of the Tonal Hierarchy: Some Comparisons Across Age and Levels of Musical Experience." *Perception and Psychophysics*, 41 (1987): 609-620.
- Cuddy, L. L. and A. J. Cohen. "Recognition of Transposed Melodic Sequences." *Quarterly Journal of Experimental Psychology*, 28 (1976): 255-270.

- Cuddy, L. L., A. J. Cohen and J. Miller. "Melody Recognition: The Experimental Application of Rules." *Canadian Journal of Psychology*, 33 (1971): 148-157.
- Demany, Laurent and Catherine Semal. "Pitch Versus Brightness of Timbre." *Music Perception*, 11 (1993): 1-14.
- Deutsch, D. "An Auditory Illusion." *Journal of the Acoustical Society of America*, 55 (1974): 518-519.
- . "The Deterioration of Pitch Response in Memory." Unpublished doctoral diss. University of California, San Diego, 1970. Cited in: D. Deutsch. "The Processing of Pitch Combinations: Memory Systems." In *The Psychology of Music*. Ed. D. Deutsch. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- . "Mapping of Interactions in the Pitch Memory Store." *Science*, 175 (1972): 1020-1022.
- . "Memory and Attention in Music." In *Music and the Brain*. Eds. M. Critchley and R. Henson. London: Heinemann, 1977.
- . "Musical Illusions." *Scientific American*, 233, no. 4 (1975): 92-104.
- . "Octave Generalization and Tune Recognition." *Perception and Psychophysics*, 11 (1972b): 411-412.
- . "The Processing of Pitch Combinations: Memory Systems." In *The Psychology of Music*. Ed. D. Deutsch. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- . "Tones and Numbers: Specificity of Interference in Short-term Memory." *Science*, 168 (1970): 1604-1605.
- , ed. *The Psychology of Music*. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- Deutsch, D. and P. L. Roll. "Separate 'What' and 'Where' Decision Mechanisms in Processing a Dichotic Tonal Sequence." *Journal of Experimental Psychology: Human Perception and Performance*, 2 (1976): 23-29.

- Dowling, W. J. "Perception of Interleaved Melodies." *Cognitive Psychology*, 5 (1973): 322-337.
- . "Scale and Contour: Two Components of a Theory of Memory for Melodies." *Psychology Review*, 85 (1978): 341-354.
- Dowling, W. J. and D. S. Fujitani. "Contour, Interval, and Pitch Recognition in Memory for Melodies." *Journal of the Acoustical Society of America*, 49 (1971): 524-531.
- Dowling, W. J. and Dane L. Harwood. *Music Cognition*. Academic Press Series in Cognition and Perception. New York: Academic Press, 1986.
- Dowling, W. J. and Ava W. Hollombe. "The Perception of Melodies Distorted by Splitting into Several Octaves: Effects of Increasing Proximity and Melodic Contour." *Perception and Psychophysics*, 21, no. 1 (1977): 60-64.
- Dowling, W. J., K. M-T. Lung and S. Herrbold. "Aiming Attention in Pitch and Time in the Perception of Interleaved Melodies." *Perception and Psychophysics*, 41, no. 6 (1987): 642-656.
- Edworthy, J. "Towards a Pitch-contour Continuum Theory of Memory for Melodies." In *Acquisition of Symbolic Skills*. Eds. J. Sloboda and D. Rogers. New York: Plenum Press, 1983.
- Forte, Allen. *Tonal harmony in Concept and Practice*. 3rd ed. New York: Holt, Rinehart & Winston, 1979.
- Fraisse, Paul. "Rhythm and Tempo." In *The psychology of music*. Ed. D. Deutsch. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- Francès, R. *The Perception of Music*. Trans. W. J. Dowling. Hillsdale, New Jersey: L. Erlbaum, 1988.
- Fuller, David. "Accent." *The New Harvard Dictionary of Music*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1986.
- . "Beat." *The New Harvard Dictionary of Music*. Cambridge, Massachusetts:

The Belknap Press of Harvard University Press, 1986.

Gardner, A. D. and R. W. Pickford. "Relation Between Dissonance and Context." *Nature*, 152 (1943): 358.

Garner, W. R. "Good Patterns Have Few Alternatives." *American Scientist*, 58 (1970): 34-42.

———. *The Processing of Information and Structure*. Potomac, Maryland: Erlbaum, 1974.

Gordon, Edwin E. "Research Studies in Audiation." *Bulletin of the Council for Research in Music Education*, (1984): 34.

Grau, James W. and Deborah G. Kemler-Nelson. "The Distinction Between Integral and Separable Dimensions: Evidence for the Integrality of Pitch and Loudness." *Journal of Experimental Psychology: General*, 117 (1988): 347-370.

Graybill, Roger. "Pedagogically Speaking: Consonance and Dissonance in the Freshman Theory Class." *In Theory Only*, 9, no. 5-6 (1987): 51-56.

Grebel'nik, S. G. "Formirovanie u Doshkol'nikov Absoliutnovo Muzykal'no Slukha [The Formation of Absolute Pitch in Pre-school Children]." *Voprosy Psikhologii*, 2 (March-April 1984): 90-98. Cited in: M. Rush. *An Experimental Investigation of the Effectiveness of Training on Absolute Pitch in Adult Musicians*. Ph.D. diss. Ohio State University, 1989.

Green, D. M. and J. A. Swets. *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966.

Gregory, A. H. "Listening to Polyphonic Music." *Psychology of Music*, 18 (1990): 163-170.

———. "Timbre and Auditory Streaming." *Music Perception*, 12, no. 2 (winter 1994): 161-174.

Guernsey, M. "The Rôle of Consonance and Dissonance in Music." *American Journal of Psychology*, 40 (1928): 173-204.

- Handel, S. and G. R. Lawson. "The Contextual Nature of Rhythmic Interpretation." *Perception and Psychophysics*, 34 (1983): 103-120.
- Hantz, Edwin C., G. C. Crummer, J. W. Wayman, J. P. Walton and R. D. Frisina. "Effects of Musical Training and Absolute Pitch on the Neural Processing of Melodic Intervals: A P3 Event-related Potential Study." *Music Perception*, 10 (1992): 25-41.
- Hartmann, W. M. and D. Johnson. "Stream Segregation and Peripheral Channeling." *Music Perception*, 9 (1991): 155-184.
- Helmholtz, H. L. F. von. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Trans. A. Ellis. New York: Dover, 1954 (Original published as *Die Lehre von der Tonempfindungen als physiologische Grundlage für die Theorie der Music*, Braunschweig, 1863).
- Hindemith, P. *Elementary Training for Musicians*. New York: Associated Music Publishing, 1946.
- Houlahan, M. and P. Tacka. "The Americanization of Solmization: A Response to the Article by Timothy A. Smith. 'A Comparison of Pedagogical Resources in Solmization Systems'." *Journal of Music Theory Pedagogy*, 8 (1994): 137-151.
- Howard, David, Stuart Rosen and Victoria Broad. "Major/minor Triad Identification and Discrimination by Musically Trained and Untrained Listeners." *Music Perception*, 10 (1992): 205-220.
- Howell, Peter, Ian Cross and Robert West, eds. *Musical Structure and Cognition*. New York: Academic Press, 1985.
- Hubbard, T. and K. Stoeckig. "Musical Imagery: Generation of Tones and Chords." *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, no. 4 (1988): 656-667.
- Idson, W. L. and D. W. Massaro. "Perceptual Processing and Experience of Auditory Duration." *Sensory Processes*, 1 (1977): 316-337.
- James, William. *The Principles of Psychology*. New York: Holt, 1890.

- Jones, M., M. Boltz and G. Kidd. "Controlled Attending as a Function of Melodic and Temporal Context." *Perception and Psychophysics*, 32 (1982): 211-218.
- Jordan, D. S. "Influence of the Diatonic Tonal Hierarchy at Microtonal Intervals." *Perception and Psychophysics*, 41 (1987): 482-488.
- Klein, M., M. G. H. Coles and E. Donchin. "People with Absolute Pitch Process Tones Without Producing a P300." *Science*, 223 (1984): 1306-1309.
- Koffka, K. *Principles of Gestalt Psychology*. 4th ed. First published in England in 1935. London: Routledge & Kegan Paul, 1955.
- Komar Arthur. *Theory of Suspensions*. Princeton: Princeton University Press, 1971.
- Kramer, J. D. *The Time of Music: New Meanings, New Temporalities, New Listening Strategies*. New York: Schirmer, 1988.
- Kristofferson, A. B. "A Quantal Step Function in Duration Discrimination." *Perception and Psychophysics*, 27 (1980): 300-306.
- Krumhansl, C. L. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press, 1990.
- . "The Psychological Representation of Musical Pitch in a Tonal Context." *Cognitive Psychology*, 11 (1979): 346-374.
- . "Tonal Hierarchies and Rare Intervals in Music Cognition." *Music Perception*, 7 (1990): 309-324.
- Krumhansl, C. L. and Paul Iverson. "Perceptual Interactions Between Musical Pitch and Timbre." *Journal of Experimental Psychology: Human Perception and Performance*, 18, no. 3 (1992): 739-751.
- Krumhansl, C. L. and E. J. Kessler. "Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys." *Psychological Review*, 89 (1982): 334-368.
- Krumhansl, C. L. and R. N. Shepard. "Quantification of the Hierarchy of Tonal Functions within a Diatonic Context." *Journal of Experimental Psychology:*

- Human Perception and Performance*, 5 (1979): 579-594.
- Langer, Susanne. *Feeling and Form*. New York: Scribner's, 1953.
- Larson, Steve. "Scale-degree Function: A theory of Expressive Meaning and its Application to Aural-skill Pedagogy." *Journal of Music Theory Pedagogy*, 7 (1993): 69-84.
- Lemoine, L. and Danhauser A. *Solfège des Solfèges*, Vol. 1A. Paris: Éditions Henry Lemoine, n.d.
- Lerdahl, F. "Timbral Hierarchies." *Contemporary Music Review*, 2 (1987): 135-160.
- Lerdahl, F. and R. Jackendoff. *A Generative Theory of Tonal Music*. Cambridge, Massachusetts: MIT Press, 1983.
- Lester J. *The Rhythms of Tonal Music*. Carbondale: Southern Illinois University Press, 1986.
- Levelt, W. J. M., J. P. van de Geer and R. Plomp. "Triadic Comparisons of Musical Intervals." *British Journal of Mathematical & Statistical Psychology*, 19 (1966): 163-179.
- Lockhead, G. R. and R. Byrd. "Practically Perfect Pitch." *Journal of the Acoustical Society of America*, 70 (1981): 387-389.
- Long, Peggy A. "Relationships Between Pitch Memory in Short Melodies, and Selected Factors." *Journal of Research in Music Education*, 25, no. 4 (1977): 272-282.
- Lundin, R. W. "Can Perfect Pitch Be Learned?" *Music Education Journal*, (1963): 49-51.
- MacDougall, R. "The Structure of Simple Rhythm Forms." *Psychological Review* (Monograph Supplements), 4 (1903): 309-416.
- Melara, R. D. and L. E. Marks. "Perceptual Primacy of Dimensions: Support for a Model of Dimensional Interaction." *Journal of Experimental Psychology: Human Perception and Performance*, 16 (1990): 398-414.

- Meumann, E. "Untersuchungen zur Psychologie und Aesthetic des Rhythmus." *Philosophische Studien*, 10 (1894): 249-322, 393-430. Cited in: Paul Fraisse. "Rhythm and Tempo." In *The psychology of music*. Ed. D. Deutsch. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- Meyer, Leonard B. *Emotion and Meaning in Music*. Chicago: University of Chicago Press, 1956.
- Meyer, M. "Is the Memory of Absolute Pitch Capable of Development by Training?" *Psychological Review*, 6 (1899): 514-516.
- Miller, G. A. "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information." *Psychological Review*, 63 (1956): 81-97.
- Miller, G. A. and G. A. Heise. "The Trill Threshold." *Journal of the Acoustical Society of America*, 22 (1950): 637-638.
- Miller, J., C. Wier, R. Pastore, W. Kelly and R. Dooling. "Discrimination and Labeling of Noise-buzz Sequences with Varying Noise-level Times: An Example of Categorical Perception." *Journal of the Acoustical Society of America*, 60 (1976): 410.
- Minsky, Marvin. "Music, Mind and Meaning." In *Music, Mind and Brain: The Neuropsychology of Music*. Ed. Manfred Clynes. New York: Plenum Press, 1982.
- Mishima, J. *Introduction to the Morphology of Human Behavior. The Experimental Study of Mental Tempo*. Tokyo: Tokyo Publishing, 1965. Cited in: Paul Fraisse. "Rhythm and Tempo." In *The psychology of music*. Ed. D. Deutsch. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- Miyazaki, K. "Absolute Pitch Identification: Effects of Timbre and Pitch Region." *Music Perception*, 7 (1989): 1-14.
- . "Musical Pitch Identification by Absolute Pitch Possessors." *Perception and Psychophysics*, 44 (1988): 501-512.
- . "Perception of Musical Intervals by Absolute Pitch Possessors." *Music Perception*, 9 (1992): 413-426.

- . "The Speed of Musical Pitch Identification by Absolute Pitch Possessors." *Music Perception*, 8 (1990): 177-188.
- Monahan, C. B. "Parallels Between Pitch and Time: The Determinants of Musical Space." Ph.D. Diss. University of California, Los Angeles, 1984.
- Mull, H. K. "The Acquisition of Absolute Pitch." *American Journal of Psychology*, 36 (1925): 469-493.
- Murphy, L. E. "Absolute Judgments of Duration." *Journal of Experimental Psychology*, 71 (1966): 260-263.
- Neisser, Ulric. *Cognition and Reality: Principles and Implications of Cognitive Psychology*. San Francisco: W. H. Freeman, 1976.
- Ornstein, R. E. *On the Experience of Time*. Harmondworth, Middlesex: Penguin Books, 1969.
- Ortmann, Otto. "Problems in the Elements of Ear-dictation." *Research Studies in Music*, 2 (October 1934): 1-95.
- Parncutt, R. *Harmony: A Psychoacoustical Approach*. Springer Series in Information Sciences. New York: Springer-Verlag, 1989.
- . "Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord." *Music Perception*, 6 (1988): 65-94.
- Paulson, D. C., R. Orlando and A. M. Schoelkopf. "Experimental Analysis and Manipulation of Auditory Generalization in Three Developmental Retardates by Discriminated-operant Procedures." *IMRID Papers and Reports*, 4, no. 13 (1967). Nashville, Tennessee: Institute on Mental Retardation and Intellectual Development, George Peabody College for Teachers.
- Pechmann, Thomas. "Memory for Chords: The Retention of Pitch and Mode." *Music Perception*, 16 (1998): 43-54.
- Pembroke, R. G. "The Effect of Vocalization on Melodic Memory Conservation." *Journal of Research in Music Education*, 35 (1987): 155-169.

- Pembroke, R. G. and H. Lee Riggins. "'Send Help!': Aural Skills Instruction in U. S. Colleges and Universities." *Journal of Music Theory Pedagogy*, 4, no. 2 (fall 1990): 231-241.
- Peterson, J. and F. Smith. "The Range and Modifiability of Consonance in Certain Musical Intervals." *American Journal of Psychology*, 4 (1930): 561-572.
- Pistou, W. *Harmony*. 4th ed. New York: Norton, 1978.
- Plomp, R. *Aspects of Tone Sensation: A Psychophysical Study*. Academic Press Series in Cognition and Perception. New York: Academic Press, 1976.
- Plomp R. and W. J. M. Levelt. "Tonal Consonance and Critical Bandwidth." *Journal of the Acoustical Society of America*, 38 (1965): 548-560.
- Plomp, R., W. A. Wagenaar and A. M. Mimpen. "Musical Interval Recognition with Simultaneous Tones." *Acustica*, 29 (1973): 101-109.
- Plotkin, L. "Stimulus Generalization in Morse Code Learning." *Archives of Psychology*, no. 287 (1943): n.p.
- Povel, Dirk-Jan. "Internal Representations of Simple Temporal Patterns." *Journal of Experimental Psychology: Human Perception and Performance*, 7 (1981): 3-18.
- Powers, Harold S. "Rhythm." *The New Harvard Dictionary of Music*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1986.
- Preusser, D. "The Effect of Structure and Rate on the Recognition and Description of Auditory Temporal Patterns." *Perception and Psychophysics*, 11 (1972): 233-240.
- Preusser, D., W. R. Garner and R. L. Gottwald. "The Effect of Starting Pattern on Descriptions of Perceived Temporal Patterns." *Psychonomic Science*, 21 (1970): 219-220.
- Rakowski, A. "Intonation Variants of Musical Intervals in Isolation and in Musical Contexts." *Psychology of Music*, 18 (1990) 60-72.
- Randel, Don Michael. "Meter." *The New Harvard Dictionary of Music*. Cambridge,

- Massachusetts: The Belknap Press of Harvard University Press, 1986.
- Riemann, Hugo. "Die Natur der Harmonik." *Sammlung musikalischer Vorträge*, 40 (1882). Cited in: James Tenney. *A History of 'Consonance' and 'Dissonance.'* New York: Excelsior Press, 1988.
- Risset, J. C. and M. V. Mathews. "Analysis of Musical Instrument Tones." *Physics Today*, 22, no. 2 (1969): 23-30.
- Roberts, L. A. "Consonance and Dissonance Judgments of Musical Chords." Paper presented at the 105th meeting of the Acoustical Society of America, 1983. Cited in: L. A. Roberts and M. L. Shaw. "Perceived Structure of Triads." *Music Perception*, 2 (1984): 95-124.
- Roberts, L. A. and M. L. Shaw. "Perceived Structure of Triads." *Music Perception*, 2 (1984): 95-124.
- Rosch, E. "Cognitive Reference Points." *Cognitive Psychology*, 7 (1975): 532-547.
- Rothstein, W. *Phrase Rhythm in Tonal Music*. New York: Schirmer Books, 1989.
- Royer, F. L. and W. R. Garner. "Perceptual Organization of Nine Element Auditory Temporal Patterns." *Perception and Psychophysics*, 7 (1970): 115-120.
- . "Response Uncertainty and Perceptual Difficulty of Auditory Temporal Patterns." *Perception and Psychophysics*, 1 (1966): 41-47.
- Rush, M. *An Experimental Investigation of the Effectiveness of Training on Absolute Pitch in Adult Musicians*. Ph.D. diss. Ohio State University, 1989.
- Schachter, Carl. "Rhythm and Linear Analysis: Aspects of Meter." *The Music Forum*, 6, no. 1 (1987) 1-59.
- Scharf, B. "Critical Bands." In *Foundations of Modern Auditory Theory*. Ed. J. V. Tobias. New York: Academic Press, 1970.
- Schoenberg, Arnold. *Harmonielehre*. Vienna: Universal, 1911.
- . *Summer Morning by a Lake*. (Five Orchestral Pieces Op. 16 No. 1) New

- York: C. F. Peters, 1973.
- Schouten, J. F. "On the Perception of Sound and Speech; Subjective Time Analysis." 4th International Congress on Acoustics, Copenhagen Congress Report II. (1962): 201-203. Cited in: D. Deutsch. "Memory and Attention in Music." In *Music and the Brain*. Eds. M. Critchley and R. Henson. London: Heinemann, 1977.
- Sergeant, D. "Experimental Investigation of Absolute Pitch." *Journal of Research in Music Education*, 17 (1969): 135-143.
- Siegel, J. A. and W. Siegel. "Absolute Identification of Notes and Intervals by Musicians." *Perception and Psychophysics*, 21 (1977): 143-152.
- Simon, H. A. and R. K. Summer. "Pattern in Music." In *Formal Representation of Human Judgment*. Ed. B. Kleinmuntz. New York: John Wiley, 1968.
- Singh, Punita G. and Ira J. Hirsh. "Influence of Spectral Locus and Fo Changes on the Pitch and Timbre of Complex Tones." *Journal of the Acoustical Society of America*, 92 (1992): 2650-2661.
- Slawson, W. *Sound Color*. Berkeley, California: University of California Press, 1985.
- Sloboda, J. A. *The Musical Mind: The Cognitive Psychology of Music*. New York: Oxford University Press, 1985.
- Sloboda, J. A., and D. H. H. Parker. "Immediate Recall of Melodies." In *Musical Structure and Cognition*. Eds. P. Howell, I. Cross and R. West. London: Academic Press, 1985.
- Smith, J., S. Hausfield, R. P. Power and A. Gorta. "Ambiguous Musical Figures and Auditory Streaming." *Perception and Psychophysics*, 32 (1982): 454-464.
- Smith, Timothy A. "A Comparison of Pedagogical Resources in Solmization Systems." *Journal of Music Theory Pedagogy*, 5, no. 1 (1991): 1-23.
- Speer, J. R. and W. E. Adams. "The Effects of Musical Training Upon the Development of the Perception of Musical Pitch." Paper presented at the meetings of the Society for Research in Child Development, 1985. Cited in: C. L. Krumhansl.

- Cognitive Foundations of Musical Pitch*. New York: Oxford University Press, 1990.
- Spender, Natasha. "Absolute Pitch." *The New Grove Dictionary of Music and Musicians*. New York: Macmillan, 1980.
- Stevens, S. S. and M. Guirao. "Subjective Scaling of Length and Area and the Matching of Length to Loudness and Brightness." *Journal of Experimental Psychology*, 66 (1963): 177-186.
- Strader, C. R. *The Effects of Variations of Duration on Musicians' Perceptions of Tonic in Tonally Ambiguous 3 Notes Set*. Ph.D. diss. Ohio State University, 1996.
- Stuckey, R. "Music in Cyclic Twelfths." Unpublished manuscript. London, 1983. Cited in: R. Parncutt. "Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord." *Music Perception*, 6 (1988): 65-94
- Studdert-Kennedy, M., A. M. Lieberman, K. Harris and F. S. Cooper. "The Motor Theory of Speech Perception: A Reply to Lane's Critical Review." *Psychological Review*, 77 (1970): 234-249.
- Studdert-Kennedy, M. and D. Shankweiler. "Hemispheric Specialization for Speech Perception." *Journal of the Acoustical Society of America*, 48 (1970): 579-594.
- Stumpf, C. "Konsonanz und Konkordanz." *Beiträge zur Akustik und Musikwissenschaft*, 6 (1911): 116-150. Cited in: N. Cazden. "The Definition of Consonance and Dissonance." *International Review of Aesthetics and Sociology of Music*, 11 (1980): 123-168.
- . *Tonpsychologie: Vol. I* [The Psychology of Music: Vol. I]. Leipzig: S. Hirzel Verlag, 1883.
- Taylor, Jack A. "Perception of Melodic Intervals Within Melodic Context." Unpublished doctoral dissertation. University of Washington, 1971. Cited in: Peggy A. Long. "Relationships Between Pitch Memory in Short Melodies, and Selected Factors." *Journal of Research in Music Education*, 25, no. 4 (1977): 272-282.
- Telesco, Paula. "Contextual Ear Training." *Journal of Music Theory Pedagogy*, 5,

no. 2 (fall 1991): 179-190.

Tenney, James. *A History of 'Consonance' and 'Dissonance.'* New York: Excelsior Press, 1988.

Terhardt, E. "The Concept of Musical Consonance: A Link Between Music and Psychoacoustics." *Music Perception*, 1 (1984): 276-295.

———. "The Impact of Computers on Music: An Outline." In *Music, Mind and Brain*. Ed. M. Clynes. New York: Plenum Press, 1982.

Terhardt, E. and M. Seewann. "Aural Key Identification and its Relationship to Absolute Pitch." *Music Perception*, 1 (1983): 63-83.

Terhardt, E., G. Stoll, R. Schermbach and R. Parncutt. "Tonhöhenmehrdeutigkeit, Tonverwandschaft und Identifikation von Sukzessivintervallen." *Acustica*, 61 (1986): 57-66.

Terhardt, E., G. Stoll and M. Seewann. "Algorithm for Extraction of Pitch and Pitch Salience from Complex Tonal Signals." *Journal of the Acoustical Society of America*, 71 (1982): 679-688.

Terhardt, E. and W. D. Ward. "Recognition of Musical Key: Exploratory Study." *Journal of the Acoustical Society of America*, 72 (1982): 26-33.

Tervaniemi, M., K. Alho, R. Paavilainen, M. Sams and R. Naatanen. "Absolute Pitch and Event-related Brain Potentials." *Music Perception*, 1, no. 3 (spring 1993): 305-316.

Thompson, W. F. "Judgments of Key Change in Bach Chorale Excerpts: An Investigation of the Sensitivity to Keys, Chords, and Voicing." Unpublished doctoral dissertation. Department of Psychology, Queen's University. Kingston, Ontario, Canada, 1986. Cited in: C. L. Krumhansl. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press, 1990.

Tighe, T. J. and W. J. Dowling, eds. *Psychology and Music: The Understanding of Melody and Rhythm*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, 1993.

- Tisserand, M. and J. Guiolt. "Etude du Tempo de 335 Sujets Masculins dans la Région Parisienne." *Biotypologie*, 10, no. 11 (1949-1950): 89-94. Cited in: Paul Fraisse. "Rhythm and Tempo." In *The psychology of music*. Ed. D. Deutsch. Academic Press Series in Cognition and Perception. New York: Academic Press, 1982.
- Trotter, J. R. "The Psychophysics of Melodic Interval: Definitions, Techniques, Theory and Problems." *Australian Journal of Psychology*, 19 (1967): 13-25.
- Tulving, Endel. *Elements of Episodic Memory*. Oxford: Clarendon Press, 1983.
- Unyk, A. and J. Carlsen. "The Influence of Expectancy on Melodic Perception." *Psychomusicology*, 7 (1987): 3-23.
- Van Noorden, L. A. P. S. "Temporal Coherence in the Perception of Tone Sequences." Unpublished doctoral dissertation. 1975. Cited in: D. Deutsch. "Memory and Attention in Music." In *Music and the Brain*. Eds. M. Critchley and R. Henson. London: Heinemann, 1977.
- Vernon, P. E. "Absolute Pitch: A Case Study." *British Journal of Psychology*, 68 (1977): 485-489.
- Vos, J. "Purity Ratings of Tempered Fifths and Major Thirds." *Music Perception*, 3 (1986): 221-258.
- Vos, P. G. "Identification of Metre in Music." *Report 76 ON 06*. University of Nijmegen, 1976.
- . "Temporal Duration Factors in the Perception of Auditory Rhythmic Patterns." *Scientific Aesthetics*, 1 (1977): 183-199.
- Ward, W. D. "Musical Perception." In *Foundations of Modern Auditory Theory*. Ed. J. V. Tobias. New York: Academic Press, 1970.
- Ward, W. D. and E. M. Burns. "Absolute Pitch." In *The Psychology of Music*. Academic Press Series in Cognition and Perception. Ed. D. Deutsch. New York: Academic Press, 1982.
- Warren, R. M. and C. J. Obusek. "Identification of Temporal Order within Auditory

- Sequences." *Perception and Psychophysics*, 12 (1972): 86-90.
- Waters, J. "[Development of] Absolute Pitch." Unpublished paper. University of Minnesota, n.d. Cited in: M. Rush. *An Experimental Investigation of the Effectiveness of Training on Absolute Pitch in Adult Musicians*. Ph.D. diss. Ohio State University, 1989.
- Wedell, C. H. "The Nature of the Absolute Judgment of Pitch." *Journal of Experimental Psychology*, 17 (1934): 485-503.
- Wertheimer, M. "Untersuchungen zur Lehre von der Gestalt." *Psychologische Forschung*, 2 (1923): 301-350. Cited in: K. Koffka. *Principles of Gestalt Psychology*. 4th ed. First published in England in 1935. London: Routledge & Kegan Paul, 1955.
- Wing, H. "Tests of Musical Ability and Appreciation." *British Journal of Psychology*, 27 (1948): Monograph Supplements.
- Woodrow, H. "Time Perception." In *Handbook of Experimental Psychology*. Ed. S. S. Stevens. New York: Wiley, 1951.