THE ROLE OF NON-DIATONIC CHORDS IN PERCEPTION OF HARMONY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

NART BEDİN ATALAY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

JUNE 2007

Approval of the Graduate School of Informatics

____

Assoc. Prof. Dr. Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

____

Prof. Deniz Zeyrek
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

____

Prof. Hasan Gürkan Tekman
Supervisor

Examining Committee Members

Assist. Prof. Dr. Bilge Say (METU, COGS) ________________

Prof. Dr. Hasan Gürkan Tekman (Uludağ U., PSY) ________________

Prof. Dr. Türev Berki (Hacettepe U., MUS) ________________

Assist. Prof. Dr. Annette Hohenberger (METU, COGS) ________________

Assoc. Prof. Dr. Emre Özgen (Bilkent U., PSY) ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last Name : Nart Bedin, Atalay

Signature : __________________

ABSTRACT

THE ROLE OF NON-DIATONIC CHORDS IN PERCEPTION OF HARMONY

Atalay, Nart Bedin
Ph.D., Department of Cognitive Science
Supervisor: Prof. Dr. Hasan Gürkan Tekman

June 2007, 185 pages

The perceptual reality of the music theoretical relation between the Neapolitan chord and the dominant; and the secondary dominant chord and its diatonic associate was investigated within the chord priming paradigm. In Experiment 1, expectation towards the dominant chord after the Neapolitan chord was observed in Turkish musicians and non-musicians with piano timbre. In Experiment 2, expectation towards the dominant chord after the Neapolitan chord was observed in European musicians but not in European non-musicians. In Experiment 3, Turkish non-musicians were tested with Shepard tones; but it was not possible to observe any priming effects. To understand effects of cultural background on the difference between the results of Experiments 1 and 2 further studies are necessary. In Experiments 4-5, the perceptual reality of the relation between the secondary dominant chord and its diatonic associate was investigated in Turkish non-musicians. In Experiment 4, chord sequences that included secondary dominant chords were played with Shepard tones; and they were scrambled with 2by2 scrambling algorithm. Experiment 5 was identical with Experiment 4, except chord sequences were played with the piano timbre. Experiment 6 was identical with Experiment 5, except chord sequences were scrambled with 4by4. However, in Experiments 4-6 detrimental effects of scrambling sequences that include secondary dominant chords on the priming of chords were not observed. Turkish non-musicians did perceive the relation between the secondary dominant chord and its diatonic associate. In neural network simulations of this thesis it was shown that statistical learning from the musical environment with self-organization could be achieved without committing the questionable assumptions of previous studies.

Keywords: Music Perception, Harmony, Non-diatonic Chords, Self-Organizing Maps

ÖZ

DİYATONİK OLMAYAN AKORLARIN ARMONİ ALGISINDAKİ ROLÜ

Atalay, Nart Bedin
DOKTORA, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Prof. Dr. Hasan Gürkan Tekman

Haziran 2007, 185 sayfa

Napolitan ve dominant akorları ile ikincil dominant ve onun diyatonik eşliği arasında bulunan müzik teorik ilişkilerin algısal olarak gerçek olup olmadığı problemi akoru hazırlama yöntemi ile incelenmiştir. Birinci deneyde akor yürüyüşleri Shepard sesi ile çalınmış ve Napoliten akorundan sonra dominant için beklenti hem müzisyen ve hem de müzisyen olmayan Türk katılımcılardan gözlemlenmiştir. İkinci deneyde akor yürüyüşleri piyano sesi ile çalınmış ve Napoliten akorundan sonra dominant için beklenti sadece Avrupalı müzisyenlerden gözlenmiş, müzisyen olmayan Avrupalı katılımcılardan gözlemlenememiştir. Üçüncü deneyde uyaranlar Shepard sesi ile çalınmışsa da müzisyen olmayan Türk katılımcılardan akora hazırlanma etkisi gözlemlenememiştir. Bu nedenle kültürel farklılıkların birinci ve ikinci deneyin sonuçlarına etkisini anlamak için yeni çalışmalar yapılmalıdır. Dördüncü, beşinci ve altıncı deneylerde ikincil dominant ve onun diyatonik eşliği arasındaki müzik teorik ilişkinin algısal olarak gerçek olup olmadığı müzisyen olmayan Türk katılımcılarla incelenmiştir. Dördüncü deneyde ikincil dominant içeren akor yürüyüşleri Shepard sesleri ile çalınmış ve yürüyüşler 2ye2 sıra karıştırma algoritması ile karıştırılmıştır. Beşinci deney dördüncü deneyin aynıdır; tek farrk beşinci deneyde uyaranlar piyano sesi ile çalınmıştır. Altıncı deney de beşinci deneyin aynıdır; tek fark altıncı deneyde 4de4 sıra karıştırma algoritması kullanılmıştır. Dördüncü, beşinci ve altıncı deneylerde ikincil dominant içeren yürüyüşlerin akorlarının sıralarını karıştırmanın akora hazırlanma üzerine etkisi gözlemlenememiştir. Bu nedenle müzisyen olmayan Türklerin ikincil dominant ve onun diyatonik eşliği arasındaki ilişkiyi algılamadığı sonucuna varılmıştır. Bütün bunlara ek olarak, bu tezde yapılan yapay sinir ağları modelleri göstermiştir ki ağların bir müzikal çevreye dayalı olarak kendi kendini örgütlemesi ile gerçekleşen istatistiki öğrenme için önceki çalışmaların sorgulanabilir varsayımlarına dayanmak şart değildir.

Anahtar Kelimeler: Müzik Algısı, Armoni, Diyatonik Olmayan Akorlar, Kendiliğinen Örgütlenen Yapay Sinir Ağları

ACKNOWLEDGEMENTS

I would like to thank my advisor Hasan Gürkan Tekman and to Petri Toiviainen for their continued interest and kind support. This work was partially supported by Turkish Academy of Sciences (TÜBA) HGT-TUBA-GEBIP-2001-1-1 and Middle East Technical University (ODTÜ) BAP-08-11-DPT-2002-K-120510.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Music perception is the study of mind/brain mechanisms that enable us to distinguish, appreciate and respond to musical stimuli. It is the study of the musical faculty as well as the study of faculties that are not specialized in music. Music perception gets insights from and uses methods of musicology, ethnomusicology, psychophysics, neuroscience, cognitive psychology, and psychology of emotions. Music perception is one of the central topics of cognitive science.

Music is a universal ability of the human species. All known human populations have their own musical style (Krumhansl, 1992). Furthermore, there are universals between musical styles (Harwood, 1976; Dowling & Harwood, 1986). Music perception depends on innate capacities and acquired knowledge about music (Justus & Bharucha, 2002). Innate constraints bring forth basic perceptual skills upon which elaborated musical structures are built. According to many, innate musical capacities result from mechanisms responsible for the domain of general cognition, auditory cognition, or language (Justus & Bharucha, 2002; McDermott & Hauser, 2005; Pinker, 1997; Schwartz, Howe, & Purves, 2003; Terhardt, 1974; Trehub, 2000). Knowledge about a musical style is acquired by exposure to it. It is proposed that acquired knowledge of music stems from the encounter of a general learning mechanism and a highly structured musical environment (Bharcha, 1987; Krumhansl, 1990, 1992; Tillmann, Bharucha & Bigand, 2000).

A description of musical knowledge requires hierarchical organization of structural units, which are generated by a musical grammar (Bernstein, 1976; Horton, 2001; Lerdahl & Jackendoff, 1983). Many aspects of musical knowledge are described with grammars; such as metrical structure, melodic structure, grouping structure, sensation of stability, tension and relaxation, tonal hierarchies (see Deutsch & Feroe, 1981; Horton, 2001; Krumhansl, 1990; Lerdahl, 2001; Lerdahl & Jackendoff, 1983, Narmour, 1990). The psychological reality of these music theoretical descriptions has been a popular topic (see Deliège, 1987; Deutsch & Feroe, 1981; Krumhansl, 1990; Krumhansl, 1996; Lerdahl & Krumhansl, 2001, Lerdahl, Krumhansl, Hanon, & Fineberg, 2000).

Musical styles differ with respect to their grammar (Lerdahl & Jackendoff, 1983; Temperley, 2000). The hierarchical organization of musical elements and the grammar of music are acquired by listeners of the style (Bharcha, 1987; Krumhansl, 1990, 1992; Lerdahl & Jackendoff, 1983; Tillmann, Bharucha & Bigand, 2000; Trainor & Trehub, 1992, 1993, 1994; Trehub, Cohen, Thorpe, & Morrongiello, 1986). Grammaticality of music and the acquisition of a musical grammar remind us of the language faculty.

There are structural similarities between music and language. For example, both of them depend on processing of signals extended in time; and the hierarchical organization of the musical surface resembles syllabic stress (see Lerdahl & Jackendoff, 1983). Music and language share some neural circuits (Patel, 2003; Patel, Gibson, Ratner, Besson, & Holcomb, 1988; Maess, Koelsch, Gunter, & Friederici, 2001). However, similarity between music and language is limited. Musical and linguistic abilities may dissociate as a result of dysfunction of neural tissues (Luria, Tsvetkova, & Futer, 1965; Peretz, 1993). Contrary to language, music perception is free from long distance dependencies[1] (Deliège, Melen, Stammers, & Cross, 1996; Tillmann, Bigand, & Madurell, 1998; Tillmann & Bigand, 1998). Meaning in language is quite different from meaning in music. Linguistic

---

[1] Long distance dependency is a type of grammatical relation between functional elements, according to which there may be unlimited number of constituents between the dependent elements. For example, subject-verb agreement in English is a long distance dependency.

expressions refer to something other than themselves; i.e. they are symbolic. Linguistic meaning is based on meaning of atomic units and their syntactic combination (Fodor & Pylyshyn, 1988). There are musical pieces that function as symbols (e.g. the Wedding March is a symbol of wedding); but such instances are rare. Musical elements are not symbols, neither are their syntactic combinations. There are discussions about the nature of musical understanding. Parsing of musical structure, association of structural patterns with emotions, and experience of motion are proposed models of musical understanding (Bharucha, Curtis, & Paroo, 2006; Krumhansl, 1992; Lerdahl & Jackendoff, 1983; Meyer, 1956; Sloboda, 2004).

The cognitive psychology of music has primarily been interested in effects of musical knowledge on perception. There were laborious investigations to understand effects of musical knowledge on expectation, memory, conscious experience, categorical perception, and others (see Bharucha, Curtis, & Paroo, 2006; Deutsch, 1982, 1999, 2006; Deliege & Sloboda, 1997; Krumhansl, 1990, 1991, 2000, Sloboda, 1985 for reviews). But what is music perception? What do we perceive when we perceive music?

There are many aspects of music perception[2], perception of pitch, tuning, rhythm, tension and relaxation etc. The interaction between different musical percepts is rather complex. Figure 1.1 summarizes aspects of musical listening and their complex interactions. Pitch is the basic building block of music perception. The human voice, environmental sounds, and sound of musical instruments are composed of complex wave patterns. Pitch perception is the pattern recognition mechanism of sound waves. It is an important element of parsing sounds into distinct objects and streams (Bregman, 1990). Pitch height, pitch timbre and pitch class are elements of pitch perception. Pitch height is directly related with the

---

[2] For a detailed coverage of aspects of musical listening please refer to Bharucha, Curtis, & Paroo (2006) and Justus & Bharucha (2002).

frequency of the sound wave (see Moore, 1995; 2003). Perception of pitch is the basis of perception of melodies. The sensation of timbre enables us to associate sound waves with sound sources. For example the sound of a piano is differentiated from the human voice with respect to its timbre. Octave equivalence is the perceptual identity of two pitches whose frequency ratio is 2/1. Musical pitch is categorical; i.e. the frequency of every pitch category is predetermined. The match between frequencies and pitch categories is called tuning. The number of pitch categories of a musical style is finite. Different musical cultures have different set of pitch categories (Dowling & Harwood, 1986).

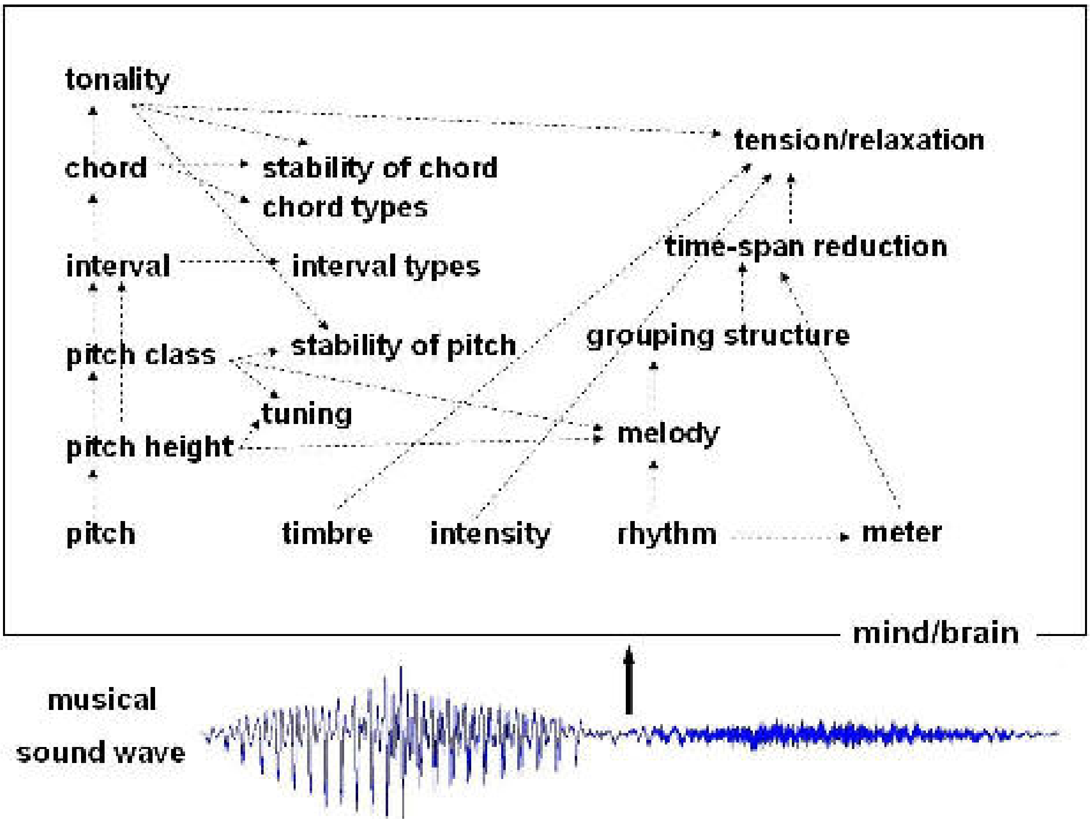


Figure 1.1 Aspects of musical listening and their interaction. Arrows represent information exchange between musical percepts (see also Bharucha, Curtis, Paroo, 2006; Lerdahl, 2001; Koelsch & Siebel, 2005).

An interval is the simultaneous or sequential sounding of two pitches. Consonance and dissonance are particular perceptual qualities of intervals. Consonance is the perceptual fusion or merger of two pitches; dissonance is the contrary. Consonance/dissonance may depend on the frequency of pitches (sensory consonance), or musical context in which pitches occur (Krumhansl, 1991). A chord is the simultaneous or sequential sounding of more than two pitches. There are different types of chords in Western tonal music; and they differ with respect to the sense that they induce. Chord is an important concept in the theory of Western tonal music. A branch of music theory, called harmony, is specifically involved with the study of chord relations.

Stability of pitch is determined by musical context. Stable pitches give us the sense of closure (Krumhansl and Shepard, 1979) whereas unstable pitches pop out (Janata, Birk, Tillmann, & Bharucha, 2003; see Krumhansl, 1990 for review). Stability of pitch is correlated with the number of occurrences of the pitch in the musical context and the musical style (Castellano, Bharucha. & Krumhansl, 1984; Krumhansl, 1990; Oram & Cuddy, 1995). It has been proposed that listeners of a musical style internalize information about typical frequency of occurrences of pitches of the style. This knowledge gives rise to the perception of stability (Krumhansl, 1990; Tillmann, Bharucha, & Bigand, 2000). Stability of pitches is organized hierarchically (Krumhansl and Shepard, 1979). This organization is named 'tonality'.

Hierarchical organization of the sense of stability applies to chords as well. Stability of chords is correlated with stability of pitches and frequency of occurrences of chords (Krumhansl, 1990). Stability of chords affects memory and expectation for chords and perception of relations between chords (Bharucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986; Krumhansl, 1990). The relation between chords is a subject of a music-theoretical study called harmony. The study of harmony is unique to Western tonal music. It is the result of highly structured chord relations of this genre. The acquisition and perception of chordal relations (harmony) and its effect on expectation is a well-studied topic of music perception (see Bharucha, 1987; Bharucha & Stoeckig, 1986, 1987; Bigand & Pineau, 1997; Bigand, Madurell, Tillmann & Pineau, 1999; Bigand, Poulin, Tillmann, & D'Adamo, 2003; Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001; Bigand, Tillmann, Poulin-Charronnat & Manderlier, 2006; Leman, & Carreras, 1997; Parncutt, 1988, 1994; Tekman & Bharucha, 1992, 1998; Tillmann & Bigand, 1998; Tillmann, Bigand, & Pineau, 1998; Tillmann, Janata, Birk, & Bharucha, 2003).

Music is composed of sequences of pitches or pitch related events. Therefore, there are temporal aspects of music perception as well[3]. The musical signal is divided into regular intervals, or beats. Beat is a perceptual phenomenon. It may or may not be equal to the onset and duration of pitch events. Beat is derived from the musical signal by the listener. Tempo is the speed at which musical events occur; technically it is the number of beats per minute. Meter is a hierarchical organization of beats. A rhythmic pattern consists of the musical elements structured in time. Pitch events combined with rhythmic patterns compose melodies. Elements of melodies are grouped hierarchically with respect to the rhythmic and metric structure (Lerdahl & Jackendoff, 1983). Event hierarchies are hierarchies of groups of musical elements. Event hierarchies refer to the specific musical events. Tonal hierarchies are abstract hierarchical relations between pitches. Tonal hierarchy and tonality are equivalent in this sense. Event hierarchies and tonal hierarchies give rise to musical tension and relaxation patterns. (Lerdahl, 2001; Lerdahl & Jackendoff, 1983).

Music perception studies have implications for other faculties of cognition, or cognitive science in general. For example, Bharucha & Mencl (1996) have showed that octave equivalence, a musical universal, may develop as a result of a general learning mechanism. This result contributed to the nurture vs. nature debate of

[3] For a detailed account of temporal aspects of music perception see Fraisse (1982) and Clarke (1999).

cognitive science. By interpreting results on absolute pitch perception Deutsch (2006) has conjectured a language acquisition mechanism, which associates names with absolute pitches, and shuts off with the absence of the cues in the linguistic environment and maturation.

In this thesis the perception of relations between chords (or the perception of harmony) in Western tonal music is investigated. The perception of the relation between chords has been investigated with subjective reports (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha & Kessler, 1982; Krumhansl, Bharucha & Castellano, 1982), recognition memory experiments (Bharucha & Krumhansl, 1983; Krumhansl & Castellano, 1983; Krumhansl, Bharucha & Castellano, 1982) and the priming paradigm (Bharucha & Stoeckig, 1986, 1987; Bigand & Pineau, 1997; Bigand, Madurell, Tillmann & Pineau, 1999; Tekman & Bharucha, 1992, 1998).

Chord priming is an indirect method of investigating the perception of the relation between chords. It measures the expectation induced by a musical context towards a target chord. Musical events (the priming context) that precede the target lead participants to develop expectations for the target chord. Participants are asked to make a binary decision about a perceptual aspect of the target chord. Reaction time and accuracy of responses reflect the amount of expectancy. According to Meyer (1956) satisfied and dissatisfied expectations in music are the basis of musical meaning; for Tillmann (2005) expectations shape the listener's interaction with music by facilitating and delaying the processing of sound events. The chord priming paradigm is adequate for collecting data from both musicians and non-musicians. It was the experimental paradigm used in this thesis.

It has been repeatedly shown that priming of chords follows along a music theoretical construct, the circle of fifths (Bharucha & Stoeckig, 1986, 1987; Bigand & Pineau, 1997; Bigand, Madurell, Tillmann & Pineau, 1999; Bigand, Poulin, Tillmann, & D’Adamo, 2003; Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001; Tekman & Bharucha, 1992, 1998; Tillmann & Bigand, 1998; Tillmann, Bigand, & Pineau, 1998). On the circle of fifths all possible major and minor chords of Western tonal music are located with respect to their stability[4]. The distance on the circle of fifths between the priming context and the target chord indicates the amount of expectation towards the target; the greater the distance, the less the expectancy for the target chord. This holds when the priming context is composed of a single chord (Bharucha & Stoeckig, 1986, 1987; Tekman & Bharucha, 1992, 1998) or a chord sequence (Bigand & Pineau, 1997; Bigand, Madurell, Tillmann & Pineau, 1999; Bigand, Poulin, Tillmann, & D'Adamo, 2003; Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001; Tillmann & Bigand, 1998; Tillmann, Bigand, & Pineau, 1998). Priming of chords has been modeled with a neural network called MUSACT (Bharucha, 1987). MUSACT is a connectionist model of tonal knowledge of Western music. It simulates quite a range of chord priming experiments (Tillmann, Bigand, & Bharucha, 2000).

Results of priming experiments, subjective reports (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha & Kessler, 1982; Krumhansl, Bharucha & Castellano, 1982) and recognition memory judgments (Bharucha & Krumhansl, 1983; Krumhansl & Castellano, 1983; Krumhansl, Bharucha & Castellano, 1982) reflect the fact that the perception of the relation between chords is related with the stability of chords (Krumhansl, 1990). Stable chords are perceived as related. Relatedness between stable and unstable chords depends on their relative position in time. Perceived relatedness increases when the stable chord follows the unstable chord, compared to the opposite positioning.

Based on results of experiments obtained so far, it is safe to conclude that the perception of a relation and expectation between chords depend on stability of

---

[4] For a detailed description of major and minor chord types and the circle of fifths please refer to Section 2.1.

chords (Bharucha, 1987; Krumhansl, 1990; Tillmann, Bharucha, Bigand, 2000). The main objective of the present thesis work is to challenge this conclusion. By investigating perceptual relations between stable and unstable chords the aim is to show that stability of chords is not the only determinant of the expectation between chords. The study of harmony investigates the music theoretical relation between chords (Piston, 1978). The proposed music-theoretical relations between chords may be seen as hypotheses to be tested empirically. Music theoretically, the Neapolitan chord and the dominant are related; as well as the secondary dominant chord and its diatonic associate[5]. In this thesis the perceptual reality of these relations was investigated. The Neapolitan chord regularly precedes the dominant chord in musical pieces according to music theory (Piston, 1978). The Neapolitan chord is an unstable chord; whereas the dominant is a stable chord (Krumhansl, 1990). In this thesis, priming of the dominant chord by the Neapolitan chord was investigated. Music theoretically, the secondary dominant chord has to precede, and must not succeed, its diatonic associate (Piston, 1978). Secondary dominant chords are unstable and diatonic chords are stable chords (Krumhansl, 1990). In this thesis, the perceptual reality of this sequential relation was investigated by its effect on the priming of tonic and subdominant chords. The results of experiments were compared with predictions of the theories of harmonic knowledge, namely predictions of MUSACT (Bharucha, 1987) and Tonal Pitch Space (Lerdahl, 2001). By studying the perceptual relations between the Neapolitan chord and the dominant and between the secondary dominant chord and its diatonic associate, it is possible to understand whether specific associations and sequential order between chords are coded in the implicit knowledge of tonal harmony.

Chord priming is governed by cognitive processes, i.e. learned knowledge about Western tonal music. This is a reasonable conclusion given the fact that chord priming is observed for both musicians and non-musicians (Bharucha & Stoeckig,

---

[5] For a detailed definition of these chord types see Section 3.1.

1986, 1987; Bigand & Pineau, 1997; Tekman & Bharucha, 1992, 1998; Bigand et al. 1999). It is not governed by sensory processes, i.e. acoustical[6] properties of musical sound (Bigand, et al. 2003; Bigand, et al. 2005; Tekman & Bharucha, 1998; Tillmann, Koelsch, Escoffier, Bigand, Lalitte, Friederici, & vonCramon, 2006). Chord priming and statistical distribution of chords in the tonal music match: frequency of occurrence of chords in musical pieces correlates with perceived stability of chords (Krumhansl, 1990). The circle of fifths summarizes co-occurrence relations between chords (Justus & Bharucha, 2002). It is proposed that listeners of Western musical styles acquire statistical regularities of chords by being exposed to musical pieces; and that knowledge governs the expectation and perception of relation between chords (Bharucha, 1992; Justus & Bharucha, 2002; Leman, & Carreras, 1997; Tillmann et al. 2000). Statistical learning is acquiring implicit knowledge about statistical properties of events in the environment without external feedback (Patel, 2005). MUSACT (Bharucha, 1987) represents the end state of such a learning process.

The mechanism of statistical learning of chord relations has been studied with neural network simulations (Bharucha & Todd, 1989; Leman, & Carreras, 1997; Tillmann et al. 2000). The learning mechanism has been modeled with self-organizing neural network architectures (Kohonen, 1982; Rumelhart & Zipser, 1985) and feed-forward neural networks with back-propagation (Rumelhart, Hinton & Williams, 1986). The neural network models are trained with a representative corpus of Western tonal music to obtain detectors of pitches, chords and chord relations. The learning architecture is coupled with a short term memory (STM) model of pitch. Sometimes the output of the STM model is the input of the statistical learning algorithm (Bharucha & Todd, 1989; Leman, & Carreras, 1997; Tillmann et al. 2000). This way, the learning architecture becomes sensitive to the frequency of occurrence of

[6] Acoustical properties are properties of sound structure.

musical events. The STM model has been implemented as a leaky integrator[7] (Bharucha & Mencl, 1996; Bharucha & Todd, 1989; Leman, & Carreras, 1997; Tillmann et al. 2000; see also Krumhansl, Louhivuori, Toiviainen, Järvinen, & Eerola, 1999; Krumhansl, Toivanen, Eerola, Toiviainen, Järvinen, & Louhivuori, 2000). Frequent and infrequent events are represented with higher and lower activation patterns respectively; and frequent events find higher chances of being internalized by the learning mechanisms compared to infrequent events.

Tillmann et al. (2000) studied the self-organization of MUSACT (Bharucha, 1987). They trained self-organizing neural network architectures with typical chord progressions of Western tonal music. In order to obtain the architecture of MUSACT, Tillmann et al. (2000) applied pruning and scaling algorithms on the weights of self-organized networks. In the end, they simulated many important studies on perception of chords with their networks. However, pruning and scaling of weights were based on many questionable assumptions[8]. This can doubt the possibility of the self-organization of chord schema. In this thesis, self-organizing models of chord perception that were not based on pruning and scaling of weights were studied. Important results on perception of chords were simulated. The possibility of the self-organization of chord-schema was justified.

The thesis is organized as follows: In Chapter 2 music theoretical concepts related with the thesis work will be introduced; studies on the perception of tonal aspects of music will be summarized, and theories formulating mechanism of perception of chords will be described. In Chapter 3, the rationale behind studying non-diatonic chords will be explained; studies focused on perception of non-diatonic chords will

---

[7] Leaky integrator equation is a differential equation of the form $\Delta x/\Delta t=Ax+C$, where A and C are constants. It describes a system that integrates input but gradually leaks a small amount over time (http://en.wikipedia.org/wiki/Leaky_integrator). A leaky integrator model of STM represents events as a function of their recency.

[8] See Section 5.2 for the discussion of these assumptions.

be summarized, and the research plan of the thesis work will be described. In Chapter 4, behavioral experiments conducted to investigate the perception of the Neapolitan and secondary dominant chords will be described. Predictions of MUSACT (Bharucha, 1987) and Tonal Pitch Space (Lerdahl, 2001) will be compared with the results obtained. In Chapter 5, the rationale behind studying self-organizing maps of chord schema will be described; self organizing maps will be introduced; and the simulation results will be presented. In Chapter 6, the results of behavioral experiments and self-organization map simulations will be discussed.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

In this chapter music theoretical concepts related with the thesis work are clarified; studies on the perception of tonal aspects of music are summarized; and theories formulating mechanisms of perception of chords are described.

2.1 PITCH, INTERVAL, CHORD, AND TONALITY

In this section, music theoretical concepts essential to follow the thesis work are presented. These concepts were developed to describe the structure of music by music theorists (see Randel, 2003). Perceptual aspects of some of these theoretical concepts will be introduced in Section 2.2.

In the domain of physics, music is sound waves. A sound wave is a pattern of air pressure differences. Musical waves are periodic sound waves, whose air pressure patterns are regular. Sound waves are described by frequency, intensity and shape. Frequency of a wave is measured by Hertz (Hz), the number of periodic cycles per second. Frequency is associated with pitch. Even though frequency is a continuous feature, musical pitch (henceforth pitch) is categorical, which means that the frequency of every pitch is pre-determined. Tuning is adjusting the frequencies of an instrument to set them in agreement with some pitch (see Randel, 2003).

Different musical cultures have different pitch categories. In Western music, there are 12 pitch categories (or pitch classes), namely A, A# (read A sharp)/Bb (read B flat[1]), B, C, C#/Db, D, D#/Eb, E, F, F#/Gb, G, G#/Ab. The chroma circle (Figure 2.1) defines the cyclic sequential relation between pitch categories. Conventionally,

[1] In music theory, flat and sharp are together called 'accidentals'.

frequency of A is 440 Hz etc. The frequency ratio of adjacent pitches on the chroma circle is $2^{1/12}$. For example, A#/Bb is ~466.16 Hz, B is ~493.88 Hz. Pitches with a frequency ratio of 2/1 are categorized as same. This is called octave equivalence. For example, pitches with 27.5 Hz, 55 Hz, 110 Hz, 220 Hz, 880 Hz are all As.

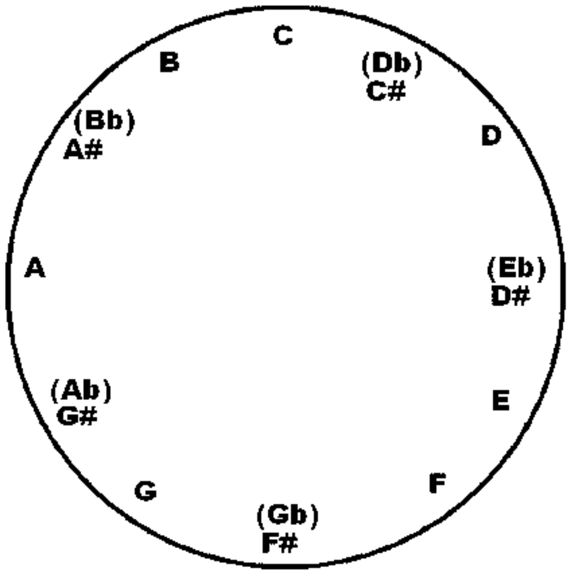


Figure 2.1 Chroma circle.

An interval specifies the relationship between two pitches. An interval is classified according to a) the number of steps between the two pitches on the chroma circle, which is indicated by an adjective (major, minor, perfect, diminished, augmented) b) the number of steps between two pitches on a diatonic scale, which is indicated by an ordinal number (see below for a definition of diatonic scale). An interval can be described by the ratio of frequencies of two pitches (see Randel, 2003). For example, C-D is a major second (8:9), C-E is a major third (27:32), C-F is a perfect fourth (3:4), C-G a perfect fifth (2:3), C-A forms a major third (16:27), C-B is a major seventh (128:243).

Chords are more than two pitches sounding simultaneously. There are four basic chord categories i.e., major, minor, augmented, diminished. A major chord is composed of a base pitch (this may be any pitch), the pitch a major third above the base, and the pitch a perfect fifth above the base. A minor chord is composed of a base pitch (this may be any pitch), the pitch a minor third above the base, and the pitch a perfect fifth above the base. The only difference between major and minor chord is the third: there is a major third interval with the base in the major chord, minor third interval with the base in the minor chord. A diminished chord is composed of a base pitch (this may be any pitch), the pitch a minor third above the base, and the pitch a diminished fifth above the base. An augmented chord is composed of a base pitch (this may be any pitch), the pitch a major third above the base, and the pitch an augmented fifth above the base. Chords are named with reference to the base and the category of the chord. For example pitches of C major are C, E and G; and pitches of C minor chord are C, Eb and G. A branch of music theory, called harmony, investigates the relations between chords. The study of harmony is unique to Western music. This is because only Western musical cultures somehow developed complex chord and chord relation structures.

Tonality is the hierarchical organization of pitch categories according to a central pitch category. Tonalities of Western tonal music are based on seven pitch categories. Tonalities are classified according to the central pitch category and mode. The most common modes are major and minor. A major tonality typically includes major intervals; and a minor tonality typically includes minor intervals. Pitch categories of a major tonality include the central pitch category (this may be any pitch), and pitch categories major second, major third, perfect fourth, perfect fifths, major sixth, and major seventh above it. Pitch categories of a minor tonality include the central pitch category (this may be any pitch) and pitch categories major second, minor third, perfect fourth, perfect fifths, minor sixth, and minor seventh above it. For example pitch categories of C major tonality are C, D, E, F, G, A, B; pitches of C minor tonality are C, D, Eb, F, G, Ab, Bb. Given that chords are composed of pitches, a tonality consequently determines a set of chords. For example, chords of the C major tonality are C major, D minor, E minor, F major, G major, A minor, and B diminished. Pitches (chords) that are members of a tonality are called diatonic; pitches (chords) that are not members of a tonality are called non-diatonic.

A scale is the collection of pitches sequenced from lower to higher or from higher to lower. Order between pitches is determined by the chroma circle. For example, 'C, D, E, F, G, A, B' is the scale of C major. Scales sequenced on the clockwise direction of the chroma circle are called ascending scales. The previous example is the ascending scale of C major. Scales sequenced on the counterclockwise direction are called descending scales. For example, 'B, A, G, F, E, D, C' is the descending scale of C Major. Scales are ascending by default. Diatonic scale is the scale of pitch categories of a tonality. Non-diatonic scale includes non-member pitch categories.

Degree is the numbered position of pitch categories on a major or minor scale (see Randel, 2003). Degrees are referred by names and roman numerals. A chord built on a pitch category is also named with the same degree name of the pitch. The first degree (I, the central pitch or the chord based on this pitch) is tonic; the second degree (II, the second pitch or the chord based on this pitch) is supertonic; the third degree is mediant (III); the fourth degree is subdominant (IV); the fifth degree is dominant (V); the sixth degree is submediant (VI); the seventh degree is leading (VII).

If a musical piece is written in a tonality, pitch categories of the tonality (i.e. pitches of the diatonic scale, or simply diatonic pitches) are primarily used to compose the piece. In addition, tonality imposes a hierarchy of importance to the pitch categories. Diatonic pitches are more important than non-diatonic pitches. Among diatonic pitches, tonic is the most important, followed by mediant and dominant, followed by other degrees of the tonality. The importance of pitches is reflected in the frequency of occurrence of pitches in musical pieces. Important degrees of a tonality occur frequently in a piece written in that tonality. Tonality imposes a hierarchy of importance to chords as well. Among diatonic chords, tonic is the most important, followed by dominant and subdominant, followed by other chord degrees. Diatonic chords are important than non-diatonic chords.

Tonalities are related to each other different levels. Relatedness is based on the number of shared pitches. For example, tonalities C major and G major are closely related because they share six pitches; but C major and B major are unrelated because they share only a single pitch.

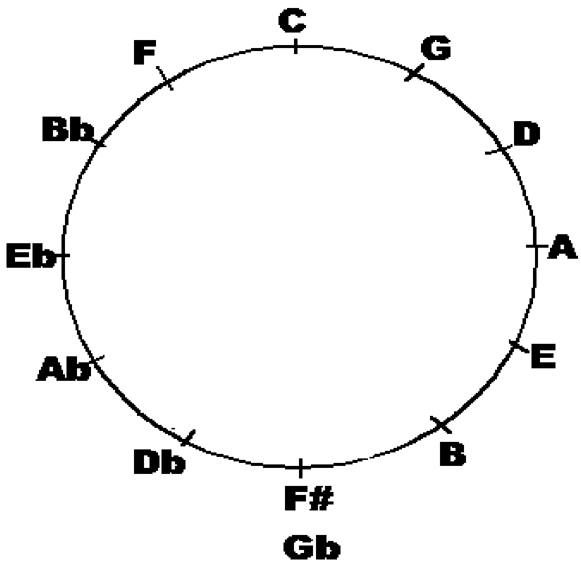


Figure 2.2 Circle of fifths.

A graph called circle of fifths (Figure 2.2) describes various aspects of pitches, chords and tonalities in a concise form. If we read the symbols of the circle as pitch categories, seven adjacent pitches on the circle form a major tonality. For example, F, C, G, D, A, E, and B form the C major tonality. If we read the symbols of the circle as major chords, three adjacent chords are subdominant, tonic and dominant of a major tonality. For example, F major, C major, and G major are subdominant, tonic and dominant of the C major tonality. If we read the symbols of the circle as minor chords, three adjacent chords are supertonic, submediant and mediant of a major tonality. For example, F minor, C minor, and G minor are supertonic, submediant and mediant of the Eb major tonality. If we read the symbols of the circle as major tonalities, the distance on the circle reflects the relatedness between tonalities. For example, C major and G major are closely related and they are adjacent on the circle; C major and B major are unrelated and they are distant on the circle.

2.2 PERCEPTUAL STUDIES ON PITCH, INTERVAL, CHORD, AND TONALITY

In this section, perceptual and cognitive studies on tonal aspects of musical listening are surveyed, by giving emphasis on studies related to this thesis work.

Sensation of pitch is essential to appreciate music. Every human being with a normal brain can sense pitch. Pitch perception is a pattern recognition mechanism to parse auditory scenes into sound sources (Bregman, 1984). Pitch perception has an important role in speech perception as well: perception of intonation of speech depends on pitch perception (Vaissière, 2005). The sense of pitch corresponds to the frequency of the sound wave. Fundamental frequency of a pitch is the frequency of the sinusoidal sound wave that induces the sense of the pitch. A wide range of sound waves could give sense of the same pitch; even the fundamental frequency of the pitch may be absent in the complex sound wave pattern. This is the central problem of understanding mechanism of pitch perception (see Moore, 2003). The perception of pitch without a fundamental frequency resembles the perception of illusory contours (Justus & Bharucha, 2001). According to Zatorre (1988) Heschl's gyrus has a role in the perception of pitch. Infants are supposed to learn frequency patterns of speech signals and store them as templates. The pitch sensed was a function of the match between the template and the sound signal (Terhardt, Stoll & Seewan, 1982a, 1982b).

Pitch height and pitch category are aspects of musical pitch perception. Pitch height is the logarithmic function of the fundamental frequency of musical sound waves. Perception of pitch class is based on the fact that pitches with a frequency ratio of 2/1 are perceived as the same. (In music theory this is called octave equivalence.) Shepard (1964) demonstrated the separation of perception of pitch class from pitch height with Shepard tones. A Shepard tone is composed of sinusoidal sound waves, and frequency ratio between the sinusoidals is the powers of 2. In other words, they are octave equivalent (they belong to the same pitch category). Shepard tones (1964) have an ambiguous pitch height, but definite pitch category (Justus & Bharucha, 2001). If an ascending (or descending) scale is played with Shepard tones (1964) again and again, same pitch categories will be perceived repeatedly, while pitch height increases (or decreases) constantly. Such an illusory percept is reminiscent of the lithograph print works of M.C. Escher, "Ascending and Descending". The lithograph depicts an old monastery and monks. Some of the monks continuously climb a never-ending stair located on the top of the monastery. Shepard tones create the same kind of illusion in the auditory domain.

Musical pitches are perceived categorically (see Section 1.2 for the description of pitch categories). A small percentage of people are talented for being aware of musical pitch categories i.e. they name the pitch of the sound immediately and correctly without a reference pitch (Ward and Burns, 1982). Those people are absolute pitch possessors. According to Deutsch (2006), a language acquisition mechanism is involved for the acquisition of absolute pitch perception. Categorical pitch perception was also observed with relative pitch possessors, who can name the pitch given a reference pitch (Burns and Campbell, 1994). Both musicians and non-musicians retained absolute pitch of musical pieces that they know well (Terhardt & Ward, 1982); and absolute pitch of melodies influenced short term memory of melodies (Cuddy, Cohen and Miller, 1979).

Shepard (1982) introduced geometric models to describe the cognitive representation of pitch. Perception of pitch height is directly related to the frequency of the pitch: as the frequency goes up, so does the pitch height. Perception of pitch class is based on the fact that pitches that are multiples of an octave apart are perceived as the same. In addition, pitch perception is affected by the context in which pitches occur (Krumhansl & Shepard, 1979; see below). Shepard (1982) proposed a torus[2], which is defined by pitch height, chroma circle (Figure 2.1) and circle of fifths (Figure 2.2), as a geometric model of pitch. This representation captured many aspects of pitch perception (see Shepard, 1982). Even this multidimensional representation is inadequate to capture relations between pitches within a musical context. Pitch relation in context is dynamic and asymmetric (see below) (Krumhansl, 1979, 1990). Naturally, it is impossible to capture these features with an Euclidean geometric shape.

Some intervals sound pleasing or non-disturbing but others sound ugly or disturbing when heard in isolation. This is called 'tonal consonance/dissonance of intervals'. For example, perfect fifth and major third are consonant intervals; major seventh and minor third are dissonant intervals. Dissonance of an interval was shown to be a function of the frequency difference between sinusoidal components of complex wave patterns (Helmholtz, 1954 cited in Moore, 1990; Plomp & Levelt, 1965). Burns and Ward (1978) showed that musicians, but not non-musicians, perceive musical intervals categorically. On the other hand, both musicians and non-musicians utilize intervals to code melodies given the fact that the melody is the same even if it starts with a different pitch. For example, one can play the same melody by starting with a different key on the piano. This way, pitches of the melody change while intervals between pitches of the melody are kept constant. Therefore, intervals, but not pitches, play a role in memory for melodies.

---

[2] Torus is a surface having the shape of a doughnut (Pickett, 2003).

Perception of a chord is more than just the addition of senses of pitches that compose the chord. Musically experienced listeners perceive major chords as category prototypes (Acker, Pastore & Hall, 1995); and components of the major chords are perceived integrally (Acker & Pastore, 1996). Perceptual fusion of chords has been explained by the template matching model of pitch perception (Terhadt, 1974; Parncutt, 1988). On the other hand, according to Tillmann, Bharucha & Bigand (2000) the perceptual fusion of chords is the result of a mechanism that develops feature detectors for chords.

The musical context also affects perception of pitches and chords. The probe tone technique was designed to investigate pitch perception within a musical context (Krumhansl & Shepard 1979). In this technique, participants listen to a musical context, followed by one or two pitches (probe tone(s)). The context can be a scale, a chord sequence or a musical excerpt. Participants are asked to report a specific aspect of their perceptual experience about probe tone(s) by designating a value on a scale.

Krumhansl and Shepard (1979) investigated effects of major and minor tonalities on the perception of pitches. They utilized major and minor scales as contexts, followed by a probe tone that was one of the chromatic pitches. Participants were asked to rate how well the probe tone completed the scale. For musicians, ratings were hierarchical; tonic received the highest rating followed by dominant and mediant, followed by other members of the scale, followed by non-scale (non-diatonic) pitches. For non-musicians, ratings were not hierarchical; they were determined by the pitch height of the probe tone. However, various follow-up studies obtained hierarchical ratings from non-musicians by using instructions clear to them (Cuddy & Bardertscher, 1987; Krumhansl & Keil, 1982; Speer & Meeks, 1985). Figure 2.3 shows probe tone ratings (also called stability profiles) for major and minor tonalities (Krumhansl & Kessler, 1982). Stability profiles match well with the music theoretical description of hierarchy of importance of pitches within tonality (Krumhansl, 1990).

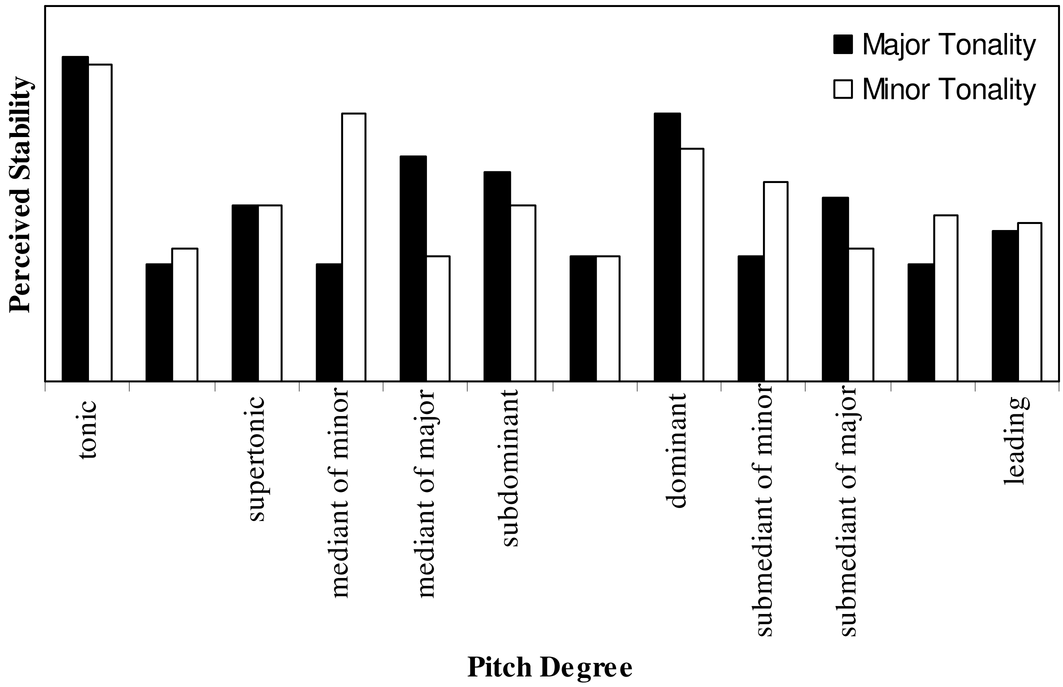


Figure 2.3 Stability profiles of pitches for major tonality (adapted from Krumhansl, 1990). On the horizontal axis only diatonic pitches are marked.

The hierarchical organization of pitches according to their stability is called the sensation of tonality. Studies showed that the sense of tonality develops rapidly; it may shift within a piece; and the shift between related tonalities is quicker than the shift between unrelated tonalities (Krumhansl & Kessler, 1982; Toiviainen & Krumhansl, 2003). It is possible to sense more than one tonality at once; but is not possible to selectively attend to only one of them (Krumhansl & Schmuckler, 1982). Cross-cultural studies observed that tonalities of different musical cultures induce different hierarchical organization of pitch stability (Castellano, Bharucha, Krumhansl, 1984; Kessler, Hansen, Sheapard, 1984; Nam, 1998). Hierarchical organization of pitch stability predicts various other percepts, such as the sense of the end of musical phrases (Palmer & Krumhansl, 1987a, 1987b), melodic expectations (Schmuckler, 1989), the pop-out of non-diatonic pitches (Janata et al., 2003).

Musical context affects the perceived relation between pitches as well. Krumhansl (1979, 1990) utilized major and minor scales as contexts followed by two probe tones. Probe tones were all possible combinations of the 12 chromatic pitches. Participants were asked to rate how well the second pitch followed the first one. Results showed that the perceived relation between pitches was dynamic and asymmetric. It was dynamic because pitches higher in the hierarchy were perceived as more related compared to pitches lower in the hierarchy; it was asymmetric because the hierarchy of the second pitch contributed more to perceived relatedness than the hierarchy of the first pitch. Krumhansl & Kessler (1982) analyzed relatedness scores with a multidimensional scaling technique to picture perceived relations between pitches after musical context. The solution was a cone. Position in the hierarchy (stability value) and position on the chroma circle determined the coordinates of pitches on the cone.

Krumhansl (1990) investigated effects of major and minor tonalities on perception of chords with experiments similar to the probe tone studies. This time probes were major and minor chords. Participants were asked to rate how well the probe chord fit to the context. The ratings of chords were hierarchical: the highest rating was given to the tonic chord, followed by subdominant, dominant, and mediant, followed by other diatonic chords (Krumhansl, 1990). The ratings of chords matched well the music theoretical description of the hierarchy of importance of chords within a tonality (Krumhansl, 1990). The hierarchy of chords and the hierarchy of stability of pitches that make up the chords were correlated significantly; but the correlation did not explain all of the variation in the ratings of chords (Krumhansl, 1990). As a result, perception of chords was determined by the perception of the component pitches; but there were other influences as well. Perceived relations between chords were dynamic and asymmetric, parallel to what has been found for pitches (Krumhansl, Bharucha & Castellano, 1982; Krumhansl, Bharucha & Kessler, 1982; Bharucha & Krumhansl, 1983).

Tonalities share chords (see Section 1.2). Depending on the tonality, the same chord may be at different positions in the hierarchy of stability. Listeners of Western tonal music are aware of different functions of the same chord in different tonalities (Krumhansl, Bharucha & Kessler, 1982). In addition, a single chord is sufficient to induce a hierarchical organization of chords (Bharucha & Krumhansl, 1983).

Prototype theory of category perception (Rosch, 1973; 1975) explains perception of pitches and chords within context well. Eleanor Rosch proposed that some members of a category are better representatives of it (they are prototypes of the category); they constitute basic building blocks around which categories are formed (Rosch, 1973). Furthermore, they function as reference points by which other instances of the category are interpreted (Rosch, 1975). Rosch showed, for example, that horizontal and vertical lines are natural reference points for the perception of tilted lines (Rosch, 1975). In a similar fashion, pitches and chords higher in the hierarchy of stability are reference points for those lower in the hierarchy. For example, tonic, which is the highest in the hierarchy, is a reference point for all other degrees.

Reference points of some domains of perception may be static; for example, for the perception of lines, reference points are always horizontal and vertical lines. On the other hand, for the perception of pitches and chords references change occasionally; different tonalities have different tonics, and the sense of tonality may change within a piece, and from piece to piece.

Reference points of some domains of perception may be natural (Rosch, 1973); they depend on the evolved structure of the perceptual system. Contrarily, reference points of pitches and chords are learned; and they depend on the acquired structure of the perceptual system. Different musical cultures have different pitch categories and different hierarchical organization of pitches. Therefore, children are supposed to learn pitch categories and hierarchical organization of pitches. This knowledge, called tonal schema (Bharucha, 1987; Leman 1995), is the source of perception and appreciation of tonal aspects of music.

Bharucha (1987) envisioned that tonal schemata might affect processing of chords. He proposed that if a tonal schema is activated while listening to music, chords that conform to the active schema must be processed easily. The chord priming paradigm was introduced to test this idea (Bharucha & Stoeckig, 1986, 1987). In chord priming studies effects of tonal context on the processing of the target chord are examined. The tonal context may be a single chord (Bharucha & Stoeckig, 1986) or a chord sequence (Bigand & Pineau, 1997). A target chord follows the context immediately. Participants are asked to make a binary judgment with an objectively correct answer on the target chord. The binary judgment used for this purpose has usually been a consonance/dissonance or in-tune/out-of-tune discrimination on the target chord. Consonant major chords have been contrasted with chords that were made dissonant by either replacing a component pitch with a non-diatonic pitch (Bharucha & Stoeckig, 1986; Bigand, Madurell, Tillmann & Pineau, 1999; Tillmann, Bigand, & Pineau, 1998) or adding a non-diatonic pitch to the chord (Bigand & Pineau, 1997) or by altering the frequency of a pitch to make it out-of-tune (Bharucha & Stoeckig, 1986, 1987; Tekman & Bharucha, 1992, 1998).

In Bharucha and Stoeckig (1986), primes and targets were single chords. Participants were asked to decide whether the target was in-tune or out-of-tune. Participants were faster and more accurate for their judgments when the prime was close to the target on the circle of fifths compared to when it was distant. Chord priming was not due to acoustical similarity[3] between prime and target (Bharucha & Stoeckig, 1987); chord priming dominated acoustical similarity (Tekman & Bharucha, 1992, 1998); chord priming was not affected from the short term memory of and explicit knowledge about the target chord (Justus & Bharucha, 2001); and explicit musical knowledge of participants, and their musical training did not change the advantage of closer prime over the distant one (Bharucha & Stoeckig, 1987; Tekman & Bharucha, 1992, 1998).

Because chord priming is not due to the acoustical similarity between prime and target chords, it must not be governed by sensory processes (i.e. structure and function of the ear and peripheral nervous systems), but it must be a result of some cognitive processes (i.e. higher level brain processes, such as association, categorization, statistical learning etc.). In addition, it must be a result of irrepressible, automatic (or modular) processes, since it cannot be eliminated with explicit knowledge about the target chord (Justus & Bharucha, 2002). Because the amount of musical training and explicit knowledge about music do not change the chord priming (Bharucha & Stoeckig, 1987; Tekman & Bharucha, 1992, 1998), chord priming must be a result of an implicit learning process. Bharucha (1987) proposed that chord priming is due to implicit knowledge about the relationship between pitches, chords and tonalities, which can be implemented with neural networks (his model is called MUSACT, see section 2.3 for a detailed description of MUSACT). Such knowledge is learnable (in other words MUSACT can self-organize) by mechanisms acting as feature detectors that are sensitive to statistical properties of events in the environment (Bharucha, 1992; Tillmann, Bharucha & Bigand, 2000).

---

[3] Acoustical properties are sound-related properties. In this context, acoustical similarity refers to the similarity of sound structure of prime and target.

MUSACT (Bharucha, 1987) simulates chord priming. Chord priming is related with the distance on the circle of fifths. The circle of fifths is a music theoretical device to describe stability relations between chords. It correlates with the frequency of occurrence of chords (Krumhansl, 1990; Justus & Bharucha, 2001). Stability of chords also correlates with the frequency of the occurrence of pitches and chords (Krumhansl, 1990). Therefore, the circle of fifths, priming and stability of chords seem to reflect the same phenomena: the statistical regularities of chords and the acquired knowledge about them. This knowledge has been modeled with MUSACT (Bharucha, 1987).

Bigand & Pineau (1997) initiated global chord priming studies in which primes were chord sequences and targets were single chords. In these studies, two chord sequences were arranged in different tonalities. The final two chords of the sequences were identical. Depending on the tonality of the sequences the final chord (target) was either the tonic or the subdominant of the sequence. The overall distance on the circle of fifths between the priming chords and tonic target is smaller, compared to the subdominant target. Participants were faster and more accurate for their judgments when the target was the tonic compared to when it was the subdominant (Bigand & Pineau 1997, Bigand, Madurell, Tillmann & Pineau 1999, Experiment 1). This effect was called global priming. The chord that immediately preceded the target (penultimate) was identical in different sequences. Therefore, factors that may cause priming could not be local (within the immediate vicinity of the target). Global priming of chords has been repeatedly observed for both musicians and non-musicians (Bigand & Pineau 1997; Tillmann & Bigand, 1999; Bigand, Poulin, Tillmann, & D'Adamo, 2003). Global and local factors of priming interacted with each other (Tillmann, Bigand & Pineau,1998); global priming developed gradually as the number of chords in the priming sequence increased (Bigand, et. al. 1999, Experiment 2); global priming interacted with temporal aspects of the priming sequence (Bigand, et. al. 1999, Experiment 3); global priming was not affected by changing the order of chords in the priming sequence (Tillmann & Bigand, 1999); global priming dominated acoustical similarity between the priming sequence and the target (Bigand, Poulin, Tillmann, & D'Adamo, 2003); and this was true even if the target was identical to the penultimate chord (Bigand, Tillmann, Poulin-Charronnat & Manderlier, 2006). Musical context was advantageous for the processing of tonic and disadvantageous for the processing of subdominant targets (Tillmann, Janata, Birk, & Bharucha, 2003). Global priming was shown to be effective for tasks that were not related to musical processing, such as phoneme monitoring[4] (Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001).

Global chord priming has been observed both in musicians and non-musicians; and it is insensitive to acoustical similarity. Therefore, it is reasonable to conclude that global chord priming is governed by cognitive processes that represent the acquired knowledge about chords. In addition, MUSACT (Bharucha, 1987) predicted results of many global chord priming experiments (Bigand & Pineau 1997, Bigand, Madurell, Tillmann & Pineau 1999, Experiment 1; Bigand, Poulin, Tillmann, & D'Adamo, 2003; Tillmann & Bigand, 1999; Tillmann, Bigand & Pineau,1998; Tillmann, Janata, Birk, & Bharucha, 2003). The typical finding from chord priming experiments (both local and global) was that primes closer to the target on the circle of fifths facilitated responses compared to more distant primes. The primary aim of this thesis is to investigate whether it is possible to observe chord priming effects that are contrary to this typical finding.

2.3 MECHANISMS OF PERCEPTION OF CHORDS WITHIN MUSICAL CONTEXT

In this section, mechanisms proposed for explaining effects of tonal context on perception of chords are described. These are principles of chord perception

---

[4] Phoneme monitoring is a task of detecting target phonemes appearing in the stimuli. In this study, target phonemes were heard simultaneously with target chords.

(Bharucha & Krumhansl, 1983), Tonal Pitch Space (Lerdahl, 2001), MUSACT (Bharucha, 1987), and a sequential neural network model (Bharucha & Todd, 1989). All of these mechanisms, except the one proposed by Bharucha & Todd (1989), were designed to cover tonal aspects of perception broader than the chord perception. However, descriptions of mechanism in this section are limited to chord perception.

In order to explain effects of tonality on perception of chords, Bharucha and Krumhansl (1983) introduced three principles of chord perception; namely, contextual identity, contextual distance, and contextual asymmetry. These are functions that map stabilities of chords to relative distance between chords within a hypothetical space of chord perception. Stabilities of chords are determined by the tonal context. Relative distances between chords in this hypothetical space are supposed to reflect perceptual similarities/dissimilarities between chords. The closer two chords are in this space, the more similar they are perceptually, and the easier to be perceptually processed one after the other. In other words, distance in this cognitive space is supposed to reflect perceptual similarity and expectation between chords.

Contextual identity states that the higher the stability of a chord, the closer the cognitive distance between different instantiations of it. For example, two instances of C major chord are closer when they are tonic (sounded in C major tonality) compared to when they are non-diatonic (sounded in F# major tonality). Contextual distance states that the higher the stability of two chords, the closer the cognitive distance between them. For example, C major and G major chords are closer when they are tonic and dominant (sounded in C major tonality), compared to when they are non-diatonic chords (sounded in F# major tonality). Contextual asymmetry covers the effect of temporal order. This principle states that the higher the stability of the second chord, the closer the cognitive distance between them. For example, C major and F# major chords are closer when they are sounded in F# major tonality compared to when they are sounded in C major tonality. Three principles of chord perception formalized the observed relations between chords after a tonal context (Bharucha & Krumhansl, 1983). However, means to implement these functions were not described.

Another theory conceptualized upon a hypothetical cognitive space is Tonal Pitch Space (Lerdahl, 2001). Tonal Pitch Space is a symbolic theory of tonal perception. Symbolic theories describe mental representations as complex symbol structures and mental processes as operations on these representations (Fodor & Phylyshyn, 1988). Tonal Pitch Space (Lerdahl, 2001) provides context sensitive representations for chords and algebraic operations on them. Representations of chords are hierarchical in form. Operations on representations provide measures of chord distance and harmonic attraction.

"Basic Space" provides the form (syntax) of context sensitive representation. Basic space is composed of pitches. There are five levels in this space, which correspond to different levels of stability. "Level E" is the lowest level, "Level A" is the highest. Stable pitches occur at higher levels. "Level E" represents 12 pitch categories. "Level D" represents pitches of a tonality. "Level C" represents pitches of a chord. "Level B" represents the base pitch and the fifth of the chord. "Level A" represents the base pitch of the chord. Lerdahl (2001) depicted basic space on a 12X5 grid (see Figure 2.4, 2.5 and 2.6). On this grid horizontal and vertical axes stand respectively for pitch categories and levels of stability (Figure 2.4, 2.5 and 2.6).

Context sensitive representations of chords are obtained by the combination of "Level D" on the one hand and "Levels C, B and A" on the other hand. "Levels C, B and A" form an undivided unit for context free representation of a chord. "Level D", as a unit by itself, supplies representation a tonality, or context. Context sensitive representation of a chord is obtained by combining "Levels D, C, B and A". For example, context sensitive representations of different of forms C major (i.e. C major as tonic, as a dominant, or as a non-diatonic chord etc.) are obtained by combining the representation of C Major (at "Levels C, B and A") with representations of different tonalities (at "Level D") (see Figures 2.4, 2.5 and 2.6). Typographical description of chords in Tonal Pitch Space (Lerdahl, 2001) is straightforward. Seven upper-case Roman numerals (I, II, III, IV, V, VI, VII) are reserved for major degrees (from tonic to leading); seven lower-case Roman numerals (i, ii, iii, iv, v, vi, vii) are reserved for minor degrees (from tonic to leading); seven upper-case bold letters (**A, B, C, D, E, F, G**) are reserved for major tonalities; seven lower-case bold letters (**a, b, c, d, e, f, g**) are reserved for minor tonalities. Symbols of context sensitive representations are obtained by concatenating numerals and letters with '/'. For example, tonic of C major tonality (C major chord sounded in C major tonality) is represented as I/**C** (see Figure 2.4).

Figure 2.4 shows the Tonal Pitch Space representation for C major chord in C major tonality. C major tonality is represented in Level D by marking the pitches of C major tonality, which are C, D, E, F, G, A, B. The C major chord is represented in Levels C, B and A by marking the pitches of C major chord, which are C, E, and G[5]. Figure 2.5 shows the Tonal Pitch Space representation for the C major chord in F major tonality (dominant of F major tonality). Figure 2.5 is almost identical to Figure 2.4 except Level D now represents pitches of F major tonality, which are C, D, E, F, G, A, A#/Bb. Figure 2.5 shows representation of C major in B major tonality (the Neapolitan sixth of B major); now Level D represents pitches of B major tonality, while Levels C, B and A represents C major chord[6].

---

[5] It is clear from Figure 2.4 that there is a hierarchy between pitches of the C major chord as well. It is better to ignore such a detail because reporting every detail of the theory may difficilitate the understanding of important aspects of the theory.

[6] Pitches of B major tonality are B, C#/Db, D#/Eb, E, F#/Gb, G#/Ab, A#/Bb. However, not all of them are represented in Level D of Figure 2.6. This is because a pitch that is represented in a level on the basic space must also be represented in lower levels as well. C and G are pitches of the C major

level a	o											
level b	o							o				
level c	o				o			o				
level d	o		o		o	o		o		o		o
level e	o	o	o	o	o	o	o	o	o	o	o	o
	C	C#	D	D#	E	F	F#	G	G#	A	A#	B

Figure 2.4 Basic space representation of C major chord in C major tonality (I/**C**).

level a	o											
level b	o							o				
level c	o				o			o				
level d	o		o		o	o		o		o	o	
level e	o	o	o	o	o	o	o	o	o	o	o	o
	C	C#	D	D#	E	F	F#	G	G#	A	A#	B

Figure 2.5 Basic space representation of C major chord in F major tonality (V/**F**).

level a	o											
level b	o							o				
level c	o				o			o				
level d	o			o	o		o	o			o	o
level e	o	o	o	o	o	o	o	o	o	o	o	o
	C	C#	D	D#	E	F	F#	G	G#	A	A#	B

Figure 2.6 Basic space representation of C major chord in B major tonality (bII/**C**).

---

chord; but they are not pitches of the B major tonality. Therefore, C#/Db and G#/Ab is replaced by C and G in Level D.

Tonal Pitch Space (Lerdahl, 2001) provides operations on representations of chords to obtain metrics of cognition. The chord distance rule (Lerdahl, 2001 p. 60) measures similarity between representations of chords. Chord distance reflects perceived similarity between chords. The symbol of chord distance is Δ. Distance between two chords is calculated by the following formula

$$\Delta(\mathrm{V}/\mathbf{X}\rightarrow \mathrm{Z}/\mathbf{Y}) = \mathrm{cof}(\mathbf{X},\mathbf{Y}) + \mathrm{dcof}(\mathrm{V}/\mathbf{X}, \mathrm{Z}/\mathbf{Y}) + \text{half-ham}(\mathrm{V}/\mathbf{X}, \mathrm{Z}/\mathbf{Y}) \qquad (2.1)$$

where cof(**X**,**Y**) is number of steps between tonalities on the circle of fifths; dcof(V/**X**, Z/**Y**) is number of steps between chords on the diatonic circle of fifths (circle of fifths on which accidentals[7] are removed); half-ham(V/**X**, Z/**Y**) is number of o's that is not marked in the basic space representation of the V/**X**, but marked in the basic space representation of the Z/**Y** (see Figures 2.4-6). (This is equal to the half of the hamming distance[8] between representations of two chords.) Distance between chords is symmetric; the higher the distance metric, the more dissimilar the chords. For example, distance between C Major as tonic and as dominant is Δ(I/**C**→V/**F**)=5 (0 + 1 + 4).

The harmonic attraction rule (Lerdahl, 2001 p. 175) compares stability of pitches of two representations. The harmonic attraction rule measures expectation between chords. Attraction between chords is asymmetric; the higher the attraction metric, the more the second chord is expected after the first one. The symbol of harmonic attraction is α_{rh}. Attraction between two chords is calculated by the following formula

$$\alpha_{rh}(\mathrm{V}/\mathbf{X}\rightarrow \mathrm{Z}/\mathbf{Y}) = 10 * \sum^{n} [\alpha(n(\mathrm{V}/\mathbf{X})\rightarrow n(\mathrm{Z}/\mathbf{Y})))] \,/\, \Delta(\mathrm{V}/\mathbf{X}\rightarrow \mathrm{Z}/\mathbf{Y}) \qquad (2.2)$$

---

[7] Accidentals refer to the flat and sharp symbols that appear adjacent to pitch names (see Figure 2.2).

[8] Hamming distance is the minimum number of symbols that must be changed in order to convert one symbol combination into another.

where n(V/X) gives the nth pitch of the chord V/X; $\alpha(x \rightarrow y)$ is attraction from pitch x to pitch y; $\Delta(V/X \rightarrow Z/Y)$ is distance between chord V/X and chord Z/Y; 10 is a constant, whose function is simply to express the distance metric with numbers that are not too small. In other words harmonic attraction is the total attraction between pitches of two chords divided by distance between chords. Attraction between the pitches is calculated by the following formula

$$\alpha(x \rightarrow y) = [\mathrm{level}(y)/\mathrm{level}(x)] \,/\, \mathrm{chroma}(x,y)^2 \qquad (2.3)$$

where level(x) gives the level of representation of the pitch x on the basic space representation of the tonic chord of the tonality of the music; chroma(a,b) is the number of steps between x and y on the chroma circle. For example, attraction between tonic and dominant of C Major is

$$\alpha_{rh}(V/\mathbf{C} \rightarrow I/\mathbf{C}) = 10 * [\alpha(F \rightarrow E)+\alpha(D \rightarrow C)+\alpha(B \rightarrow C)+\alpha(G \rightarrow G)] \,/\, 5 \qquad (2.4)$$

where

$$\begin{gathered} \alpha(F \rightarrow E)=(3/2)/1^2=1.5 \\ \alpha(D \rightarrow C)=(4/2)/2^2=0.5 \\ \alpha(B \rightarrow C)=(4/2)/1^2=2 \\ \alpha(G \rightarrow G)= \mathrm{NULL}^{9} \end{gathered} \qquad (2.5)$$

making $\alpha_{rh}(V/\mathbf{C} \rightarrow I/\mathbf{C}) = 8$.

Tonal Pitch Space is a model for senses of tonal aspects of music (i.e. pitch, chord, and tonality) (Lerdahl, 2001, p. v). Tonal Pitch Space is a musicological theory that

---

[9] According to Lerdahl (2001) there is no attraction between the occurrences of the same pitch, and it is not included in the calculation of harmonic attraction.

is based on results of empirical research on perception of pitch, chord and tonality. In this sense musicology is a branch of theoretical psychology that studies the description of the musical mind. The formulations of the Tonal Pitch Space are both based on empirical results on perception of stability and expectation and intuitions of Fred Lerdahl. The quantization of tonal aspects of music completes the Generative Theory of Tonal Music of Lerdahl & Jackendoff (1983). Generative Theory of Tonal Music describes intuitions of listeners of Western tonal music by proposing symbolic structures and rules to obtain such structures. Generative Theory of Tonal Music focused on the perceptual of organization of musical events. It formulates event hierarchies. Tonal Pitch Space on the other hand formalizes tonal hierarchies, i.e. hierarchical relation between pitches, chords and tonalities[10] (Bharucha, 1984). Generative Theory of Tonal Music (Lerdahl & Jackendoff, 1983) consults tonal hierarchies to formulate event hierarchies. Therefore, Tonal Pitch Space (Lerdahl, 2001) completes the Generative Theory of Tonal Music (Lerdahl & Jackendoff, 1983). Perceptual hierarchies of Generative Theory of Tonal Music and Tonal Pitch Space were in agreement with empirical results (Deliège, 1987; Krumhansl, 1990; Krumhansl, 1996; Lerdahl & Krumhansl, 2001, Lerdahl, Krumhansl, Hanon, & Fineberg, 2000).

MUSACT (MUSical ACTivations) is a connectionist model of tonal harmonic knowledge developed by Bharucha (1987). In connectionist models, there are units to represent mental states; there are links between them to represent associations; and cognitive processes are propagation of signals between units through links. Connectionism was proposed as an alternative to symbolic models of cognition by Rumelhart & McClelland and PDP Research Group (1986). They claimed that mechanisms of the mind are analogous to mechanisms of the brain, not to

[10] Such a distinction may also be expressed in the domain of language as the difference between the fact that sentence is composed of noun and verb phrases; and the fact that "John" is the noun phrase, and "loves Mary" is the verb phrase of the sentence "John loves Mary".

mechanisms of Turing computers. In this sense Turing computers[11] represent a theoretical position of cognition, according to which mental states are complex symbol structures and mental operations operate on syntactic structures of these symbols. Connectionism challenges this proposal.

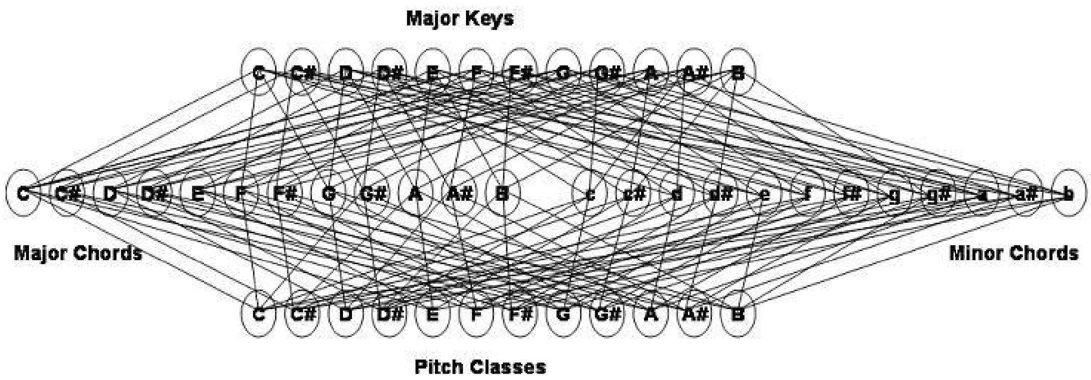


Figure 2.7 MUSACT architecture (Bharucha, 1987).

MUSACT is a network consisting of three groups of interconnected units, which represents pitch categories (12 units), major and minor chords (12 units for each), and major tonalities (12 units) (see Figure 2.7). Every chord unit is connected to the units representing the three pitches contained in that chord. Every major tonality unit is connected to the units representing the three major and three minor chords that are possible in that tonality. All connections are bi-directional. Connections and connection weights of MUSACT are hand-wired. The functioning of the network is based on phasic spreading of activation (see below). Each unit has a level of activation associated with it. The activation levels of pitch units can be changed by external input. All other changes in activation result from activation transmitted through the weighted connections and decay of activation towards zero.

---

[11] There are different senses of computation, as a competence of human beings, as a mathematical formalism, as a simulation technology and as a theoretical position about mind (Beer, 1995). Turing computers are the formal models of the last sense of computation.

Transmission of activation from one unit is proportional to the difference between what the unit holds and receives as a signal, i.e. it is phasic. The spread of activation continues until equilibrium is reached. Equilibrium is defined as an activation change below a threshold value.

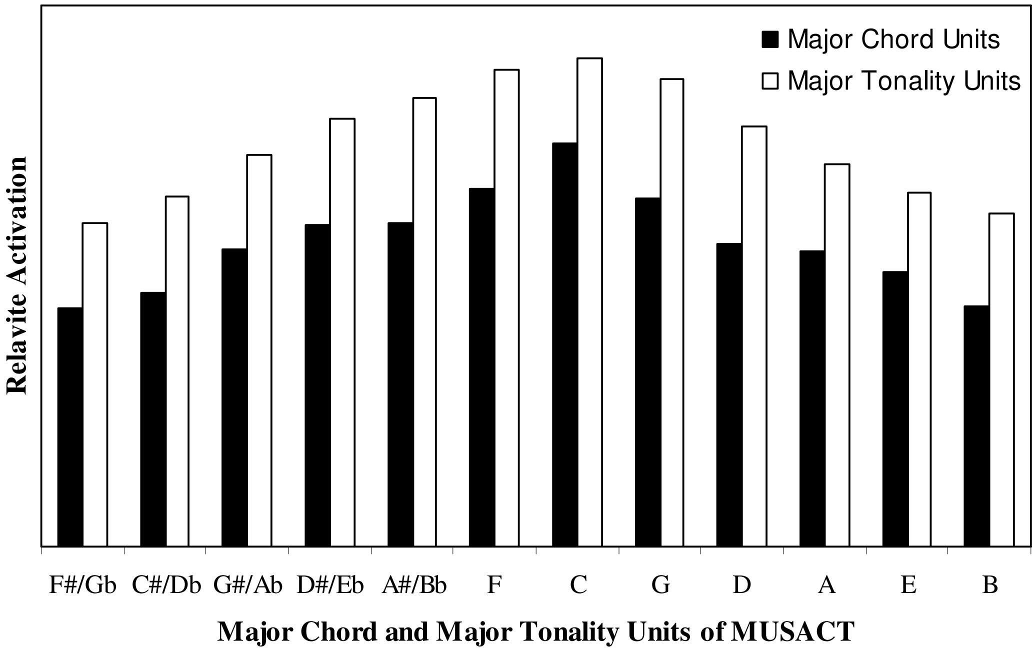


Figure 2.8 Relative activation of major chord and major tonality units of MUSACT (Bharucha, 1987) after C major chord input.

MUSACT (Bharucha, 1987) recovers the circle of fifths in the sense that after giving a major chord as an input, activation levels of major chord and tonality units are inversely related to their distance from the input chord on the circle of fifths (Figure 2.8). In order to predict performance on the chord priming task from MUSACT, it is assumed that targets that have higher activation levels before they are presented will be processed faster and with fewer errors. Tillmann, Bharucha, & Bigand (2000) showed that MUSACT simulates the principles of chord perception, i.e. contextual identity, contextual distance, and contextual asymmetry. As a result MUSACT (Bharucha, 1987) models the perceptual relation and expectation between chords.

Tillmann, Bharucha and Bigand (2000) have demonstrated that MUSACT (Bharucha, 1987) can develop through self organization as a result of exposure to combinations of tones and combinations of chords that are typical in Western music. Tillmann et.al. (2000) trained two self organizing maps consecutively. (See Chapter 5 for a detailed description of self-organizing maps in general; and self-organizing maps of Tillmann et.al. (2000).) The first network learned major and minor chords by being exposed to simultaneous combination of tones. The second network learned major tonalities by being exposed to simultaneous combination of chords, which were represented as the output of the first network. MUSACT (Bharucha, 1987) was obtained by combining the two networks and pruning and scaling of the weights. Pruning and scaling were necessary to reach the equilibrium state (see Chapter 5 for a detailed description of pruning and scaling). Tillmann et.al. (2000) gave accounts of various empirical studies on perception of chords in terms of patterns of activation and their change.

Tillmann et.al. (2000) investigated self organization for simultaneous combinations of tones and chords. Music, on the other hand, is sequential by its nature. In their simulation study, Bharucha and Todd (1989) emphasized the sequential aspects of chord perception. They studied two networks that learned chord sequences to model expectation of chords. Two types of expectations were investigated: schematic (expectations for typical events) and veridical (expectations for specific events).

The first network that Bharucha and Todd (1989) studied was a feed-forward network with three layers. The first and the third layers of the network were composed of six units each, representing three major and three minor degrees. (Number of units in the second layer was not specified.) There were fixed self-feedback connections for the units of the first layer. The network was trained with back-propagation through time algorithm (Rumelhart, Hinton & Williams, 1986). Chord sequences typical of tonal harmonic music (Piston, 1978) were presented to the network. The network was supposed to predict the next chord. Results showed that network extracted bi-gram probabilities of chords of the training set. In other words, it learned schematic expectations for chord sequences (Bharucha & Todd, 1989).

Bharucha & Todd (1989) compared the performance of the first network with the results of Bharucha & Krumhansl (1983). In Bharucha & Krumhansl (1983) participants rated how well a chord sound after another chord[12]. Output activation (activation in the third layer of the network) was correlated with the results of Bharucha & Krumhansl (1983), but Bharucha & Todd (1989) did not give the statistical results of the correlation. The second network of Bharucha and Todd (1989) was the same as the first one, except there were additional units to distinguish different sequences, and the network was supposed to learn the sequences themselves. This network was trained with 50 sequences, all of which is learned successfully. In other words, it learned veridical expectations for chord sequences. Bharucha and Todd (1989) showed that the second network also extracted the schematic expectations.

The behavioral experiments in this thesis challenge the hypothesis that expectations of chords are determined by the stability of chords. This hypothesis is formalized with the principles of chord perception (Bharucha & Krumhansl, 1983). MUSACT (Bharucha, 1987) is used to simulate these principles. Tonal Pitch Space theory (Lerdahl, 2001) also based on the notion of stability. Predictions of Tonal Pitch Space theory (Lerdahl, 2001) and MUSACT (Bharucha, 1987) were compared with the experimental data.

---

[12] This response is supposed to reflect the stability of the second chord given the first chord. Bharucha & Todd (1989) proposed that such a response should also reflect the expectation of chords. This proposal is not rootless given the assumption that stability and expectation of chords are related (Bharucha, 1987; Krumhansl, 1990).

Simulations of this thesis investigated the self-organization of chord schema. Tillmann, Bharucha and Bigand (2000) showed that MUSACT (Bharucha, 1987) may develop as a result of a self-organizing mechanism that is sensitive to statistical properties of events in the environment. Tillmann et.al. (2000) applied pruning and scaling algorithms to the weights of the network after the self-organization step. The possibility of obtaining self-organizing chord schema that simulates important empirical results of chord perception (in other words that is functionally equivalent to MUSACT) without pruning and scaling of weights was investigated.

CHAPTER 3

RESEARCH PROBLEM

In this chapter, the rationale behind studying non-diatonic chords is explained, studies focused on perception of non-diatonic chords are summarized, and the research plan of the thesis work is described.

3.1 THE NEAPOLITAN AND SECONDARY DOMINANT CHORDS

In this section, the merit of studying perception of non-diatonic chords, specifically of the Neapolitan and secondary dominant chords is explained, and these two chords are introduced.

Studies on perception of chords have shown that expectation between chords depends on stability of chords (see Chapter 2; Bharucha, 1987; Bharucha & Krumhansl, 1983; Krumhansl, 1990; Tillmann, Bharucha & Bigand, 2000). The principles of chord perception (Bharucha & Krumhansl, 1983) and MUSACT (Bharucha, 1987; Tillmann et.al., 2000) are based on this proposed relation between stability and expectation (see Section 2.3 for a description of principles of chord perception and MUSACT). The circle of fifths, which is a music theoretical formulation of stability of chords, has been suggested as an explanatory figure for the expectation between chords (Bharucha, 1987; Bigand et.al., 1999; Tillmann et.al., 2000).

On the other hand, tonal factors other than stability may also govern perceptual relation between chords. Best candidates of such factors may be found in music theoretical studies of harmony (relation between chords). Theories of harmony mainly depend on intuitions of music theorists. Intuitions of music theorists and empirical results on music perception have corresponded well in some cases. For example, Lerdahl & Jackendoff (1987) described intuitions of listeners of Western tonal music about hierarchical groupings of musical events, which were confirmed experimentally by Deliège (1987). Perception of pitches and chords, and relation between tonalities was shown to be parallel to the music theoretical descriptions (Kruhmansl, 1990). Music theoretical descriptions between chords (harmony theories) may be seen as perceptual theories that are to be tested empirically.

In theory of harmony, some chord relations are described with factors other than stability. These relations are particularly between specific types of diatonic (stable) and non-diatonic (unstable) chords. In this thesis the role of two non-diatonic chords, the Neapolitan and secondary dominant chords, on the perception of harmonic progressions was investigated. These two non-diatonic chords are associated with specific diatonic chords; and they must precede their diatonic associates. Specific associations and sequential relations between chords can not be reduced to stability of chords. Several studies investigated the perception of the Neapolitan chord (Koelsch, Gunter, Friederici, and Schroger, 2000; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). However, no study so far investigated the perceptual relation between these non-diatonic chords and their diatonic associates. The single exception was the study of the expectation induced by the Neapolitan chord (Atalay, 2002).

The Neapolitan chord is a major triad[1] constructed on the chromatically lowered supertonic degree of a major or minor tonality. It includes pitches that do not belong to the tonality. For example, in C major tonality, the Neapolitan chord is the Db major chord, which consists of the pitches Db/C#, F and Ab/G#. Db/C# and Ab/G# are not members of the C major key. They are members of Ab/G#, Db/C# and Gb/F# major keys. The Neapolitan chord of C major key is the subdominant of the Ab/G# major key, the tonic of the Db/C# major key and the dominant of the Gb/F#

---

[1] Chords composed of three pitches are called triads.

major key. Ambiguity of its membership enables the Neapolitan chord to act as a pivot (bridge) chord for modulation[2]. On the other hand, "it is a chord of strongly subdominant character, progressing most frequently to some form of the dominant chord" (Piston, 1978, p. 401). The third of the Neapolitan chord is the base pitch of the subdominant. The subdominant has a tendency to progress to dominant (Piston, 1978). The base pitch and fifth of the Neapolitan chord are chromatically lowered; and they have a tendency to progress downward, towards the third and fifth of the dominant (Piston, 1978). Remarkably, the Neapolitan chord precedes the dominant chord but the dominant is diametrically opposite to the Neapolitan chord on the circle of fifths. For example, in the tonality of C major the Neapolitan chord is the Db/C# major chord and the dominant is the G major chord. On the circle of fifths Db/C# and G are diametrically opposite (see Figure 3.1). According to music theory, the most distant chord pair on the circle of fifths is associated. In this thesis, the perceptual reality of this association was examined.

Secondary dominants are non-diatonic major or diminished chords. All diatonic major and minor chords are associated with specific secondary dominant chords. If secondary dominant chord is a major chord, the base pitch of the secondary dominant chord is right adjacent to the base pitch of the diatonic chord on the circle of fifths. For example, the secondary dominant of C major is G major; the secondary dominant of G minor is D major (see Figure 3.2). If the secondary dominant chord is a diminished chord, the base pitch of the secondary dominant chord is left adjacent to the base pitch of the diatonic chord on the chroma circle. For example, the secondary dominant of C#/Db major is C diminished; the secondary dominant of D minor is C#/Db diminished (see Figure 3.3).

---

[2] Modulation is the change of tonality. Pivot (bridge) chord initiates the modulation and it is generally member of both tonalities.

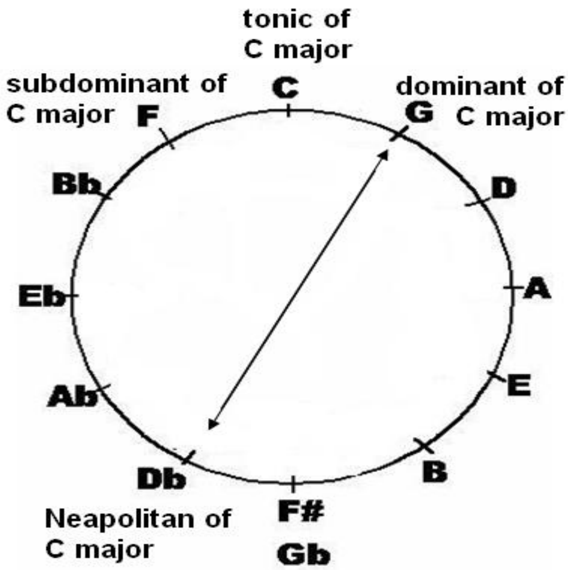


Figure 3.1 Circle of fifths. Members of the circle represent major chords. Major chords of C major are marked on the circle.

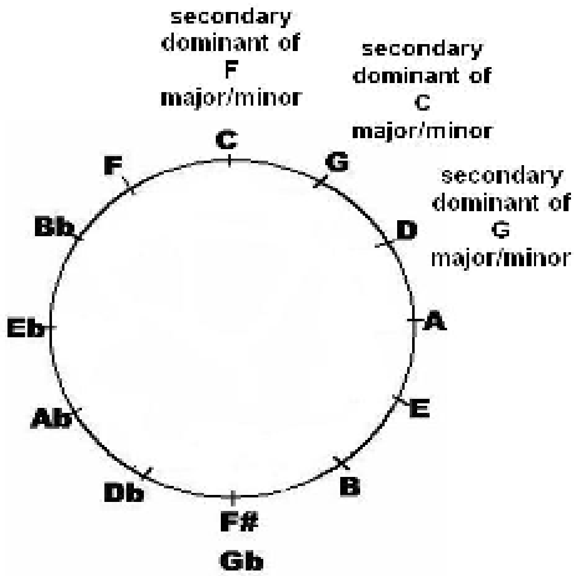


Figure 3.2 Circle of fifths. Members of the circle represent major chords. Relation between major secondary dominant chords and their associate major/minor chords are marked for some members of the circle. This relation holds for unmarked members as well.

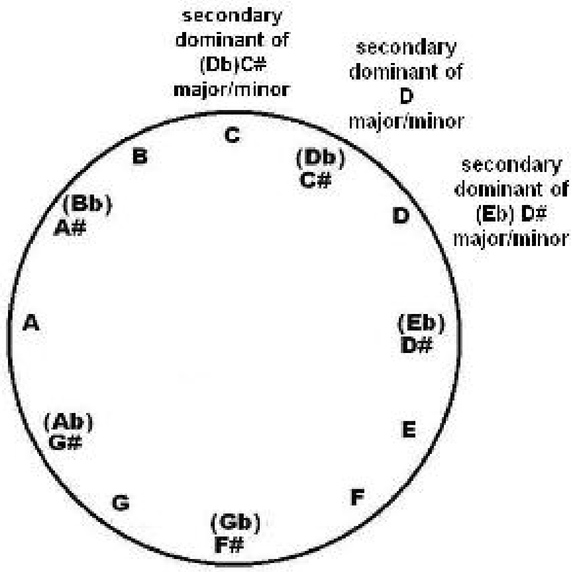


Figure 3.3 Chroma circle. Members of the circle represent diminished chords. Relation between diminished secondary dominants and their associate major/minor chords are marked for some members of the circle. This relation holds for unmarked members as well.

Secondary dominant chords do not disturb the sense of tonality (Piston, 1978). The relation between dominant and tonic is similar to the relation between secondary dominant and its diatonic associate. Therefore, secondary dominant chords are supposed to temporarily tonicize[3] the following diatonic chord (Piston, 1978). The importance of the secondary dominant chord comes from the fact that it must precede, and it must not succeed, the associated diatonic chord (Piston, 1978, p.246). According to music theory, there is a sequential order between the secondary dominant chord and the associated diatonic chord. In this thesis, the perceptual reality of the sequential relation between the secondary dominant chord and its diatonic associate was examined.

---

[3] Chords that are tonicized become as stable as the tonic. Chord that are temporarily tonicized become more stable for a short amount of time.

The music theoretical relation between the Neapolitan chord and the dominant is important psychologically. The Neapolitan chord is a non-diatonic chord. Dominant chord is a diatonic chord. The Neapolitan and the dominant are diametrically opposite on the circle of fifths. Perceptually, dominant is one of the most stable chords; and the Neapolitan chord is one of the most unstable chords. According to the principles of chord perception (Bharucha & Krumhansl, 1983) and MUSACT (Bharucha, 1983), the stability of chords determine the expectation between chords; therefore there should be no expectation towards the dominant after the Neapolitan chord. It is important to investigate this disagreement between music theory and music psychology. In addition, the investigation of the perceptual relation between the Neapolitan and the dominant enables to compare predictions of Tonal Pitch Space theory (Lerdahl, 2001) and MUSACT (Bharucha, 1983) with novel stimuli.

The music theoretical relation between the secondary dominant chord and its diatonic associate is also important psychologically. Music theory specifies a sequential order between the secondary dominant chord and its diatonic associate. None of the theories of chord perception describe a sequential order between specific chord types. It is important to investigate this relation music psychologically; and test the predictions of Tonal Pitch Space theory (Lerdahl, 2001) and MUSACT (Bharucha, 1983).

3.2 STUDIES ON THE NEAPOLITAN CHORD

In this section, studies that investigated the perception of the Neapolitan chord are summarized.

In the field of neuropsychology of music, the perception of the Neapolitan chord was a popular topic for a while. The Neapolitan chord is the most distant chord to the tonic on the circle of fifths. Therefore it is one of the most unstable chords. Focusing on this property, neuropsychologists recorded and located neural activations elicited by listening to the Neapolitan chord. By this way, they investigated the correlation of brain activity and stability of chords.

Patel, Gibson, Ratner, Besson, & Holcomb (1998) compared Event Related Potentials[4] (ERP) elicited by stable, moderately-unstable, and extremely-unstable chords embedded in chord sequences. Stable chords were tonic chords; extremely-unstable chords were Neapolitan chords; moderately-unstable chords were chords three steps away from the tonic on the circle of fifths. Patel et.al. (1998) reported a late positive component, whose amplitude was inversely related to the stability of the chord; and right antero-temporal negativity elicited by unstable chords. According to Patel et.al. (1998) these components signal the fit of the chord into the musical context.

Koelsch, Gunter, Friederici, and Schroger (2000) investigated ERP components elicited by the Neapolitan chord. They composed chord sequences, in which the Neapoliten chord was either the third or the fifth chord. When it was the third chord, it followed the tonic and preceded the dominant chord; when it was the fifth chord, it followed the dominant chord. The Neapolitan chord elicited early right anterior negativity and late positivity. Amplitudes of these components were higher for the Neapolitan chord at the fifth position compared the Neapolitan at the third position. Koelsch et.al. (2000) suggested that early right anterior negativity signal the violation of the expectation of the listeners, which increased as the length of the sequence increased; and the late positivity signal the musical integration processes. It was shown that the early right anterior negativity signal was elicited preattentively (Koelsch, Schroger, & Gunter, 2002), and generators of it were located at Broca's

[4] Event Related Potential in this thesis, refers to the electric potentials of human brain that are induced by an external stimuli and recorded above the skull. Event Related Potentials reflect brain processes. Event Related Potentials are important tools of cognitive psychology.

area and its right-hemisphere homologue (Maess, Koelsch, Gunter, & Friederici, 2001).

The above mentioned studies did not examine the distinctive relation between the Neapolitan chord and the dominant. The distinctive role of the Neapolitan chord in perception was suggested by Horton (2002). In this experiment dominant-tonic chord pairs (cadances[5]) in different tonalities were presented following chord progressions. The task of the participants was to rate how well each cadence completed the progression. Progressions were of two types: Either the progression started on the tonic and moved to the tonality five steps counterclockwise on the circle of fifths (the tonality of the Neapolitan chord) or it moved to the tonality five steps clockwise. For example, in the former case a progression starting on C major would move to C#/Db major and in the latter case a progression that started on C#/Db major would move to C major. In the former case cadences in both the initial tonality (G-C) and the second tonality (G#/Ab –C#/Db) were rated higher compared to neighboring tonalities on the circle of fifths. In the latter case, however, only the cadence in the second tonality (G-C) received higher ratings. The progression in the latter case could be interpreted as starting on the Neapolitan chord and moving to the home tonality. This would not have created expectancy for returning to the initial tonality.

Atalay (2002) investigated expectation for the dominant chord after the Neapolitan chord. The Neapolitan and dominant chords are closely associated according to music theory; but they are the most distant pair on the circle of fifths. The typical finding from chord priming experiments has been that primes closer to the target on the circle of fifths facilitated responses compared to more distant primes (Bharucha & Stoeckig, 1986; Bigand et. al., 1999; Tekman & Bharucha, 1992, 1998).

---

[5] In music theory, dominant-tonic chord pair at the end of a chord sequence is called cadence. Cadences induce the sense of musical ending.

Atalay (2002) investigated the expectation for the dominant chord induced by single chords and chord sequences. For the processing of the dominant target, Atalay (2002) observed the advantage of the closest prime over the most distant one when the prime was a single chord; and the elimination of this advantage when the prime was a chord sequence and the penultimate (next to last) chord was the Neapolitan chord. In addition, Atalay (2002) investigated effects of different types of penultimate chords on the processing of dominant target. Chord progressions were kept identical except for the penultimate chord, which was realized with the Neapolitan chord (the most distant), the closest chord, or with chords at intermediate distances from the target on the circle of fifths. Atalay (2002) observed comparable facilitative effects of the Neapolitan chord (the most distant) and the closest chord on the processing of the dominant target. The Neapolitan chord and the closest chord created an advantage for the processing of the dominant target over chords three-steps-away from the dominant.

3.3 CURRENT STUDY

In this thesis work, the psychological reality of the music theoretical relations between the Neapolitan chord and the dominant, and between the secondary dominant chord and its diatonic associate was investigated. These relations are specific relations between diatonic (stable) and non-diatonic (unstable) chords. They specify a sequential order between chord types.

Perceptual relations between chords have been investigated with memory tasks (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha & Castelleno, 1982), similarity ratings (Bharucha & Krumhansl, 1983; Krumhansl, 1990; Krumhansl, Bharucha & Castelleno, 1982; Krumhansl, Bharucha & Kessler, 1982;) and priming studies (Bharucha, 1987; Bharucha & Stoeckig, 1986, 1987; Bigand & Pineau, 1997; Bigand, et. al. 1999; Tekman & Bharucha, 1992, 1998).

In this thesis not all aspects of perception of chord relations are investigated. The focus was given to expectation. Expectation was studied with priming paradigm. (See Section 2.2 for description of the paradigm.) The nature of expectation induced by the Neapolitan chord for the dominant was investigated by comparing conditions with different prime and target types, by eliminating melodic factors, by introducing critical chords after different musical contexts, and by comparing performances of musically educated and uneducated participants (See Sections 4.1, 4.2 and 4.3). The perceptual reality of the relation between secondary dominant chords and diatonic associates was investigated by its effect on global priming. Global priming is facilitative processing of tonic targets over subdominant targets (Bigand & Pineau, 1997; Bigand, et. al. 1999). The effect of the perceptual relation between the secondary dominant chord and its diatonic associate on the global chord priming was investigated by scrambling sequences, and by eliminating melodic factors (See Sections 4.4, 4.5 and 4.6).

The first three experiments of this thesis work investigated the expectation for the dominant after the Neapolitan chord. Experiment 1 (Section 4.1) was conducted to investigate whether the relation between the Neapolitan and the dominant is unique, or it is simply based on the distance on the circle of fifths. Experiment 2 (Section 4.2) was conducted to investigate whether the expectation towards the dominant after the Neapolitan chord is observed when the melodic factors are removed. Experiment 2 was conducted in Finland with European participants. Experiment 3 (Section 4.3) was conducted with Turkish participants with stimuli similar to the Experiment 2. This was done to investigate effects of cultural factors on the results of Experiment 2. In addition, in Experiment 3 the necessity of listening to chord sequences before the Neapolitan-dominant chord pair in order to observe the relation between the Neapolitan and the dominant was investigated.

The last three experiments of this thesis work investigated the perception of the relation between the secondary dominant and its diatonic associate. This was done by observing the global priming effect (the facilitative responses given to tonic targets compared to subdominant targets). Experiment 4 (Section 4.4) was conducted to investigate whether the global priming effect is reduced or eliminated when secondary dominant chords succeed their diatonic associates (i.e. when the secondary dominant – diatonic associate relation is ill-formed). Experiment 5 (Section 4.5) was conducted to investigate whether it is possible to observe the disruption of global priming effect due to the ill-formed secondary dominant chords with more natural musical stimuli. Experiment 6 (Section 4.6) was conducted to investigate whether the global priming effect is reduced or eliminated when secondary dominant chords and their diatonic associates are put far apart in chord sequences.

Simulations in this thesis were conducted to investigate whether self-organizing models of chord schema are obtained without pruning and scaling of weights that Tillmann, Bharucha, Bigand (2000) performed. In addition, effects perceptual relations between pitches on the performance of self-organization of chord schema were investigated.

CHAPTER 4

BEHAVIORAL EXPERIMENTS

In this chapter, behavioral experiments conducted to investigate effects of the Neapolitan and secondary dominant chords on the perception of harmonic progressions are described. Predictions of MUSACT and Tonal Pitch Space are compared with the results of the experiments.

4.1 EXPERIMENT 1

Music theoretically, the dominant is associated with the Neapolitan chord (Piston, 1978) and the two are six-steps away on the circle of fifths. Atalay (2002) showed that the Neapolitan chord created an advantage for the processing of dominant over non-diatonic chords which were three-steps away.

Experiment 1 was conducted to replicate this finding and to investigate whether the six-steps relation is critical for priming of chords after musical context. By comparing expectations for dominant and subdominant targets induced by chords three- and six-steps away from them, the unique relation between dominant and the Neapolitan chord can be observed. Subdominant is not associated with a non-diatonic chord. Therefore, expectations for subdominant should be governed by distance on the circle of fifths (Bharucha, 1987). Dominant is associated with the Neapolitan chord (Atalay, 2002). Therefore, there should be expectation for the dominant after the Neapolitan chord.

In order to test these hypotheses, sequences of seven chords were used. The last chord (the target) of sequences was either the dominant or the subdominant of the tonality. The chord preceding the target (the penultimate chord) was either three or six steps away from the target on the circle of fifths. The first five chords ("tonic, subdominant, dominant, tonic, tonic" in the given order) were identical in all conditions. Therefore, differences in expectations for the targets were induced by the penultimate chords. If the target was the dominant, penultimate six steps away was the Neapolitan chord (see Figure 4.1). The Neapolitan chord and the dominant are music theoretically associated (Piston, 1978). For dominant targets the Neapolitan chord should create an advantage. There is no specific association between the penultimate chords and the subdominant. Therefore, for subdominant targets priming should be governed by the distance on the circle of fifths; the penultimate chord three-steps away should create and advantage over the penultimate chord six-steps away for the processing of subdominant.

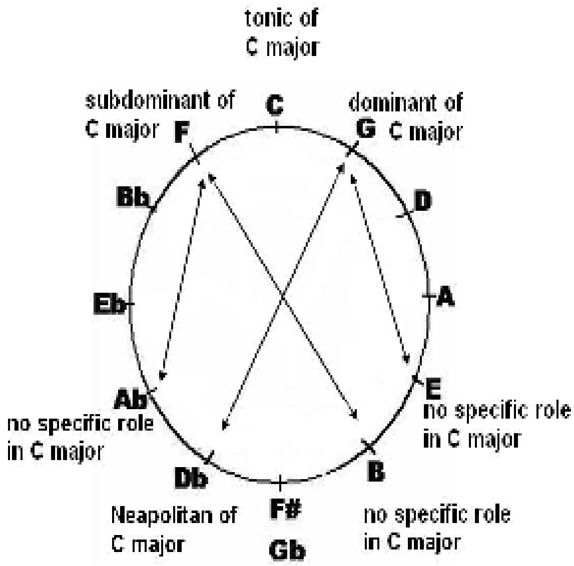


Figure 4.1 Description of the conditions of Experiment 1 on the circle of fifths.

Predictions of MUSACT (Bharucha, 1987) for the stimuli of Experiment 1 are shown in Figure 4.2. The network was implemented with the code published in the Appendix of Bigand et.al. (1999) (see Appendix A for the code). Activation of the target chord before it was presented to MUSACT modeled expectation for the target. The higher the activation, the greater the expectation. According to MUSACT

(Bharucha, 1987) expectation for dominant and subdominant targets should be governed by distance between penultimate and target on the circle of fifths.

Predictions of the Tonal Pitch Space Theory (Lerdahl, 2001) are shown in Figure 4.3. Tonal Pitch Space provided the harmonic attraction rule (Lerdahl, 2001 p. 175) to calculate the expectation between chords. The higher the harmonic attraction score, the greater the expectation. Harmonic attraction from the Neapolitan chord to the dominant is higher α_{rh}(bII/**I**→V/**I**)=6 compared to the attraction from the chord three-steps-away to the dominant α_{rh}(V/**vi**→V/**I**)=4.8. Harmonic attraction from the chord six-steps-away to the subdominant is higher α_{rh}(VI/**i**→IV/**I**)=1.2 compared to the attraction from the chord three-steps-away to the subdominant α_{rh}(V/**iii**→IV/**I**)=0.7. According to the Tonal Pitch Space Theory (Lerdahl, 2001) for both subdominant and dominant targets the penultimate chord six-steps-away should facilitate responses compared to the penultimate chord three-steps-away.

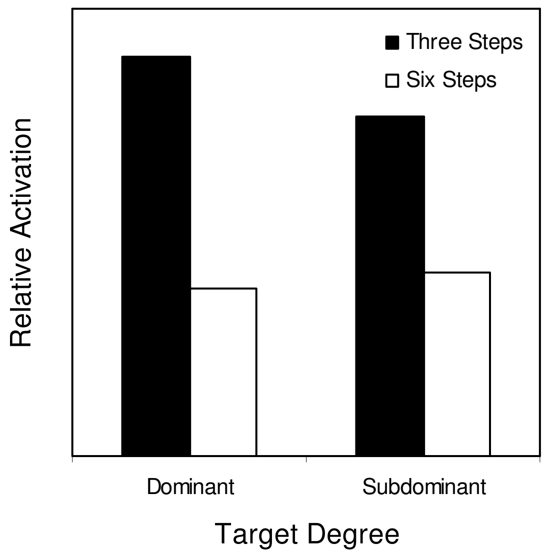


Figure 4.2 Predictions of MUSACT (Bharucha, 1987) for the conditions of Experiment 1.

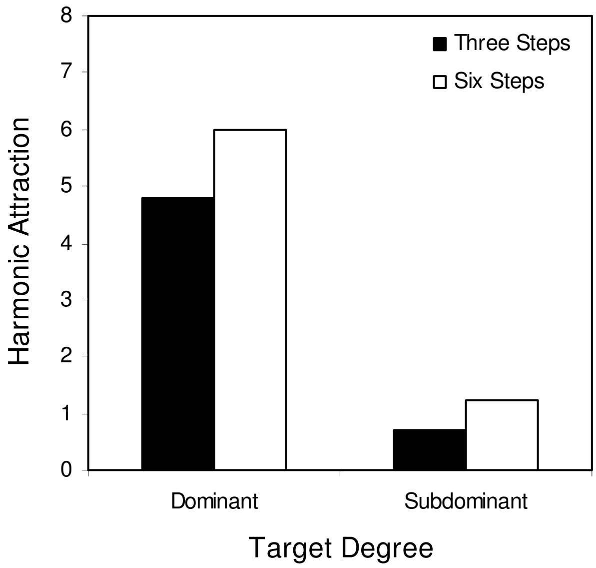


Figure 4.3 Predictions of Tonal Pitch Space (Lerdahl, 2001) for the conditions of Experiment 1.

In chord priming studies, different discrimination tasks have been assigned to participants, such as in-tune/out-of-tune (Bharucha & Stoeckig, 1986, 1987; Tekman & Bharucha, 1998), consonance/dissonance (Bigand & Pineau 1997, Bigand, Madurell, Tillmann & Pineau 1999), or timbre discrimination (Tillmann, Bigand,Escoffier, Lalitte, 2006). In Turkish non-musicians, the chord priming effect has been observed with the consonance/dissonance discrimination task (Atalay, 2002). Turkish non-musicians had difficulties in making an in-tune/out-of-tune discrimination (Atalay, 2002). Consequently, the consonance/dissonance discrimination was chosen as the task of Experiment 1.

4.1.1 METHOD

4.1.1.1 PARTICIPANTS

Thirty two participants participated in Experiment 1. Eleven participants were students of the Ankara State Conservatory of Hacettepe University. They had 11.45 years of musical training on average with a range of 6 to 19. Their average age was 23.36 years with a range of 20 to 35. The remaining 21 participants were students of the Middle East Technical University and they had 0.37 years of musical training on average with a range of 0 to 2 years. Their average age was 21.38 with a range of 18 to 24.

4.1.1.2 STIMULI

Stimuli of Experiment 1 were sequences of four part chords. Stimuli were composed with MIDI piano timbre; and MIDI was converted to WAVE format with TiMidity++ software[1]. Each chord lasted 1000 milliseconds (msec), except for the last chord (the target chord) which was played for 2000 msec. Sequences consisted of seven chords: five diatonic chords of a major tonality (henceforth global context) followed by one non-diatonic chord (penultimate chord) followed by target chord.

The target was the dominant or the subdominant. The penultimate chord was either three-steps or six-steps away from the target on the circle of fifths (on clockwise direction for dominant and counterclockwise direction for subdominant targets). For example, if the dominant target was the G major chord, it was preceded by either the E major chord (three-steps) or the C#/Db major chord (six-steps, the Neapolitan chord) (see Figure 4.1). If the subdominant target was the F major chord, it was preceded by either the G#/Ab major chord (three-steps) or the B major chord (six-steps). Global context (see above) was composed of "tonic, subdominant, dominant, tonic, tonic" played in the given order. Scores of example stimuli are given in Appendix B.

Reaction times and accuracy were recorded in a consonance/dissonance discrimination task for the target chords. Consonant targets were major chords. Dissonant targets were augmented chords, whose fifth degree was raised one semitone compared to the major chord. For example, the dissonant target contrasted with the C major (C, E,G) was composed of C, E, and G#/Ab. Chord sequences

---

[1] http://www.onicos.com/staff/iz/timidity/

were repeated from 12 tonalities, once with a consonant target and once with a dissonant target, which made 96 sequences in total.

4.1.1.3 PROCEDURE

Participants sat facing a computer screen. Musically uneducated participants participated in the experiment in the Experimental Psychology Laboratory of the Department of Cognitive Science, METU. Musically educated participants were recruited in their institutions. Test procedure and task demands (consonance vs. dissonance distinction) were explained to participants.

The experiment started with a test session in which participants listened to single chords and decided whether they were consonant or dissonant. Musically uneducated participants gave their responses by pressing one of two designated response keys on the response box of Cedrus SuperLab Pro software. Musically educated participants gave their responses by pressing one of two designated response keys on a Compaq Presario 2800 laptop. Feedback was given on the accuracy of each response on a computer screen. Participants listened to 28 consonant and 28 dissonant chords mixed randomly.

A practice session followed the test session. In the practice session participants listened to chord sequences; and they were instructed to judge the consonance of the final chord. A set of question marks appeared on the computer screen together with the final chord, which required a response, to make it more distinct. Responses were given with the same keys of the test part. Participants listened to 24 chord sequences selected from the stimuli of the experiment. The practice trials were selected from among the experimental trials such that equal numbers of trials of each condition were repeated between the practice session and the experiment. Responses of this session were not included the analysis.

The experimental session was the same as the practice session, except that responses given in the experiment session were included the analysis. Reaction times and accuracy were recorded for the target chords. Responses were given on the same keys. There were 96 sequences in the experimental session. Stimuli were repeated in twelve tonalities. Two types of penultimate (three-steps vs. six-steps), factored by two types of target degree (dominant vs. subdominant) factored by two types of target timbre (Shepard vs. saw-tooth) factored by twelve tonalities made 96 stimuli in total. Stimulus presentation and response collection were controlled by Cedrus SuperLab Pro software. Stimuli were presented through Sennheiser HD-433 headphones.

4.1.2 RESULTS AND DISCUSSION

The average reaction time (RT) of correct responses and the percentage of error (PE) were calculated for each participant. Group averages of musician and non-musician participants for different conditions of the experiment are presented in Figures 4.4 and 4.5, respectively. RT and PE data were analyzed with separate Analyses of Variance (ANOVA) with musical education (musician vs. non-musician) as a between subject factor, and target degree (dominant vs. subdominant), penultimate distance (three vs. six steps), and target type (consonant vs. dissonant) as within subject factors. Only significant main effects and interactions are reported.

RTs revealed a significant effect of musical education [$\underline{F}(1, 30) = 4.762$, $\underline{MS}_e = 303625.244$, $\underline{p} < .05$, $\eta^2 = .14$]. Musicians responded faster (824 msec.) than non-musicians (982 msec.) on average. There was a main effect of target type [$\underline{F}(1, 30) = 7.773$, $\underline{MS}_e = 24541.952$, $\underline{p} < .01$, $\eta^2 = .21$]. Dissonant targets were responded faster (875 msec.) compared to consonant targets (932 msec.). There was an interaction between target degree and target type [$\underline{F}(1, 30) = 8.555$, $\underline{MS}_e = 7197.425$, $\underline{p} < .01$, $\eta^2 = .22$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 57 ms.] subdominant dissonant targets were responded faster (858 msec.) than subdominant consonant (948 msec.) and dominant consonant targets (916 msec.). There was a significant three-way interaction between target degree, penultimate distance and target type [$F(1, 30) = 18.717$, $MS_e = 4719.293$, $p < .001$, $\eta^2 = .38$] and four-way interaction between musical education, target degree, penultimate distance and target type [$F(1, 30) = 5.412$, $MS_e = 4719.293$, $p < .05$, $\eta^2 = .15$].

In order to examine the four-way interaction, RTs of musicians and non-musicians were analyzed with separate ANOVAs with target degree (dominant vs. subdominant), penultimate distance (three vs. six steps) and target type (consonant vs. dissonant) as within subject factors. RTs of musicians revealed a significant main effect of target type [$F(1, 10) = 5.025$, $MS_e = 19313.009$, $p < .05$, $\eta^2 = .33$]. Musicians responded to dissonant targets faster (791 msec.) compared to consonant targets (857 msec.). For musicians, there was also a significant three-way interaction between target degree, penultimate distance and target type [$F(1, 10) = 15.855$, $MS_e = 5018.427$, $p < .01$, $\eta^2 = .61$].

In order to examine the three-way interaction in musicians' data, RTs of consonant and dissonant targets were analyzed with separate ANOVAs with target degree (dominant vs. subdominant) and penultimate distance (three vs. six steps) as within subject factors. For consonant targets, RTs of musicians revealed significant interaction between target degree and penultimate distance [$F(1, 10) = 9.225$, $MS_e = 4730.250$, $p < .05$, $\eta^2 = .48$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 89 ms] after a penultimate chord six-steps away from the target musicians responded faster for consonant dominant targets compared to (812.13 msec.) consonant subdominant targets (903.58 msec.).

RTs of non-musicians revealed a significant interaction between target degree and target type [$F(1, 20) = 6.780$, $MS_e = 7707.757$, $p < .05$, $\eta^2 = .26$]. Posthoc comparison did not reveal [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 75 ms] any significant differences between pairs of means.

PEs revealed a significant effect of musical education [$\underline{F}(1, 30) = 77.857$, $\underline{MS}_e = .063$, $\underline{p} < .001$, $\eta^2 = .72$]. Non-musicians made more errors (% 34) than musicians. (% 5). There was a significant interaction between musical education and target type [$\underline{F}(1, 30) = 4.652$, $\underline{MS}_e = .027$, $\underline{p} < .05$, $\eta^2 = .13$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = %16] responses of musicians for consonant (%7) and dissonant targets (%3) were significantly different from responses of non-musicians for dissonant (%37) and consonant (%31) targets. There was a significant main effect of target degree [$\underline{F}(1, 30) = 4.959$, $\underline{MS}_e = .014$, $\underline{p} < .05$, $\eta^2 = .14$]. Subdominant targets were responded to with more error (%22) compared to dominant targets (%18). There was a significant main effect of penultimate distance [$\underline{F}(1, 30) = 4.457$, $\underline{MS}_e = .009$, $\underline{p} < .05$, $\eta^2 = .05$]. Targets that followed the six-steps-away-penultimate were responded to with more errors (%21) compared to targets that followed three-steps-away-penultimate (%18). There was an interaction between target degree and penultimate distance [$\underline{F}(1, 30) = 9.125$, $\underline{MS}_e = .015$, $\underline{p} < .01$, $\eta^2 = .23$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = %8] subdominant targets that followed the six-steps-away penultimate chords were responded to with more error (%25) compared to dominant targets followed three-steps-away penultimate (%17). There was also a three way interaction between target degree, penultimate distance and target type [$\underline{F}(1, 30) = 14.573$, $\underline{MS}_e = .037$, $\underline{p} < .01$, $\eta^2 = .33$]. In order to examine the three-way interaction, PEs for consonant and dissonant targets were analyzed with separate ANOVAs with musical education (musician vs. non-musician) as between subject factor, and target degree (dominant vs. subdominant) and penultimate distance (three vs. six steps) as within subject factors. PEs for consonant targets revealed a significant effect of musical education [$\underline{F}(1, 30) = 54.017$, $\underline{MS}_e = .032$, $\underline{p} < .001$, $\eta^2 = .64$]. Non-musicians responded to consonant targets with more error (%31) compared to musicians (%7). There was a significant main effect of target degree [$\underline{F}(1, 30) = 6.150$, $\underline{MS}_e = .025$, $\underline{p} < .05$, $\eta^2 = .17$]. Subdominant consonant targets were responded to with more errors (%23) compared to dominant consonant targets (%16). There was a two-way interaction between target degree and penultimate distance [$\underline{F}(1, 30) = 27.906$, $\underline{MS}_e = .022$, $\underline{p} < .001$, $\eta^2 = .48$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = %10] consonant subdominant targets that followed six-steps-away penultimate chords were responded to with more error (%31) compared to consonant subdominant targets that followed three-steps-away penultimate (%15). Contrarily, consonant dominant targets followed three-steps-away penultimate were responded with more error (%22) compared to consonant dominant targets followed six-steps-away penultimate (%9).

PE data showed that for both musicians and non-musicians processing of subdominant targets was facilitated by penultimate chord three-steps-away compared to penultimate chord six-steps-away; and for dominant targets, the advantage of a closer over a distant penultimate chord reversed. RT data for musicians were parallel with PE data. For non-musicians there were no significant effect in the RT data parallel to the PE data, but the pattern of RT for non-musicians did not conflict with PE data. Both musicians and non-musicians associated the Neapolitan chord with dominant; and this association had nothing to do with the six-steps relation.

In order to understand the relation between the musical experience and the chord priming, the correlation between the years of musical training and the difference between reaction times for the conditions of the experiments was calculated. For each participant, average reaction time for the dominant target that followed the penultimate three-steps away and average reaction time for the subdominant target that followed the penultimate six-steps were added; and from this the average reaction time for the dominant target that followed the penultimate six-steps away and average reaction time for the subdominant target that followed the penultimate three-steps were subtracted. The result was correlated with the years of musical training. The correlation was not significant (r= 0.244, p>0.05, N=32). The same calculation and correlation was done for the percentage of error of participants. This correlation was not significant either (r= -0.142, p>0.05, N=32).

Predictions of MUSACT (Bharucha, 1987) and Tonal Pitch Space (Lerdahl, 2001) were partially correct. Responses to subdominant targets were in agreement with predictions of MUSACT and contrary to the predictions of Tonal Pitch Space. Responses to dominant targets were in agreement predictions of Tonal Pitch Space and contrary to the predictions of MUSACT.

In Experiment 1, results for dominant targets showed that Turkish musicians and non-musicians perceived the relation between the Neapolitan chord and the dominant when chords were presented with the piano timbre. Tonal Pitch Space (Lerdahl, 2001), which represents the Neapolitan chord within the tonal context, was successful in predicting the result. For subdominant targets, results showed that distance on the circle of fifths governed the chord priming. MUSACT (Bharucha, 1987), which implements the circle of fifths, predicted the priming for subdominant targets. The priming between the Neapolitan chord and the dominant seem exceptional among the general rule of the distance on the circle of fifths for chord priming. Turkish musicians and non-musicians perceived the relation between the Neapolitan and the dominant.

The predictions of MUSACT (1987) and Tonal Pitch Space (Lerdahl, 2001) were contradictory for the stimuli in Experiment 1. MUSACT does not represent pitch proximity (Bharucha 1987 p. 23), which is an important aspect of tonal harmony. However, pitch proximity is represented in harmonic attraction rule of Tonal Pitch Space (Lerdahl, 2001, p.175). Therefore, it may be the case that this difference caused the contradiction between the predictions of MUSACT (1987) and Tonal Pitch Space (Lerdahl, 2001). This possibility was investigated in Experiment 2.

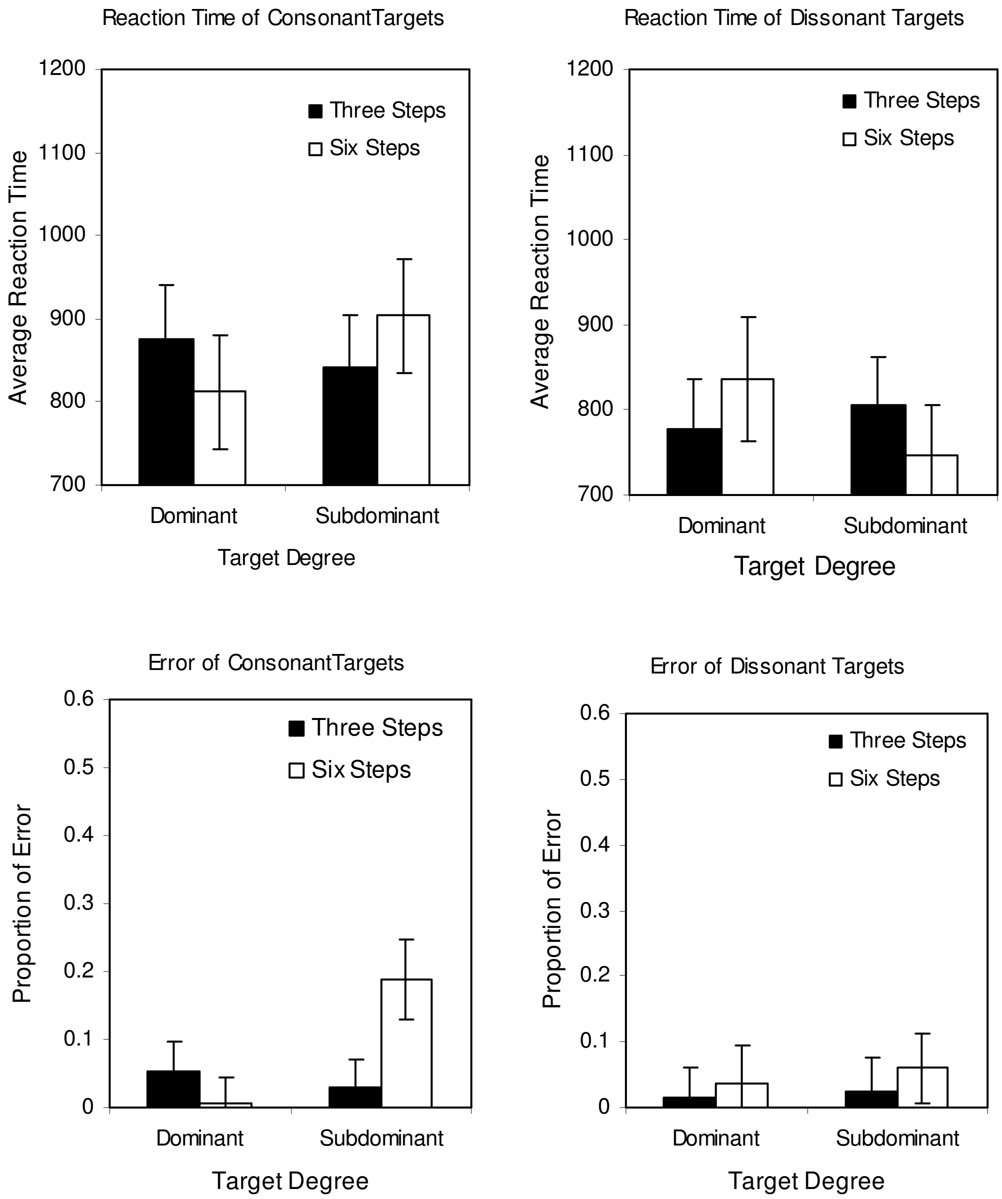


Figure 4.4 Average reaction time and percentage error of musicians for consonant and dissonant targets for the conditions of Experiment 1.

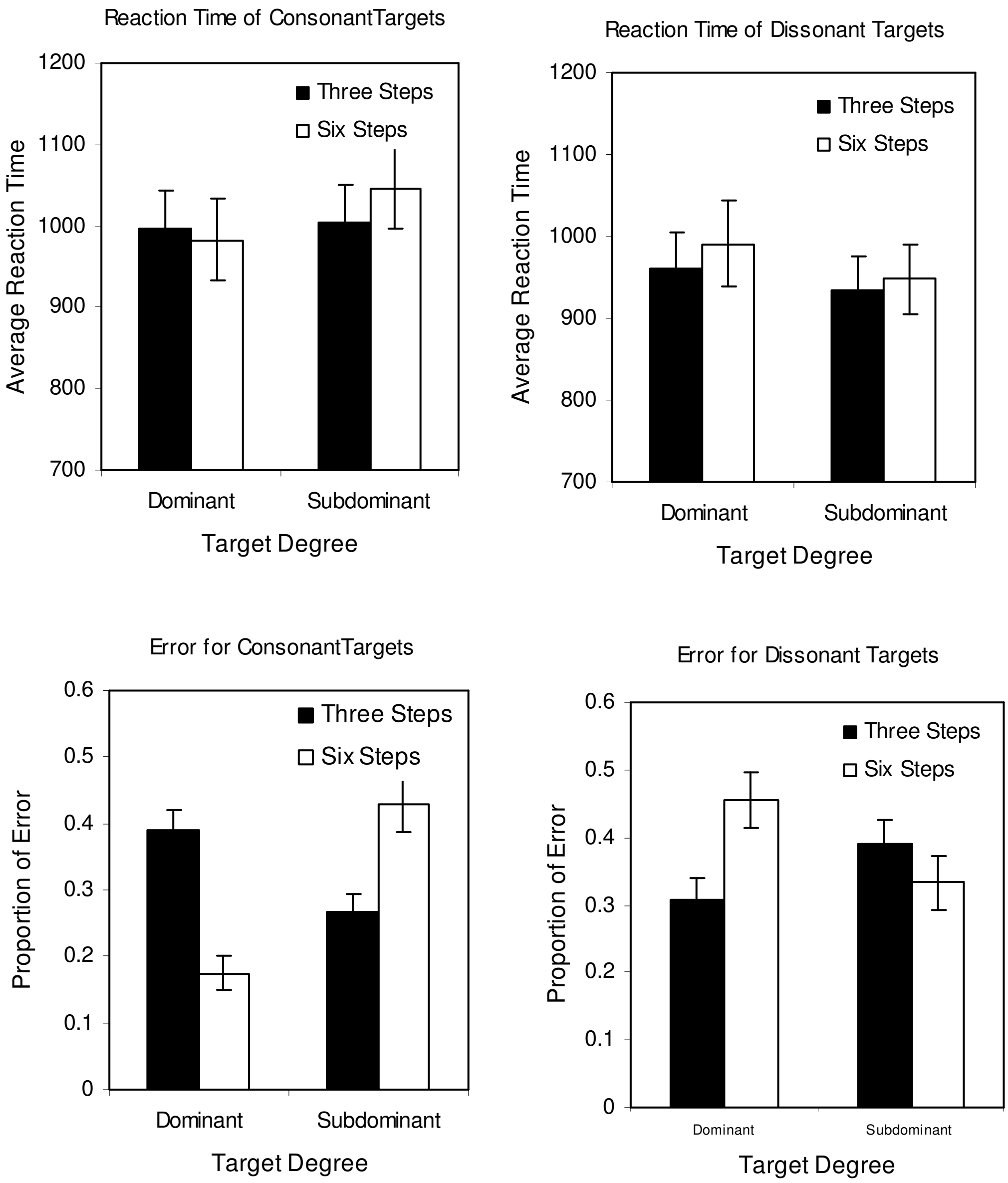


Figure 4.5 Average reaction time and percentage error of non-musicians for consonant and dissonant targets for the conditions of Experiment 1.

4.2 EXPERIMENT 2

Experiment 1 confirmed the unique perceptual relation between the Neapolitan chord and dominant, and established that this relation was not based on the six-steps relation between the two chords. On the other hand, there were limitations of the study.

In Experiment 1, one type of chord sequence was used as a global context. This put doubts on the generalizability of the results. It was necessary to investigate the effect of different instantiations of the global context on the perception of the relation between the Neapolitan sixth chord and the dominant. In addition, the piano timbre was used in Experiment 1. Stimuli composed with the piano timbre contain melodic factors (i.e. pitch height) that might effect tonal expectation (Bharucha, 1996). Therefore, it may be the case that perceptual relation between the Neapolitan chord and dominant was the result of melodic features but not of the association between the two chords per se. This possibility also can doubt the comparability of the predictions of MUSACT (Bharucha, 1987) and Tonal Pitch Space (Lerdahl, 2001). Tonal Pitch Space is governed with representations to encode melodic relations between chords but MUSACT is not. In Tonal Pitch Space (Lerdahl 2001), the calculation of expectation (or harmonic attraction) between two chords is based on the stability of pitches that compose the chords, and the distance between pitches on the chroma circle, apart from other factors (see Section 2.3 for the calculation of harmonic attraction). Stability of pitches of the chords and distance between pitches on the chroma circle calculate the melodic factors of expectation (or attraction). In MUSACT (Bharucha, 1987), expectation of a chord is predicted by the activation of the node that represent the chord; and there is no way to introduce stability of pitch and distance between pitches as operative factors of the activation of the node. Therefore, comparison of predictions of the MUSACT and the Tonal Pitch Space (Lerdahl, 2001) may fall short.

In Experiment 2 various types of chord sequences were used as global context; and chord sequences were played with Shepard tones (Shepard, 1964). Shepard tones create an auditory illusion of the circularity of pitch by emphasizing pitch class at the expense of pitch height (Section 2.2 for a description of this illusion). This illusion might reduce effects of melodic factors due to the pitch height on the results of the study. In Experiment 2, penultimate chords were either two- or six-steps away from dominant and subdominant targets; and subjects were instructed to make an intune/out-of-tune discrimination.

The tasks in Experiment 1 and Experiment 2 were different. In Experiment 1 a consonant/dissonant discrimination task was used, because Turkish non-musicians had difficulties in making an intune/out-of-tune discrimination (Atalay, 2002). In Experiment 2 non-musician participants were Europeans. The task was switched to the intune/out-of-tune discrimination in order to be able to replicate findings with a different task. Because chord priming has been observed with both tasks (for example Bharucha & Stoeckig, 1986, 1987; Tekman & Bharucha, 1998 used intune/out-of-tune discrimination; Bigand & Pineau 1997, Bigand, Madurell, Tillmann & Pineau 1999 used consonant/dissonant discrimination), the results from both experiments would be comparable.

Predictions of MUSACT (Bharucha, 1987) for the stimuli of Experiment 2 are shown in Figure 4.6. Simulation of the network and interpretation of the results were done in the same way as in Experiment 1. According to MUSACT (Bharucha, 1987), expectations for dominant and subdominant should be governed by distance between penultimate and target on the circle of fifths.

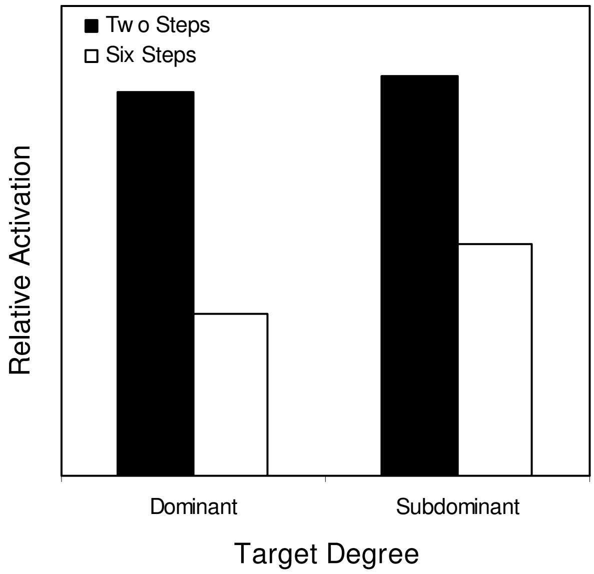


Figure 4.6 Predictions of MUSACT (Bharucha, 1987) for the conditions of Experiment 2.

Predictions of Tonal Pitch Space Theory (Lerdahl, 2001) are shown in Figure 4.6. It is reasonable to use the chord distance rule (Lerdahl, 2001 p. 60) to calculate expectation between chords without definite melody. The smaller the chord distance is, the greater the relation and the expectation between chords. The chord distance between the dominant and the Neapolitan chord (six-steps-away to the dominant) is smaller Δ(bII/**I**→V/**I**)=9 compared to the distance between the dominant and the chord two steps away from the dominant Δ(V/**ii**→V/**I**)=11. The chord distance between the subdominant and the chord six-steps-away to the subdominant is smaller Δ(bII/**I**→V/**I**)=10 compared to the distance between the subdominant and the chord three steps away from the subdominant Δ(bII/**I**→V/**I**)=14. According to Tonal Pitch Space (Lerdahl, 2001) expectations for dominant and subdominant targets should be enhanced after the six-steps-away penultimate compared to the three-steps-away penultimate.

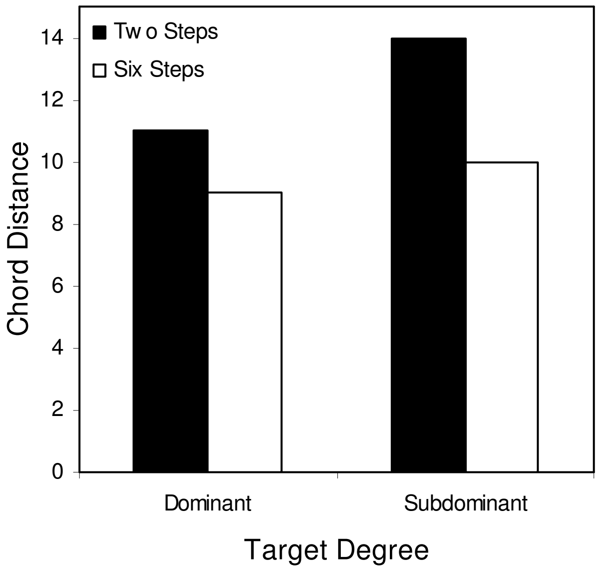


Figure 4.7 Predictions of Tonal Pitch Space (Lerdahl, 2001) for the conditions of Experiment 2.

4.2.1 METHOD

4.2.1.1 PARTICIPANTS

Twenty six participants participated in the experiment. All of them were students of Jyväskylä University, Jyväskylä, Finland. Sixteen of them took part in the experiment for course credit; the remaining ten participants were given free lunch tickets for their participation. The average age of participants was 25.9 with a standard deviation of 4.6 years. The range of age was 20 to 37 years. In order to investigate effects of musical training, participants were split into two groups (musically educated and musically uneducated) depending on years of formal musical training. The cut off point was 6.5 years, which was the median of the variable. Participants with more than 6.5 years of formal training years were classified as musically educated. Their average years of musical training was 12.8 with a standard deviation 5.08, and the range was from 7 to 25 years. For musically naive participants the average number of years of musical training was 1.7, the standard deviation was 2.01 years, and the range was from 0 to 6 years.

4.2.1.2 STIMULI

Stimuli of Experiment 2 were sequences of chords. Major and minor chord composed of Shepard tones (Shepard, 1964) were utilized. A program, which was based on the description of Shepard chords by Krumhansl, Bharucha and Kessler (1982), was implemented to generate them. Each chord was played for 1000 miliseconds (msec), except for the last chord (the target chord) which was played for 2000 msec. Sequences consisted of seven chords: five diatonic chords of a major tonality (henceforth global context) followed by one non-diatonic chord (penultimate chord) followed by target chord.

The target was the dominant or the subdominant of the tonality of the sequence. The penultimate chord was either two-steps or six-steps away from the target on the circle of fifths (clockwise for dominant and counterclockwise for subdominant targets). In Experiment 2, six-step penultimate was contrasted with two-step penultimate because the target and the penultimate two-steps away do not share a pitch. For example, if the dominant target was the G major chord, it was preceded by either the A major chord (two-steps) or the C#/Db major chord (six-steps, the Neapolitan chord) (see Figure 4.8). If the subdominant target was the F major chord, it was preceded by either the D#/Eb major chord (two-steps) or the B major chord (six-steps).

In the experiment, four different sequences were used as global context. All of them were devised to induce a sense of tonality but all ended with a different degree of tonality. Sequences utilized as global context were a) "tonic, subdominant, dominant, tonic, tonic"; b) "supertonic, subdominant, dominant, tonic, subdominant"; c) "subdominant, dominant, tonic, tonic, supertonic"; and d) "subdominant, tonic, dominant, tonic, submediant".

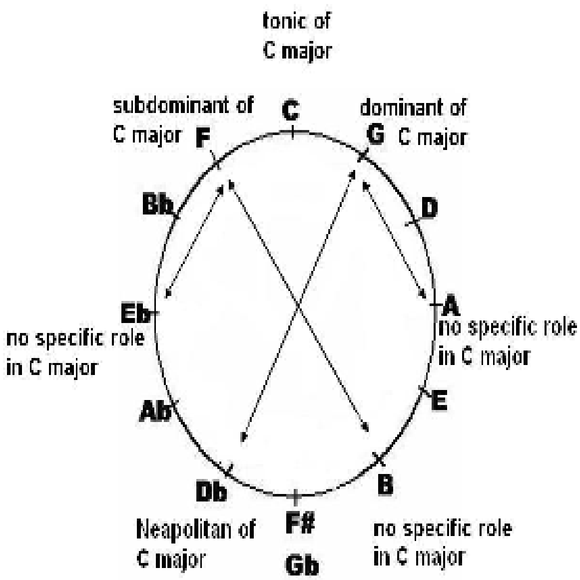


Figure 4.8 Description of the conditions of Experiment 2 on the circle of fifths.

Stimuli were repeated in four tonalities, namely C Major, A Major, F#/Gb Major and D#/Eb Major. Four types of global context factored by two types of penultimate (2 steps vs. 6 steps), factored by two types of target degree (dominant vs. subdominant) factored by two types of target tuning (intune vs. out-of-tune) factored by four tonalities made 128 sequences in total.

Reaction times and accuracy were recorded in a intune/out-of-tune judgment task for the target chords. Intune targets were major chords. Out-of-tune targets were obtained by multiplying the frequency of the fifth degree of the major chord with the factor $2^{-1/48}$.

4.2.1.3 PROCEDURE

Participants sat facing a computer screen. Test procedure and task demands (in-tune vs. out-of-tune distinction) were explained to participants. They listened to one in-tune and one out-of-tune chord example. Experiment started with a test session in which participants listened to single chords and decided whether they were intune or out-of-tune. Responses were given by pressing one of two keys on the computer keyboard, designated by red and yellow stripes. Feedback on the accuracy of each response was given on the computer screen. Participants listened to 8 intune and 8 out-of-tune chords mixed randomly for each subject.

The experimental session followed the test session. In the experimental session participants listened to chord sequences; and they were instructed to judge the intonation of the final chord. A set of question marks appeared on the computer screen together with the final chord, which required a response, to make it more distinct. Responses were given on the same keys used during the test part. Reaction times and accuracy were recorded in the in-tune/out-of-tune discrimination task for the target chords. Participants responded to 128 chord sequences in this session. Stimulus presentation and response collection were controlled by Cedrus SuperLab Pro software. Participants heard the stimuli through headphones.

4.2.2 RESULTS AND DISCUSSION

The average reaction time (RT) of correct responses and the percentage of error (PE) were calculated for each participant. Group averages of musically educated and naive participants for different conditions of the experiment are presented in Figures 4.9 and 4.10 respectively. RT and PE data were analyzed in separate Analyses of Variance (ANOVA) with musical education (musically educated vs. musically naive) as a between subject factor, and target degree (dominant vs. subdominant), penultimate distance (two vs. six steps), and target type (in-tune vs. out-of-tune) as within subject factors. Only the significant main effects and interactions were reported.

RTs revealed a significant effect of target type [$F(1, 24) = 21.723$, $MS_e = 24675.733$, $p < .001$, $\eta^2 = .48$]. Out-of-tune targets were responded to faster (804 msec.) compared to intune targets (906 msec.). Main effect of target degree was barely significant. [$F(1, 24) = 4.090$, $MS_e = 2584.092$, $p < .054$, $\eta^2 = .15$]. There was a significant three-way interaction between target degree, penultimate distance and target type [$\underline{F}(1, 24) = 7.194$, $\underline{MS}_e = 3336.924$, $\underline{p} < .05$, $\eta^2 = .23$].

In order to examine the three-way interaction, RTs of in-tune and out-of-tune targets were analyzed in separate ANOVAs with musical education (musically educated vs. musically naive) as a between subject factor, and target degree (dominant vs. subdominant) and penultimate distance (two vs. six steps) as within subject factors. RTs of in-tune targets revealed a significant interaction of target degree and penultimate distance [$\underline{F}(1, 24) = 5.902$, $\underline{MS}_e = 4975.721$, $\underline{p} < .05$, $\eta^2 = .18$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 53 ms] consonant subdominant targets were responded to faster after penultimate two-steps away (875 msec.) compared to penultimate six-steps away (916 msec.).

PEs revealed a significant effect of musical education [$\underline{F}(1, 24) = 14.759$, $\underline{MS}_e = .118$, $\underline{p} < .001$, $\eta^2 = .38$]. Musically naive participants made more errors (% 21) than musically educated participants (% 2). There was a main effect of target type [$\underline{F}(1, 24) = 12.297$, $\underline{MS}_e = .013$, $\underline{p} < .01$, $\eta^2 = .34$] and an interaction of musical education and target type [$\underline{F}(1, 24) = 6.769$, $\underline{MS}_e = .013$, $\underline{p} < .05$, $\eta^2 = .34$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = %8] musically naive participants responded to in-tune chords (%26) with more error than out-of-tune chords (%16). For musically educated participants there was no such difference (average percentage of error was %3 for in-tune and %2 for out-of-tune chords). There was a two-way interaction between penultimate distance and target type [$\underline{F}(1, 24) = 6.678$, $\underline{MS}_e = .012$, $\underline{p} < .05$, $\eta^2 = .22$]; and three-way interaction between musical education, penultimate distance and target type [$\underline{F}(1, 24) = 6.678$, $\underline{MS}_e = .012$, $\underline{p} < .05$, $\eta^2 = .22$].

In order to examine the three-way interaction, PEs of musically educated and naive participants were analyzed with separate ANOVAs with target degree (dominant vs. subdominant), penultimate distance (two vs. six steps) and target type (in-tune vs. out-of-tune) as within subject factors. PEs of musically naive participants revealed a significant main effect of target type [$\underline{F}(1, 12) = 10.122$, $\underline{MS}_e = .024$, $\underline{p} < .01$, $\eta^2 = .46$]. Musically naive participants responded to in-tune targets with more error (%25) compared to out-of-tune targets (%15). There was a significant two-way interaction between penultimate distance and target type [$\underline{F}(1, 12) = 7.143$, $\underline{MS}_e = .022$, $\underline{p} = .05$, $\eta^2 = .37$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = %17] percentage of error for intune targets (%31) was significantly different from out-of-tune targets (%13) after six-steps-away penultimate chords. PEs of musically educated participants did not reveal significant main effect or interaction.

PE data showed that response patterns of musically educated and naive participants were different, and for musically naive participants processing of dominant and subdominant targets was facilitated by the penultimate three-steps-away compared to penultimate six-steps-away. Contrarily, RT data suggested differential facilitation for dominant and subdominant targets with respect to penultimate distances, regardless to musical education.

RT data did not reveal significant differences between response patterns of musically educated and naive participants. However, in order to examine the possibility of observing different response patterns from musically educated and naive participants, which was suggested by the PE data, RT data of musically educated and musically naive participants were analyzed with separate ANOVAs with target degree (dominant vs. subdominant), penultimate distance (two vs. six steps), target type (in-tune vs. out-of-tune) as within subject factors.

RTs of musically naive participants revealed a significant main effect of target type [$\underline{F}(1, 12) = 6.104$, $\underline{MS}_e = 31789.197$, $\underline{p} < .05$, $\eta^2 = .34$]. Musically naive participants responded to out-of-tune targets faster (859 msec.) compared to in-tune targets (945 msec.). There were no other significant main effects or interactions.

RTs of musically educated participants revealed significant main effect of target degree [$\underline{F}(1, 12) = 6.437$, $\underline{MS}_e = 1154.841$, $\underline{p} < .05$, $\eta^2 = .35$]. Subdominant targets were responded to faster (799 msec.) compared to dominant targets (816 msec.) on average. In addition there was a main effect of target type [$\underline{F}(1, 12) = 20.169$, $\underline{MS}_e = 17526.269$, $\underline{p} < .001$, $\eta^2 = .62$]. Out-of-tune targets were responded to faster (750 msec.) compared to in-tune targets (866 msec.). There was a three-way interaction between target degree, penultimate distance and target type [$\underline{F}(1, 12) = 8.477$, $\underline{MS}_e = 3089.810$, $\underline{p} < .05$, $\eta^2 = .41$].

In order to examine the three-way interaction, RTs of musically educated participants for in-tune and out-of-tune targets were analyzed with separate ANOVAs with target degree (dominant vs. subdominant), penultimate distance (two vs. six steps) as within subject factors. For in-tune targets there was a two-way interaction between target type and penultimate distance [$\underline{F}(1, 12) = 7.567$, $\underline{MS}_e = 3218.818$, $\underline{p} < .05$, $\eta^2 = .39$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 66 ms] musically educated participants responded to subdominant in-tune targets faster after the two-steps-away penultimate (815 msec.) compared to all other in-tune targets (i.e. the subdominant targets after six-steps-away penultimate (887 msec.), the dominant targets after six-steps-away penultimate (874 msec.), and the dominant targets after two-steps-away (889 msec.) penultimate). For out-of-tune targets there was a two-way interaction between target type and penultimate distance [$\underline{F}(1, 12) = 4.983$, $\underline{MS}_e = 1063.869$, $\underline{p} < .05$, $\eta^2 = .29$]. Posthoc comparison did not reveal [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 37 ms] any significant differences.

In order to understand the relation between the musical experience and the chord priming, the correlation between the years of musical training and the difference between reaction times for the conditions of the experiments was calculated. For each participant, average reaction time for the dominant target that followed the penultimate two-steps away and average reaction time for the subdominant target that followed the penultimate six-steps were added; and from this the average reaction time for the dominant target that followed the penultimate six-steps away and average reaction time for the subdominant target that followed the penultimate two-steps were subtracted. The result was correlated with the years of musical training. The correlation was not significant ($r=0.088$, $p>0.05$, $N=26$). The same calculation and correlation was done for the percentage of error of participants. This correlation was not significant either ($r=-0.153$, $p>0.05$, $N=26$).

Though, the RT data did not reveal significant differences between response patterns of musically educated and naive participants, separate analysis of RT data of musically educated and naive participants suggested a picture parallel to the one found with PE data. RT data of musically naive participants did not reveal critical significant main effects and interactions. On the other hand, RT data of musically educated participants revealed critical significant main effects and interactions, which supported the perceptual association between the Neapolitan and the dominant.

In summary, responses of musically educated and naive participants for subdominant targets were facilitated by penultimate chord three-steps-away compared to penultimate chords six-steps-away. On the other hand, responses of musically educated and naive participants were different for dominant targets. For musically naive participants, facilitation was still based on distance between penultimate and dominant target. For musically educated participants distant penultimate had an advantage over the closer one for the processing of the dominant chord. MUSACT (Bharucha, 1987) predicted responses of musically naive participants and responses of musically educated participants to subdominant targets well. Tonal Pitch Space (Lerdahl, 2001) predicted only responses of musically educated participants to dominant targets.

In Experiment 2, musically naive participants did not perceive the relation between the Neapolitan and the dominant. This result was different from Experiment 1, where non-musicians perceived the relation. However, musically educated participants perceived the relation between the Neapolitan chord and the dominant in Experiment 2. This result was repeated in Experiment 1. There were important differences between Experiment 1 and Experiment 2. Participants of Experiment 2 were Europeans, whereas participants of Experiment 1 were Turkish. Stimuli of Experiment 2 were played with Shepard tones; stimuli of Experiment 1 were played with piano tones.

Results of Experiment 2 showed that only European musicians associated the Neapolitan and the dominant chords, when they were presented with the Shepard tones, which diminished the effects of the melodic factors on the results. For subdominant targets their responses were governed by the distance on the circle of fifths. In other words, the exceptional association between Neapolitan and the dominant was learned by musicians. Responses of non-musicians for dominant and subdominant targets were governed by the circle of fifths, which shows that when chords are isolated from melodic factors, non-musicians had difficulties in perceiving the relation between the Neapolitan and the dominant.

Non-musician participants in Experiment 1 were Turkish; and musically naive participants of Experiment 2 were European. Turkish musical culture is a melodic culture. Melodic factors were effective in the stimuli of Experiment 1. It may be the case that the difference between the results of Experiment 1 and Experiment 2 were due to the cultural background of participants. This hypothesis was investigated in Experiment 3.

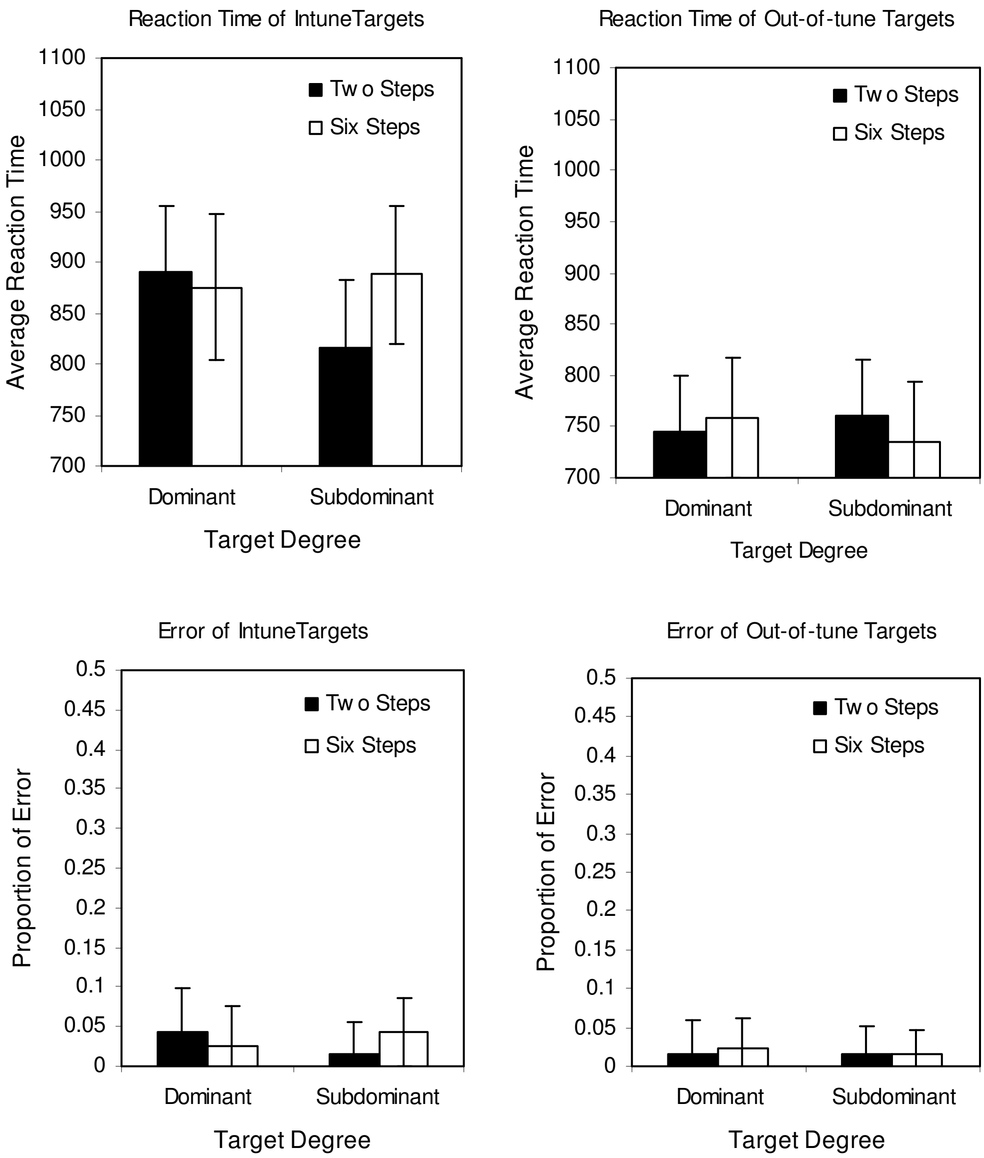


Figure 4.9 Average reaction time and percentage error of musically educated participants for intune and out-of-tune targets for the conditions of Experiment 2.

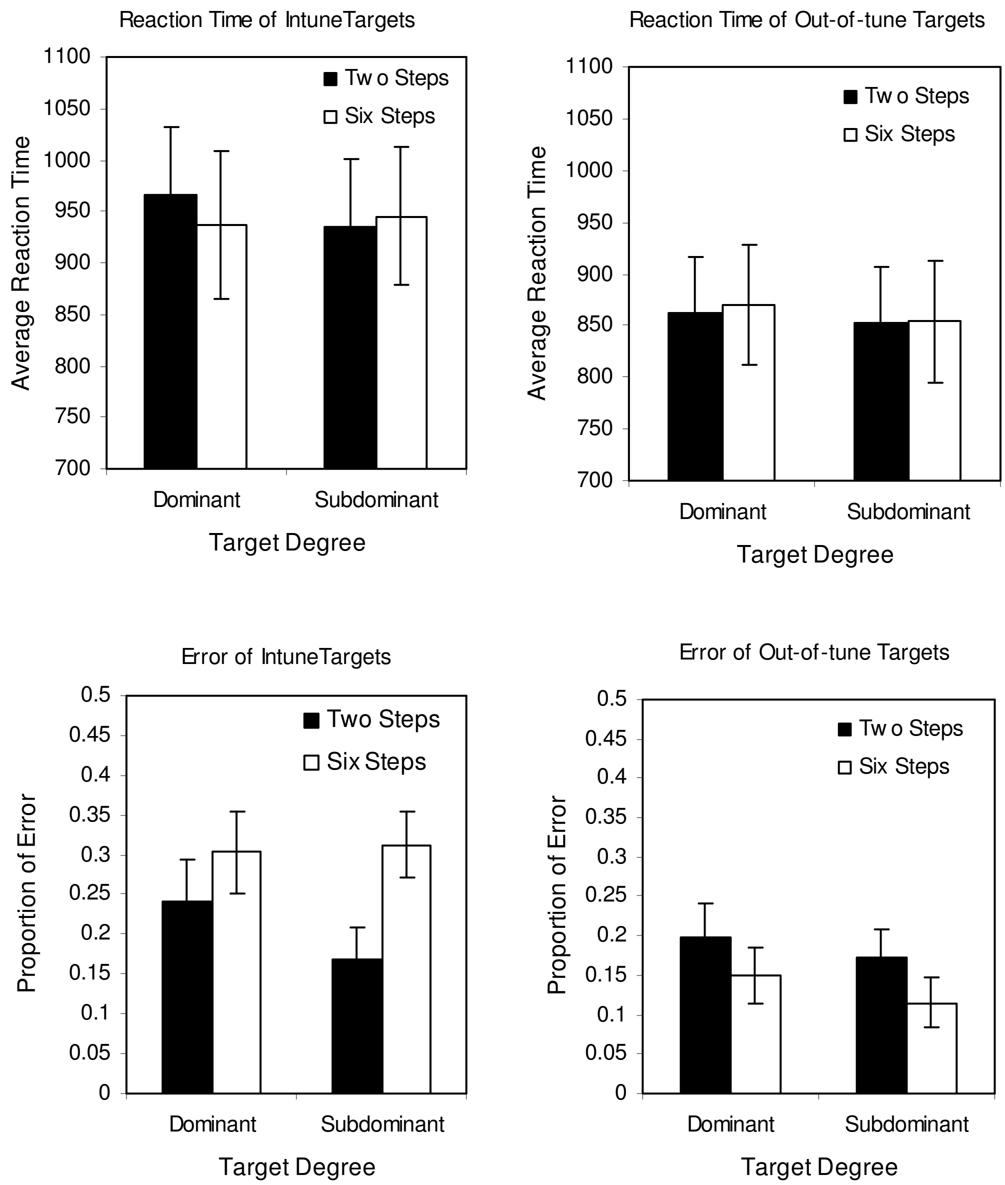


Figure 4.10 Average reaction time and percentage error of musically naive participants for intune and out-of-tune targets for the conditions of Experiment 2.

4.3 EXPERIMENT 3

Experiment 2 established that perception of the relation between the Neapolitan chord and dominant was unique to musically educated participants when melodic factors were reduced. Experiment 2 was conducted in Finland with European participants. Experiment 1 established that when melodic factors were operating, both musically educated and uneducated participants perceived the relation between the Neapolitan chord and the dominant. Experiment 1 was conducted in Turkey with Turkish participants. In order to understand effects of melodic factors on the relation between the Neapolitan chord and the dominant definitely, it was necessary to test musically uneducated Turkish subjects with stimuli composed of Shepard chords (Shepard, 1964). Experiments 1 and 2 so far provided preliminary evidence supporting the hypothesis that there is a perceptual association between the Neapolitan chord and dominant, especially for musically educated participants; and stability was not sufficient to explain this relation. On the other hand, stability was necessary because it was stability that gave a chord the identity of Neapolitan, dominant or any other function. Stability of chords is determined by tonality, which in turn is determined by musical context. In Experiment 1 and 2, chord sequences were utilized as contexts of chord priming. Chord sequences were not the only way to induce tonality perception. For example, Krumhansl and Kessler (1979) utilized scales to investigate the effect of context on perception of pitches. Therefore, it would be interesting to compare scales and chord sequences as contexts of chord priming.

In Experiment 3 scales and chord sequences were utilized as global context (see Section 4.3.1.1). Shepard tones (Shepard, 1964) were utilized to reduce effects of melodic factors on results, as it was the case in Experiment 2. Subjects were students of Middle East Technical University. Penultimate chords were either three- or six-steps away from dominant and subdominant targets.

The task of Experiment 3 was timbre discrimination, in which Shepard and saw-tooth timbres were discriminated. We conjectured that timbre discrimination task would be easier for Turkish non-musician participants compared to the consonance/dissonance discrimination (the task of Experiment 1) and in-tune/out-of-tune discrimination (the task of Experiment 2).Since chord priming has already been observed with timbre discrimination (Tillmann, Bigand,Escoffier, Lalitte, 2006), the results of the three experiments would easily be comparable.

Predictions of MUSACT (Bharucha, 1987) for the stimuli of Experiment 3 are shown in Figure 4.11. Simulation of the network and interpretation of the results were done in the same way as in Experiment 1, except the time interval parameter (implemented with the variable ti in the code of MUSACT; see Appendix A for the code of MUSACT) was adjusted for scale contexts. The time interval parameter of MUSACT sets decay with respect to the duration of the input. The smaller the duration of input is, the less the decay is applied to activations. The time interval parameter of MUSACT was set to 1 by default. Some pitches of scale context were half in duration (see 4.3.1.1 for detailed description of stimuli of Experiment 3). For those pitches the time interval parameter was set to 0.5. According to MUSACT (Bharucha, 1987), expectations for the dominant and the subdominant should be governed by distance on the circle of fifths.

Predictions of the Tonal Pitch Space Theory (Lerdahl, 2001) are shown in Figure 4.12. It is reasonable to use the chord distance rule (Lerdahl, 2001 p. 60) to calculate expectation between chords without definite melody. According to the Tonal Pitch Space Theory distance between chords does not change with the context they reside in. Therefore, chord distance between penultimate and target is equal for scales and context. The chord distance between the dominant and the Neapolitan chord (six-steps-away to the dominant) is equal $\Delta(\text{bII}/\mathbf{I} \rightarrow \text{V}/\mathbf{I})=9$ to the distance between the dominant and the chord three-steps-away from the dominant $\Delta(\text{V}/\mathbf{vi} \rightarrow \text{V}/\mathbf{I})=9$. The chord distance between the subdominant and the chord six-steps-away from the subdominant is smaller Δ(VI/**i**→V/**I**)=10 compared to the distance between the subdominant and the chord three-steps-away from the subdominant Δ(V/**iii** V/**I**)=13. According to Tonal Pitch Space (Lerdahl, 2001) there should be no difference in the expectation for the dominant target after three- or six-steps away penultimate chords. For subdominant targets the penultimate chord six-steps-away should facilitate responses compared to the penultimate three-steps-away.

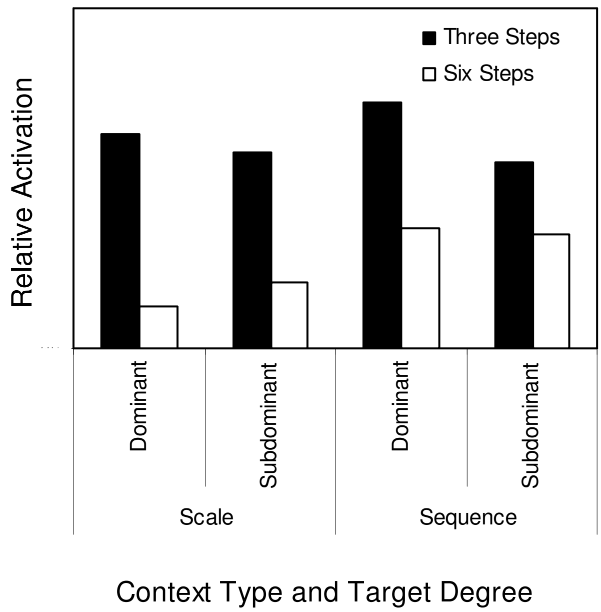


Figure 4.11 Predictions of MUSACT (Bharucha, 1987) for the conditions of Experiment 3.

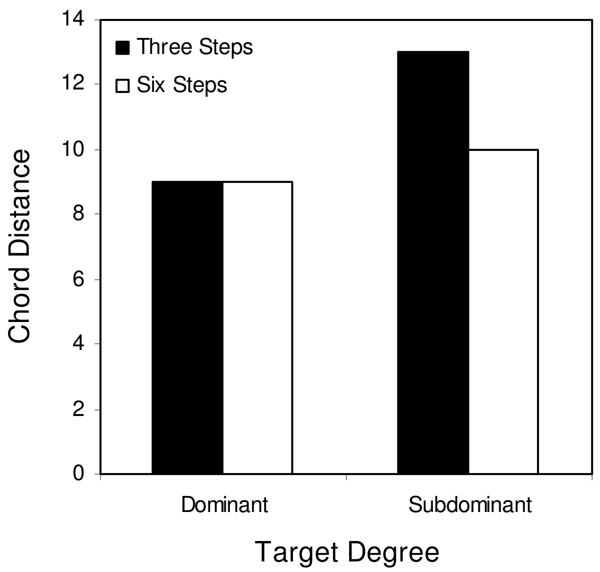


Figure 4.12 Predictions of Tonal Pitch Space (Lerdahl, 2001) for the conditions of Experiment 3.

4.3.1 METHOD

4.3.1.1 PARTICIPANTS

Thirty four participants participated in Experiment 3. They were students of Middle East Technical University. They had 1.08 years of musical training on average with a range of 0 to 12 years. Average age of the participants was 21.23 and the range was from 18 to 28.

4.3.1.2 STIMULI

Stimuli of Experiment 2 were scales and chord sequences (global context) followed by penultimate chord, and followed target chord. Major scales and major and minor chords composed of Shepard tones (1964) were utilized as global context. A program, which was based on the description of Shepard chords in Krumhansl, Bharucha and Kessler (1982), was implemented to generate them.

The target was the dominant or the subdominant of the tonality of the sequence. The penultimate chord was either three-steps or six-steps away from the target on the circle of fifths (clockwise for dominant and counterclockwise for subdominant targets). Penultimate chords followed scales or chord sequences and targets followed the penultimate chords immediately. Penultimate chords were played for 1000 msec and target chords (the last chord) were played for 2000 msec.

Three different chord sequences were used as global context. Each one consisted of eight chords. Each chord was 1000 msec long. The total duration of each sequence was 8000 msec. Chord sequences were a) "supertonic, subdominant, tonic, subdominant, tonic, supertonic, dominant, tonic" b) "submediant, dominant, dominant, tonic, tonic, supertonic, dominant, tonic" c) "tonic, subdominant, dominant, subdominant, tonic, dominant, dominant, tonic".

Three different scales were used as global context. They consisted of 15 pitches. The scales were a) ascending-descending b) descending-ascending, c) ascending-ascending. An ascending scale consisted of "tonic, supertonic, mediant, subdominant, dominant, submediant and leading" pitches. A descending scale consisted of the same pitches but in the reverse order. The last pitch of scales were always tonic. All pitches of scales were 500 msec long, except the last pitch (the tonic), which was 1000 msec. The total length of scales was 8000 msec.

Stimuli were repeated in four tonalities, namely C Major, A Major, F#/Gb Major and D#/Eb Major. Six types of global context (three for scales, three for chords) factored by two types of penultimate (three-steps vs. six-steps), factored by two types of target degree (dominant vs. subdominant) factored by two types of target timbre (Shepard vs. saw-tooth) factored by four tonalities made 192 stimuli in total.

Reaction times and accuracy were recorded in a timbre discrimination task (timbre-A/timbre-B) for the target chords. Timbre A was Shepard timbre (Shepard, 1964). Timbre B was saw tooth timbre. Saw-tooth waves were constructed by adding the first, second, fourth, and eighth harmonics with amplitudes inversely related to the harmonic number. Saw-tooth waves were generated by using the same phone to decibel and decibel to amplitude conversion tables that were used for generating Shepard chords (see Krumhansl, Bharucha and Kessler, 1982). Therefore, intensity levels of Shepard and saw-tooth targets were almost same.

4.3.1.3 PROCEDURE

Participants sat facing a computer screen. Test procedure and task demands (Shepard vs. saw-tooth timbre distinction) were explained to participants. The experiment started with a test session in which participants listened to single chords and decided whether their timbre was Shepard (Timbre A) or saw-tooth (Timbre B). Responses were given by pressing one of two designated response keys on a response box. Feedback on the accuracy of each response was given on the computer screen. Participants listened to 16 Shepard and 16 saw-tooth chords mixed randomly.

The experimental session followed the test session. In the experimental session participants listened to scales and chord sequences followed by the penultimate and target chords; and they were instructed to judge the timbre of the final chord. A set of question marks appeared on the computer screen together with the final chord, which required a response, to make it more distinct. Responses were given on the same keys used during the test part. Reaction times and accuracy were recorded in the timbre discrimination task (Shepard vs. saw-tooth) for the target chords. There were 196 stimuli in this session. Stimuli with scales and sequences as the global context were presented in different blocks. Order of blocks was counterbalanced across subjects. Stimulus presentation and response collection were controlled by Cedrus SuperLab Pro software. Stimuli were presented through Sennheiser HD-433 headphones.

4.3.2 RESULTS AND DISCUSSION

The average reaction time (RT) of correct responses and the percentage of error (PE) were calculated for each participant. Group averages of participants for different block orders are presented in Figure 4.13 and 4.14. RT and PE data were analyzed with Analysis of Variance (ANOVA) with global context (scale vs. sequence), target degree (dominant vs. subdominant), penultimate distance (three vs. six steps), target timbre (Shepard vs. saw-tooth) as within subject factors. RTs revealed a significant main effect of target timbre [$\underline{F}(1, 33) = 61.236$, $\underline{MS}_e = 10061.506$, $\underline{p} < .001$, $\eta^2 = .65$]. Saw-tooth targets were responded to faster (600 msec.) compared to Shepard targets (667 msec.). Two-way interaction between global context and target timbre was barely significant [$\underline{F}(1, 33) = 3.914$, $\underline{MS}_e = 8180.308$, $\underline{p} < .056$, $\eta^2 = .10$]. None of other main effects or interactions were significant.

In order to understand effects of block order on the results RT and PE data were analyzed with separate Analyses of Variance (ANOVA) with block order (first scale vs. first sequence) as a between subject factor, and global context (scale vs. sequence), target degree (dominant vs. subdominant), penultimate distance (three vs. six steps), and target timbre (Shepard vs. saw-tooth) as within subject factors. RT data did not reveal any significant interaction between block order and any other independent variable. PE data revealed a significant interaction between block order and global context [$\underline{F}(1, 32) = 35.831$, $\underline{MS}_e = 0.066$, $\underline{p} < .001$, $\eta^2 = .52$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = %16] participants who received scale block first responded to targets that followed scales with more error (%32) compared to targets that followed sequences (%15). In addition there was a three-way interaction between block order, target degree and penultimate distance [$\underline{F}(1, 32) = 5.782$, $\underline{MS}_e = 0.010$, $\underline{p} < .05$, $\eta^2 = .15$].

In order to examine the three-way interaction, PEs of subjects with different block order were analyzed separately with ANOVA with global context (scale vs.

sequence), target degree (dominant vs. subdominant), penultimate distance (three vs. six steps), target timbre (Shepard vs. saw-tooth) as within subject factors. PE data of subjects who received the scale block first revealed significant main effect of global context [$\underline{F}(1, 16) = 23.156$, $\underline{MS}_e = 0.079$, $\underline{p} < .001$, $\eta^2 = .59$]. They responded to targets that followed scales with more error (%31) compared to targets followed to sequences (%15). PE data of subjects who received the sequence block first also revealed significant main effect of global context [$\underline{F}(1, 16) = 12.773$, $\underline{MS}_e = 0.054$, $\underline{p} < .01$, $\eta^2 = .44$]. They responded to targets that followed sequences with more error (%21) compared to targets followed to scales (%11). PEs of subjects with different block order did not reveal any other significant main effect or interaction.

It was not possible to observe any chord priming effect from Turkish non-musician subjects in Experiment 3, which requires timbre discrimination. PE of participants decreased with practice. The results obtained in Experiment 3 were not sufficient to account for the difference between the results of Experiments 1 and 2.

Experiments 1, 2 and 3 investigated effects of non-diatonic chords on perception of harmony by observing the perceptual association between the Neapolitan chord and the dominant. Experiments 4, 5 and 6 continued the investigation by observing the perceptual association between secondary dominant chords and their diatonic associates.

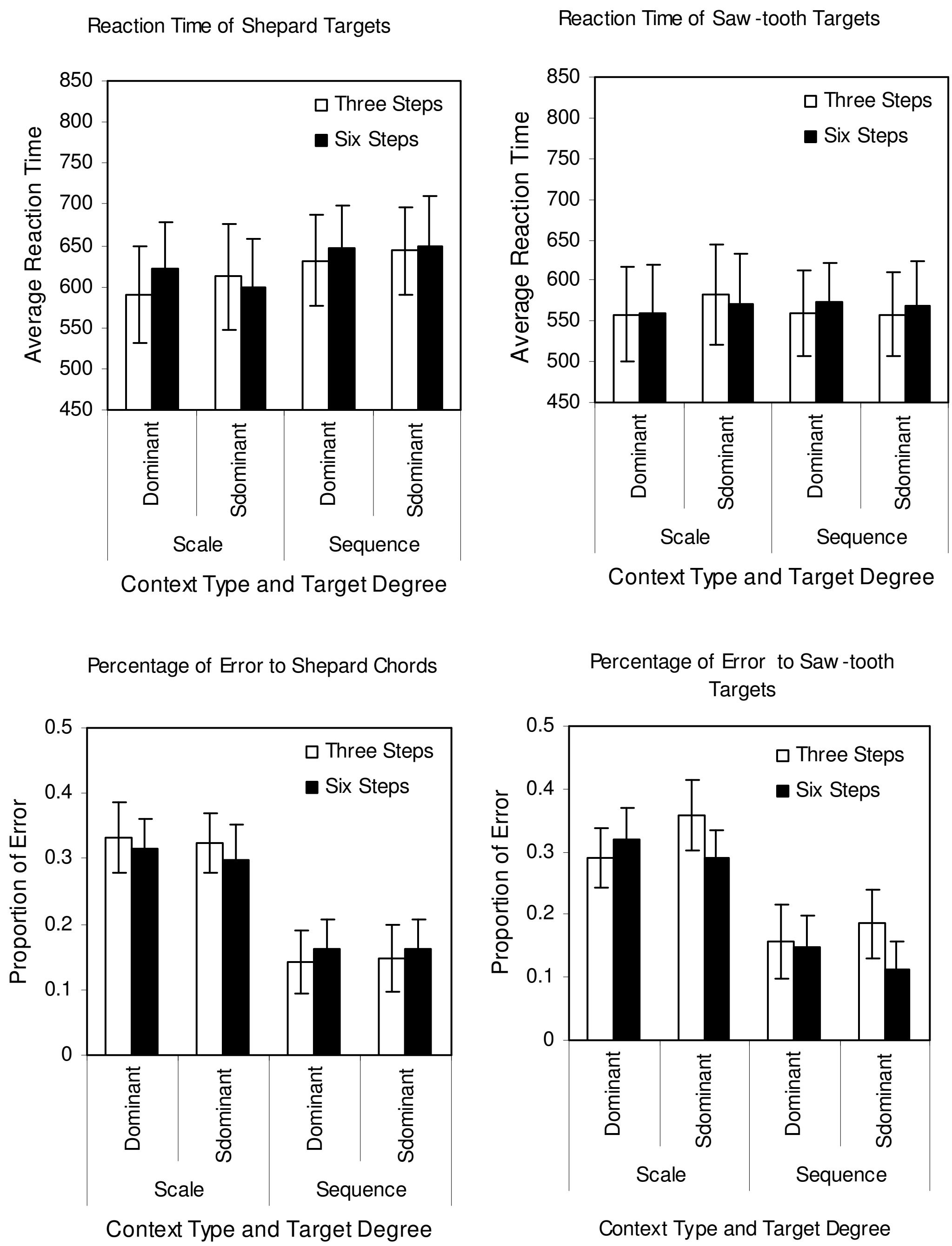


Figure 4.13 Average reaction time and percentage error of participants (non-musicians) who received scale block first for Shepard and saw-tooth targets for the conditions of Experiment 3.

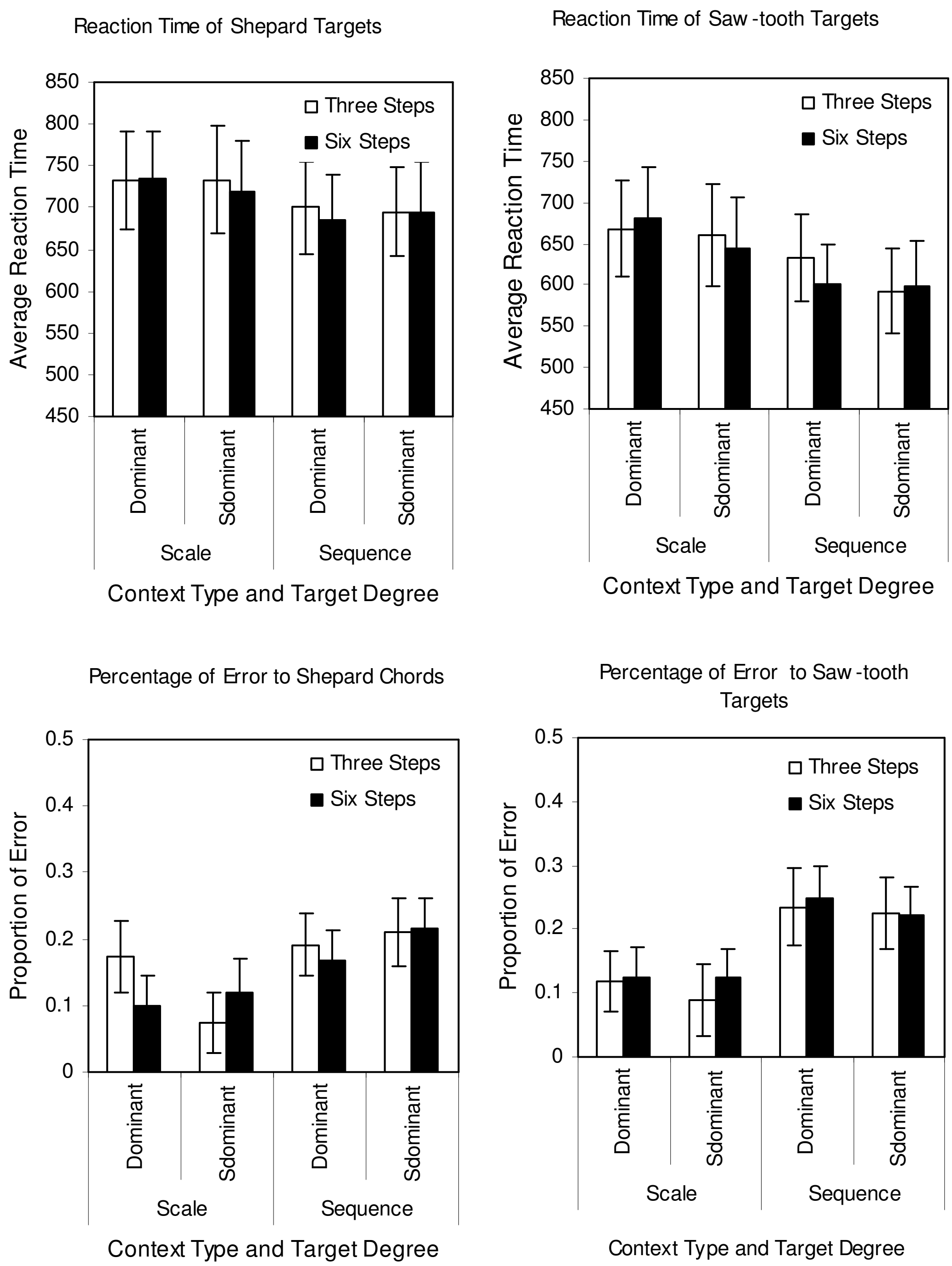


Figure 4.14 Average reaction time and percentage error of participants (non-musicians) who received sequence block first for Shepard and saw-tooth targets for the conditions of Experiment 3.

4.4 EXPERIMENT 4

Experiments 1, 2 and 3 investigated the association between the Neapolitan chord and the dominant. It was not possible to observe any priming effect in Experiment 3. Experiments 1 and 2 provided evidence supporting the hypothesis that stability was not the only determinant of expectation between chords, especially for musically educated participants. Experiments 4, 5 and 6 tested the role of stability in the perception of the chord relations for musically uneducated participants by investigating perceptual effects of secondary dominant chords.

Music theoretically, every major and minor diatonic chord is associated with the secondary dominant chords (Piston, 1978). Secondary dominants are non-diatonic major and diminished chords; and they must precede their diatonic associate. (See Section 3.2 for a detailed description of secondary dominant chords.) Secondary dominants are unstable chords; diatonic associates of secondary dominants are stable chords. Therefore, music theoretically, an unstable chord must precede specific stable chord. The psychological reality the association between secondary dominant chords and their diatonic associates was investigated in Experiment 4.

Experiment 4 investigated perceptual effects of secondary dominants by observing expectations towards tonic and subdominant targets (by observing global chord priming). Excerpts from Bach chorales (Greentree, 2005) were used as priming contexts, half of which employed secondary dominant chords in their original form. These chord sequences were also scrambled with a 2by2 algorithm. The 2by2 algorithm maps the sequence "1,2,3,4,5,6,7,8" to "2,1,4,3,6,5,7,8". Scrambled sequences violated the sequential relation between secondary dominants and their associates: secondary dominant chords succeeded their associates. In Experiment 4, effects of the original and scrambled sequences on global priming of chords were investigated. If participants perceived the sequential relation between diatonic chords and their secondary dominants, scrambling sequences that include secondary dominant chords should reduce global priming (facilitative responses to the tonic targets compared to the subdominant targets).

In Experiment 4 stimuli were composed of Shepard tones. Experiment 4 was the first study that investigated global chord priming by utilizing Shepard tones. This is critical given that predictions of MUSACT (Bharucha, 1987) and Tonal Pitch Space (Lerdahl, 2001) are different with respect to observing global chord priming with Shepard chords (see below).

Predictions of MUSACT (Bharucha, 1987) for the stimuli of Experiment 4 are shown in Figure 4.15. Simulation of the network and interpretation of the results were done in the same way as in Experiment 1. Relative activation was analyzed with Analysis of Variance (ANOVA) with target degree (tonic vs. subdominant), sequence type (diatonic vs. non-diatonic) and scrambling (original vs. scrambled) as within subject factors. Units of analysis were target unit activations before the target was presented as input and after the chord sequences utilized as the stimuli of the experiment (see Section 4.4.1.2 and Appendix D). Only significant main effects and interactions are reported. There was a main effect of target degree [$F(1, 9) = 36.68$, $MS_e = 0.0015$, $p < .001$], main effect of scrambling [$F(1, 9) = 10.0$, $MS_e = 0.0002$, $p < .05$], and main effect of sequence type [$F(1, 9) = 45.46$, $MS_e = 0.0006$, $p < .001$]. There was a three-way interaction between target degree, scrambling and sequence type [$F(1, 9) = 6.69$, $MS_e = 0.000006$, $p < .05$]. Diatonic and non-diatonic sequences were analyzed with separate ANOVA with target degree (tonic vs. subdominant) and scrambling (original vs. scrambled) as within subject factors. For diatonic sequences, there was a main effect of target degree [$F(1, 9) = 28.01$, $MS_e = 0.00087$, $p < .001$], and main effect of scrambling [$F(1, 9) = 12.13$, $MS_e = 0.00015$, $p < .01$]. For non-diatonic sequences there was a main effect of target degree [$F(1, 9) = 28.51$, $MS_e = 0.0011$, $p < .001$], and two-way interaction between target degree and scrambling [$F(1, 9) = 7.57$, $MS_e = 0.0000055$, $p < .05$]. According to MUSACT (Bharucha, 1987) tonic targets should be expected more strongly compared to subdominant targets after both diatonic and non-diatonic sequences. For diatonic sequences, scrambling should increase expectation for both tonic and subdominant targets equally. For non-diatonic sequences the difference between expectation towards tonic and subdominant targets should be increased when sequences are scrambled.

Predictions of Tonal Pitch Space (Lerdahl, 2001) were obtained by the chord distance rule (Lerdahl, 2001 p. 60). The chord distance rule is suitable to calculate expectation between chords without definite melody. Distance between chords does not change with respect to musical context that chords reside in. Chord distance between tonic and subdominant is $\Delta(\mathrm{I}/\mathbf{I} \rightarrow \mathrm{IV}/\mathbf{I})=7$, and between dominant and tonic is $\Delta(\mathrm{V}/\mathbf{I} \rightarrow \mathrm{I}/\mathbf{I})=7$. These distances are equal and do not change between conditions of the experiment. According to Tonal Pitch Space (Lerdahl, 2001) expectations for tonic and subdominant targets should be same when Shepard tones are utilized.

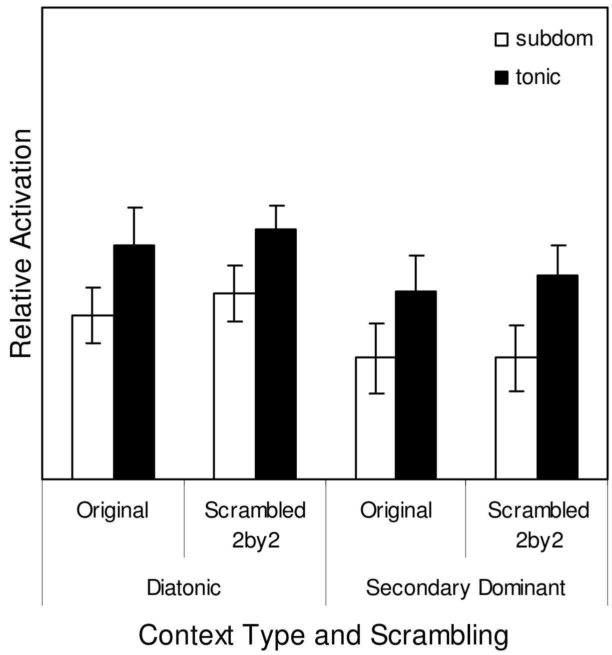


Figure 4.15 Predictions of MUSACT (Bharucha, 1987) for conditions of Experiment 4.

4.4.1 METHOD

4.4.1.1 PARTICIPANTS

Twenty eight participants participated in Experiment 4. They were students of Middle East Technical University. They had 1.53 years of musical training on average with a range of 0 to 16 years. Average age of participants was 21.71 and the range was from 17 to 25.

4.4.1.2 STIMULI

Stimuli of Experiment 4 were sequences of four part major, minor and diminished chords. Sequences were based on excerpts from Bach chorales. Excerpts were reduced to the quarter note Time Span Reduction level (Lerdahl & Jackendoff, 1983). Reduction was done by some additional rules to obtain acceptable harmonic progressions (Piston, 1978) at the quarter note level. As a result, sequences of chords were free from non-chordal pitches (such as appoggiatura, anticipation etc.); they followed the rules of harmonic progressions (Piston, 1978); they were composed of the root, the third and the fifth degrees in most cases; the base was replicated; and the base pitch was realized in the bass part. Some excerpts were transposed and the last chords of some other excerpts were changed to meet experimental needs. List of chorales utilized in the experiment are given in Appendix C. List of chord sequences used as the stimuli of the experiment are given in Appendix D.

Chord sequences were played with Shepard tones (Shepard, 1964). Shepard stimuli were generated exactly in the same way as in Experiment 2. Each chord was played for 1000 miliseconds (msec), except the last chord (the target chord) which was played for 2000 msec.

Sequences consisted of seven chords; and they were of two types: diatonic and non-diatonic. Diatonic sequences included diatonic chords only. Nondiatonic sequences included both diatonic and nondiatonic chords. Nondiatonic chords were secondary dominants. (See Section 2.1 for a description of secondary dominant chords.)

Twenty diatonic and 20 nondiatonic sequences were utilized. Half of the sequences (diatonic or nondiatonic) ended with the tonic, and the half ended with the subdominant. Sequences were scrambled according to the 2by2 algorithm (Tillmann & Bigand, 2001). The 2by2 algorithm maps the sequence "1,2,3,4,5,6,7,8" to "2,1,4,3,6,5,7,8".

The target was the tonic or the subdominant. If the target was the tonic, the penultimate chord was the dominant. If the target was the subdominant, the last two chords before the target were the dominant and the tonic. Reaction times and accuracy were recorded in a consonance/dissonance judgment task for the target chords. Consonant targets were major chords composed of seven pitches, obtained by repeating pitches of the major chord within the range of G1 (about 196 Hz) to F4 (1397 HZ). Dissonant targets were obtained by the addition of pitches a minor second and a tritone above the base pitch. Therefore, dissonant chords were composed of nine pitches. Sequences ending with the tonic were composed in A major, C major, D major, F major and G major tonalities; and sequences ending with the subdominant were composed in Bb major, C major, D major, F major, and G major tonalities. There were 160 sequences in total.

Forty new chord sequences were employed in the practice session (see Section 4.4.1.1). These sequences consisted of seven diatonic major and minor chords, and ended on the tonic or the subdominant target. These sequences were different from the experimental stimuli. Targets were consonant or dissonant chords, which were composed the same way as described above.

A new set of sequences were composed in order to diminish the transfer of the sense of tonality between experimental stimuli. These sequences consisted of twenty randomly selected chords. Chords were played with Shepard (1964) tones. Chords were of 100 msec duration. Base pitch of each chord was selected randomly from 12 pitch classes, and the chord type was selected randomly from "major, minor, or diminished".

4.4.1.3 PROCEDURE

Participants sat facing a computer screen. Test procedure and task demands (consonant vs. dissonant distinction) were explained to participants. Before the experiment they listened to three consonant and three dissonant chords as examples. In the experimental session they listened to chord sequences and they were instructed to judge the consonance of the final chord. A set of question marks appeared on the computer screen together with the final chord, which required a response, to make it more distinct. Responses were given by pressing one of two designated response keys on a response box. Feedback on the accuracy of each response was given on the computer screen.

A practice session preceded the experimental session. In the practice session participants listened to 40 sequences. Data from the practice session were not included in the analysis. Immediately after the practice, 180 experimental stimuli were presented in random order. Participants listened to random Shepard chord sequences (see Section 4.4.1.2) between stimuli in order to diminish the transfer of the sense of tonality. Stimulus presentation and response collection were controlled by a personal computer running Cedrus SuperLab Pro software. Participants heard the stimuli through Sennheiser HD-433 headphones.

4.4.2 RESULTS AND DISCUSSION

The average reaction time (RT) of correct responses and the percentage of error (PE) were calculated for each participant. Group averages for the conditions of the experiment are presented in Figure 4.16. RT and PE data were analyzed with Analyses of Variance (ANOVA) with target degree (tonic vs. subdominant), sequence type (diatonic vs. non-diatonic), scrambling (original vs. scrambled), and target consonance (consonant vs. dissonant) as within subject factors. Only significant main effects and interactions are reported.

RTs revealed a significant main effect of target degree [$\underline{F}(1, 27) = 4.952$, $\underline{MS}_e = 3178.053$, $\underline{p} < .05$, $\eta^2 = .16$]. Tonic targets were responded to faster (537 msec.) compared to subdominant targets (549 msec.) on average. There was a main effect of target consonance [$\underline{F}(1, 27) = 22.372$, $\underline{MS}_e = 9412.155$, $\underline{p} < .001$, $\eta^2 = .45$]. Dissonant targets were responded to faster (521 msec.) compared to consonant targets (565 msec.). There was a two-way interaction between target degree and sequence type [$\underline{F}(1, 27) = 4.599$, $\underline{MS}_e = 1216.074$, $\underline{p} < .05$, $\eta^2 = .15$]. Tonic targets were responded faster (536 msec.) compared to subdominant targets (555 msec.) after diatonic sequences on average. But this facilitation was decreased after non-diatonic sequences. Average response time after non-diatonic sequences for tonic targets was 539 msec., and for subdominant targets was 543 msec. Posthoc comparison did not revealed [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 25 ms] any significant differences. There was a two-way interaction between target degree and scrambling [$\underline{F}(1, 27) = 10.054$, $\underline{MS}_e = 728.558$, $\underline{p} < .001$, $\eta^2 = .27$]. Tonic targets were responded faster (539 msec.) compared to subdominant targets (543 msec.) after original sequences on average. But this difference was increased after scrambled sequences. Average response time after scrambled sequences for tonic targets was 536 msec., and for subdominant targets was 555 msec. Posthoc comparisons revealed a significant difference between tonic and subdominant targets after scrambled sequences. [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 19 ms]. There was a three-way interaction between sequence type, scrambling and target consonance [$\underline{F}(1, 27) = 10.054$, $\underline{MS}_e = 728.558$, $\underline{p} < .05$, $\eta^2 = .27$].

In order to examine the three-way interaction, RTs of consonant and dissonant targets were analyzed in separate ANOVAs with target degree (tonic vs.

subdominant), sequence type (diatonic vs. non-diatonic) and scrambling (original vs. scrambled) as between subject factors. RTs of consonant targets revealed a significant interaction between target degree and sequence type [$\underline{F}(1, 27) = 5.403$, $\underline{MS}_e = 1058.922$, $\underline{p} < .05$, $\eta^2 = .17$]. Consonant tonic targets were responded faster (556 msec.) compared to subdominant targets (577 msec.) after diatonic sequences on average. But this facilitation was eliminated after non-diatonic sequences. Average response time after non-diatonic sequences for tonic and subdominant targets was 563 msec. Posthoc comparison did not reveal any significant difference [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 23 ms]. RTs of dissonant targets revealed significant main effect of target degree [$\underline{F}(1, 27) = 5.414$, $\underline{MS}_e = 1864.880$, $\underline{p} < .05$, $\eta^2 = .17$]. Dissonant tonic targets were responded faster (514 msec.) compared to dissonant subdominant targets (528 msec.). There was a main effect of scrambling [$\underline{F}(1, 27) = 7.635$, $\underline{MS}_e = 767.554$, $\underline{p} < .01$, $\eta^2 = .22$]. Dissonant targets following original sequences were responded to faster (516 msec.) compared to dissonant targets followed scrambled sequence (526 msec.). PE data of Experiment 4 did not reveal significant main effects or interactions.

In Experiment 4 global priming of chords, i.e. facilitative responses to the tonic targets compared to the subdominant targets, was observed from Turkish subjects with Shepard tones (Shepard, 1964). Predictions of Tonal Pitch Space (Lerdahl, 2001) were disconfirmed, and predictions of MUSACT (Bharucha, 1987) were confirmed regarding the observation of global chord priming with Shepard tones. In addition, facilitative responses towards tonic targets compared to subdominant targets were reduced, if not eliminated, after non-diatonic sequences. Sequences that included non-diatonic chords had detrimental effects on global chord priming. On the other hand, the critical evidence for the perception of sequential relation between diatonic and secondary dominant chords (interaction between target degree, scrambling and sequence type), which was the main object of this experiment, was not observed. Participants did not perceive the sequential relation between diatonic chord and its secondary dominant.

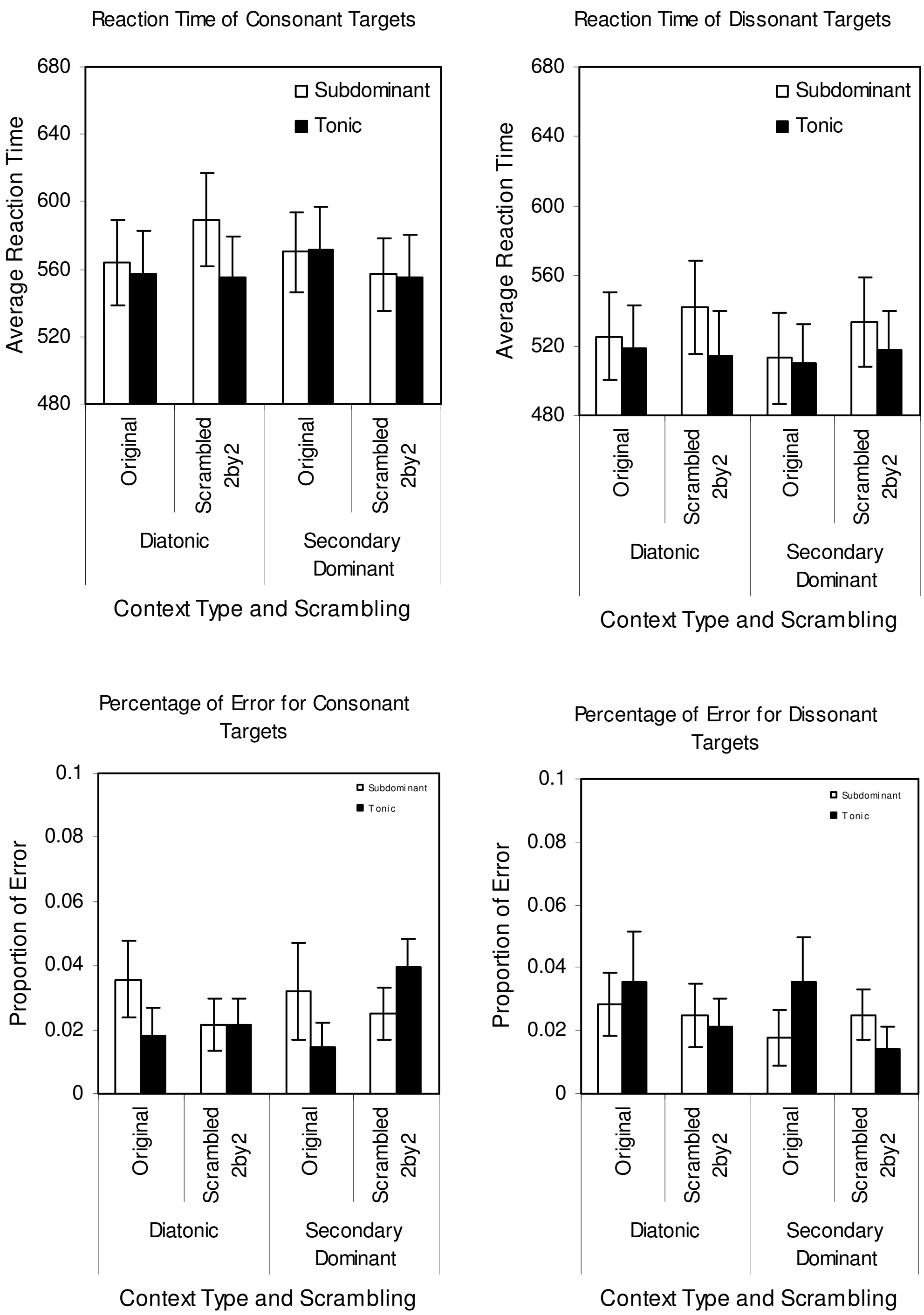


Figure 4.16 Average reaction time and percentage error of participants (non-musicians) for consonant and dissonant targets for the conditions of Experiment 4.

In Experiment 4 global chord priming was observed with diatonic sequences that were played with Shepard tones (Shepard, 1964) in Turkish non-musicians. In order words, Turkish non-musicians were able to perceive the perceptual difference between subdominant and tonic targets even if the melodic factors were diminished On the other hand, global chord priming was eliminated after non-diatonic sequences in Turkish non-musicians. The inclusion of non-diatonic chords in sequences disturbed the differential perceptual status of tonic and subdominant targets, when sequences were played with Shepard tones. Pitch height is an important factor in tonal harmony. It may be the case that Turkish non-musicians had difficulties in the integration of non-diatonic chords when melodic structures were not present in chord sequences. This possibility was investigated in Experiment 5.

4.5 EXPERIMENT 5

Experiment 4 investigated the perception of secondary dominant chords by utilizing Shepard chords and by observing expectations towards tonic and subdominant targets (by observing global chord priming). In Experiment 4 global chord priming was observed with Shepard chords from Turkish participants. Global chord priming was manifested with diatonic sequences but it was reduced, if not eliminated, with sequences that include non-diatonic chords.

By creating an illusion, Shepard tones (1964) reduced effects of melody on the results of Experiment 4. Though Shepard tones were used in many studies of music perception (Bharucha & Stoeckig, 1986, 1987; Krumhansl, 1990; Tekman & Bharucha, 1992, 1998), stimuli composed of Shepard tones lack important elements of music. Shepard tones create an auditory illusion of the circularity of pitch by emphasizing pitch class at the expense of pitch height. However, pitch height is an important factor in harmony. Pitch height imposes constraints on possible melodic movements (Piston, 1978). Experiment 5 investigated the perception of secondary dominant chords by utilizing piano timbre in order to obtain more natural musical stimuli.

Predictions of MUSACT (Bharucha, 1987) were same as in Experiment 4 (see section 4.4 and Figure 4.10). Predictions of Tonal Pitch Space Theory (Lerdahl, 2001) were obtained by the harmonic attraction rule (Lerdahl, 2001 p. 175). Attraction calculation was based on four-part chorales. Harmonic attraction from one chord to another does not change with the context that chords reside in. Harmonic attraction from tonic to subdominant is $\alpha_{rh}(I/\mathbf{I} \rightarrow IV/\mathbf{I})=1.66$, and from dominant to tonic is $\alpha_{rh}(V/\mathbf{I} \rightarrow I/\mathbf{I})=4.85$. According to Tonal Pitch Space (Lerdahl, 2001) expectations for the tonic targets should be stronger compared to the subdominant targets when piano timbre is used; and this should not change between conditions of the experiment.

4.5.1 METHOD

4.5.1.1 PARTICIPANTS

Twenty eight participants participated in Experiment 5. They were students of Middle East Technical University. They had 1.51 years of musical training on average with a range of 0 to 10 years. Average age of participants was 21.71 and the range was from 18 to 29.

4.5.1.2 STIMULI

Stimuli of the experiment were exactly same as in Experiment 4, except the case that sequences were played with piano chords. Stimuli were generated with MIDI piano timbre; and MIDI was converted to WAVE format with Midi2Wav software (http://www.midi2wav.com).

4.5.1.3 PROCEDURE

Procedure of the experiment was exactly same as in Experiment 4.

4.5.2 RESULTS AND DISCUSSION

The average reaction time (RT) of correct responses and the percentage of error (PE) were calculated for each participant. Group averages for different conditions of the experiment are presented in Figure 4.17. RT and PE data were analyzed in Analyses of Variance (ANOVA) with target degree (tonic vs. subdominant), sequence type (diatonic vs. non-diatonic), scrambling (original vs. scrambled), and target consonance (consonant vs. dissonant) as within subject factors. Only significant main effects and interactions are reported.

RT data of Experiment 5 did not reveal significant main effect or interaction. However, the main effect of target degree [$\underline{F}(1, 27) = 3.564$, $\underline{MS}_e = 6960.848$, $\underline{p} < .07$, $\eta^2 = .11$] and the two-way interaction between scrambling and sequence type [$\underline{F}(1, 27) = 3.564$, $\underline{MS}_e = 6960.848$, $\underline{p} < .07$, $\eta^2 = .11$] were close to significance.

PE data of Experiment 5 revealed a significant main effect of target degree [$\underline{F}(1, 27) = 5.095$, $\underline{MS}_e = 0.005$, $\underline{p} < .05$, $\eta^2 = .16$]. Subdominant targets were responded with more errors (%12) compared to the tonic targets (%11) on average. There was a three-way interaction between target degree, sequence type and scrambling [$\underline{F}(1, 27) = 6.170$, $\underline{MS}_e = 0.009$, $\underline{p} < .05$, $\eta^2 = .19$]. After original diatonic sequences, the subdominant targets were responded with more errors (%12) compared to the tonic targets (%8) on average. After scrambled diatonic sequences, the subdominant (%11) and the tonic (%11) targets and tonic targets were with equal percentage of errors on average. After original non-diatonic sequences, the tonic targets were responded with more errors (%12) compared to the subdominant targets (%10) on average. After scrambled diatonic sequences, on the other hand, the subdominant targets were responded with more errors (%14) compared to the tonic targets (%10) on average. In order to examine the three-way interaction, PE of targets followed diatonic and non-diatonic sequences were analyzed with separate ANOVAs with target degree (tonic vs. subdominant), scrambling (original vs. scrambled) and target consonance (consonant vs. dissonant) as between subject factors. These analyses did not reveal significant main effects or interactions.

In Experiment 5 RT data did not reveal a global priming of chords from Turkish subjects with piano tones, however PE data suggested it. According to Tonal Pitch Space Theory (Lerdahl, 2001) expectations for the tonic targets should be stronger compared to the subdominant targets when piano timbre is used; and this should not change between conditions of the experiment. However, significant interaction between target degree, scrambling and sequence type was observed with PE data. Therefore, predictions of Tonal Pitch Space (Lerdahl, 2001) were disconfirmed. Predictions of MUSACT (Bharucha, 1987) regarding the observation of global chord priming was confirmed only with PE data. For PE data a three-way interaction between target degree, scrambling and sequence type was observed, but the interaction was not in the way that MUSACT suggested. Contrary to the results of Experiment 4, detrimental effects of non-diatonic sequences on global chord priming were not observed with piano tones. Lastly, participants did not perceive the sequential relation between diatonic chord and its secondary dominant even if sequences were presented with piano tones, which was the main topic of the investigation.

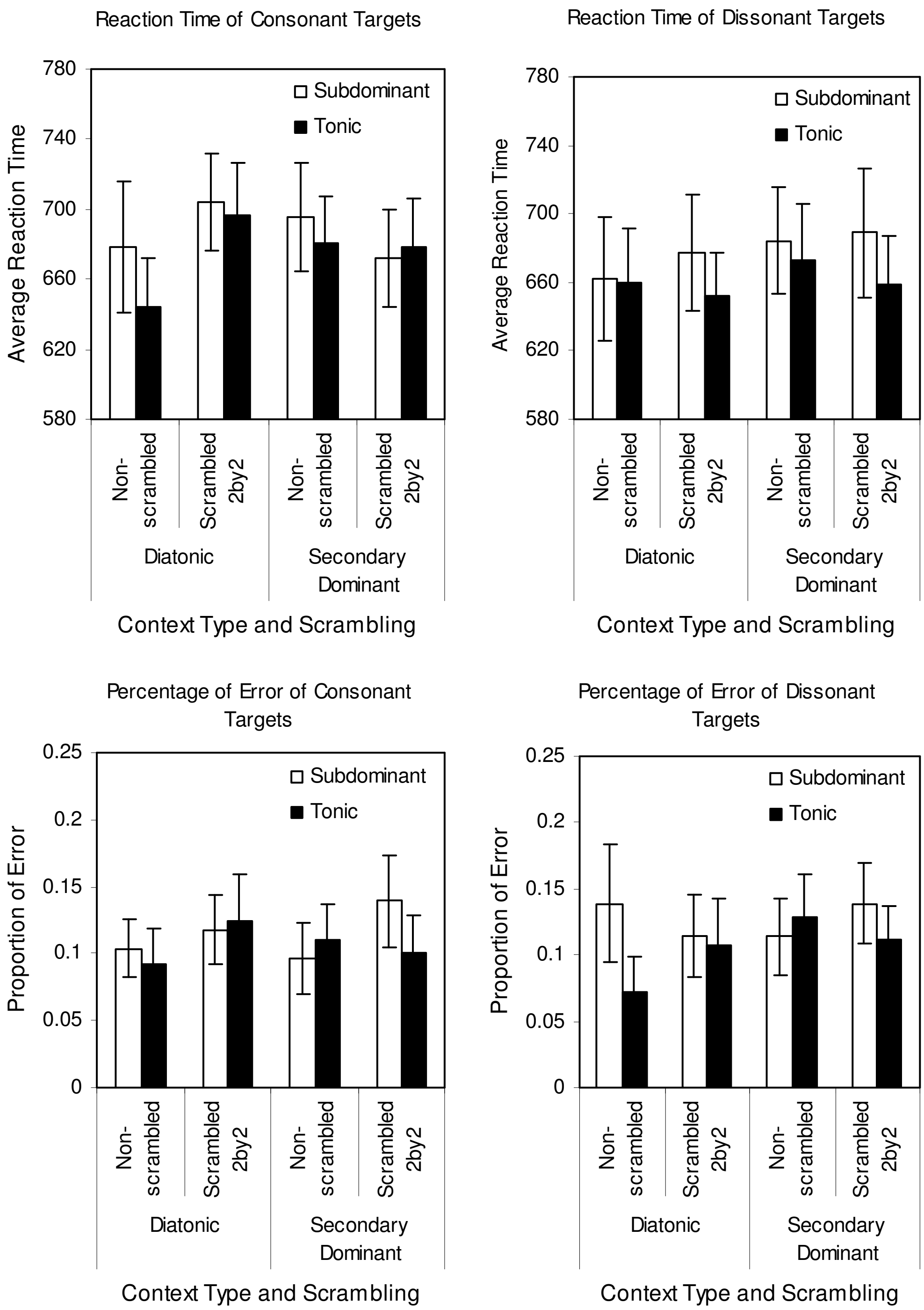


Figure 4.17 Average reaction time and percentage error of participants for consonant and dissonant targets for the conditions of Experiment 5.

In Experiment 5, it was not possible to observe global priming of chords in Turkish non-musicians. In Experiment 4, stimuli were player with Sheapard chords; in Experiment 5 it was played with piano tones. It was not possible to observe detrimental effects of non-diatonic chords on priming, contrary to the results of Experiment 4. However, this may have been due to the absence of global chord priming in Experiment 5. In Experiment 4 and 5, participants did not perceive the relation between the secondary dominant chords and their diatonic associates, which used 2by2 scrambling algorithm. With 2by2 scrambling algorithm, the sequential relation between the secondary dominants and their diatonic associates were disrupted; but these chords were still adjacent in chord sequences. It may be the case that the secondary dominants and their diatonic associates should be placed far apart in sequences to observe the disruption of global chord priming. This hypothesis was examined in Experiment 6.

4.6 EXPERIMENT 6

Experiment 4 and 5 investigated the effect of secondary dominant chords on perception of harmony by utilizing Shepard and piano tones. Global chord priming was observed definitely with Turkish participants with Shepard chords in Experiment 4 and suggestively with piano tones in Experiment 5. Global chord priming was eliminated after sequences that include non-diatonic chords in Experiment 4; but such a result was not observed in Experiment 5. Results of both experiment disconfirmed the hypothesis that sequential relation between diatonic and secondary dominant chords was perceived. A detrimental effect of scrambled non-diatonic sequences on global chord priming was not observed. However, Experiment 4 and 5 utilized 2by2 scrambling algorithm, which maps sequence "1,2,3,4,5,6,7,8" to "2,1,4,3,6,5,7,8". It may be the case that sequences were not scrambled enough to induce such results. In order to test this hypothesis, in Experiment 6 sequences were scrambled with 4by4 algorithm, which maps the sequence "1,2,3,4,5,6,7,8" to "4,1,5,2,6,3,7,8". With the 4by4 algorithm scrambling algorithm, the secondary dominants and their diatonic associates were placed far apart in chord sequences.

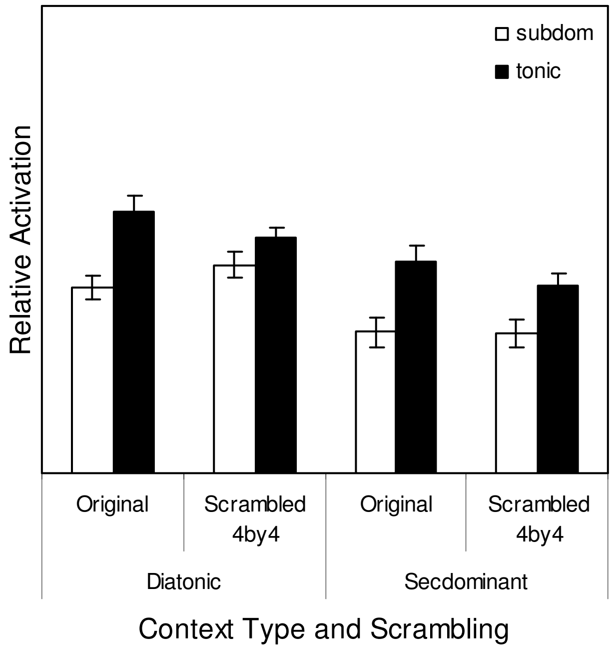


Figure 4.18 Predictions of MUSACT (Bharucha, 1987) for conditions of Experiment 6.

Predictions of MUSACT (Bharucha, 1987) for the stimuli of Experiment 6 are shown in Figure 4.18. Simulation of the network and interpretation of the results were done in the same way as in Experiment 1. Relative activation was analyzed with Analysis of Variance (ANOVA) with target degree (tonic vs. subdominant), sequence type (diatonic vs. non-diatonic) and scrambling (original vs. scrambled) as within subject factors. Units of analysis were target unit activations before the target was presented as input, and after the chord sequences utilized as the stimuli of the experiment (see Section 4.6.1.2 and Appendix D). Only significant main effects and interactions are reported. There was a main effect of target degree [$F(1, 9) = 27.694$, $MS_e = 0.0063$, $p < .001$], a main effect of sequence type [$F(1, 9) = 62.218$, $MS_e = 0.0026$, $p < .001$], and a two-way interaction between target degree, and scrambling

[$\underline{F}(1, 9) = 6.808$, $\underline{MS}_e = 0.0027$, $\underline{p} < .05$]. According to MUSACT (Bharucha, 1987) tonic targets should be expected more compared to subdominant targets after both diatonic and non-diatonic sequences. Targets following diatonic sequences should be expected more compared to targets followed non-diatonic sequences. In addition, scrambling should decrease the difference between expectation towards tonic and subdominant targets. Predictions of Tonal Pitch Space (Lerdahl, 2001) were same as in Experiment 5 (see Section 4.5).

4.6.1 METHOD

4.6.1.1 PARTICIPANTS

Thirteen participants participated in Experiment 6. They were students of Middle East Technical University. They had 0.61 years of musical training on average with a range of 0 to 4.5 years. Average age of participants was 24.41 and the range was from 18 to 30.

4.6.1.2 Stimuli

Stimuli of the experiment were exactly same as in Experiment 4, except the case that sequences were scrambled with 4by4 algorithm (Tillmann & Bigand, 2001). 4by4 algorithm maps the sequence "1,2,3,4,5,6,7,8" to "4,1,5,2,6,3,7,8".

4.6.1.3 Procedure

Procedure of the experiment was exactly same as in Experiment 4.

4.6.2 RESULTS AND DISCUSSION

The average reaction time (RT) of correct responses and the percentage of error (PE) were calculated for each participant. Group averages for different conditions of the experiment are presented in Figure 4.19. RT and PE data were analyzed with Analyses of Variance (ANOVA) with target degree (tonic vs. subdominant), sequence type (diatonic vs. non-diatonic), scrambling (original vs. scrambled), and target consonance (consonant vs. dissonant) as within subject factors. Only significant main effects and interactions are reported.

RT data of Experiment 6 revealed significant main effect of target degree [$\underline{F}(1, 12) = 4.910$, $\underline{MS}_e = 2245.357$, $\underline{p} < .05$, $\eta^2 = .29$]. Tonic targets were responded to faster (712 msec.) compared to subdominant targets (729 msec.). There was a two-way interaction between sequence type and target consonance [$\underline{F}(1, 12) = 23.568$, $\underline{MS}_e = 1141.112$, $\underline{p} < .001$, $\eta^2 = .66$]. Posthoc comparison revealed that [Tukey's Honestly Significant Difference (HSD) with p set at .05 = 39 ms] after diatonic sequences, dissonant targets were responded faster (680 msec.) compared to consonant targets (749 msec.).

In Experiment 6 global chord priming was observed from Turkish subjects with piano tones. Predictions of Tonal Pitch Space (Lerdahl, 2001) were supported by Experiment 6, contrary to the Experiment 5. According to Tonal Pitch Space there should be no difference between conditions of the experiment with respect to the observation of global chord priming with piano tones. Although significant interactions of Experiment 5 were not replicated in Experiment 6, low statistical power may have hidden such effects. Predictions of MUSACT (Bharucha, 1987) regarding the observation of global chord priming was confirmed. On the other hand, it was not possible to observe the two-way interaction between target degree and scrambling that MUSACT suggested. Again low statistical power may have hidden such effects. Most importantly, it was not possible to observe detrimental effects of scrambled non-diatonic sequences on the global chord priming, which was the main topic of the investigation. It may be that low statistical power may also have hidden such effects. But given the pattern of RT and PE data of Experiment 6, and results of Experiments 4 and 5, it is highly unlikely to observe such an interaction by collecting more data.

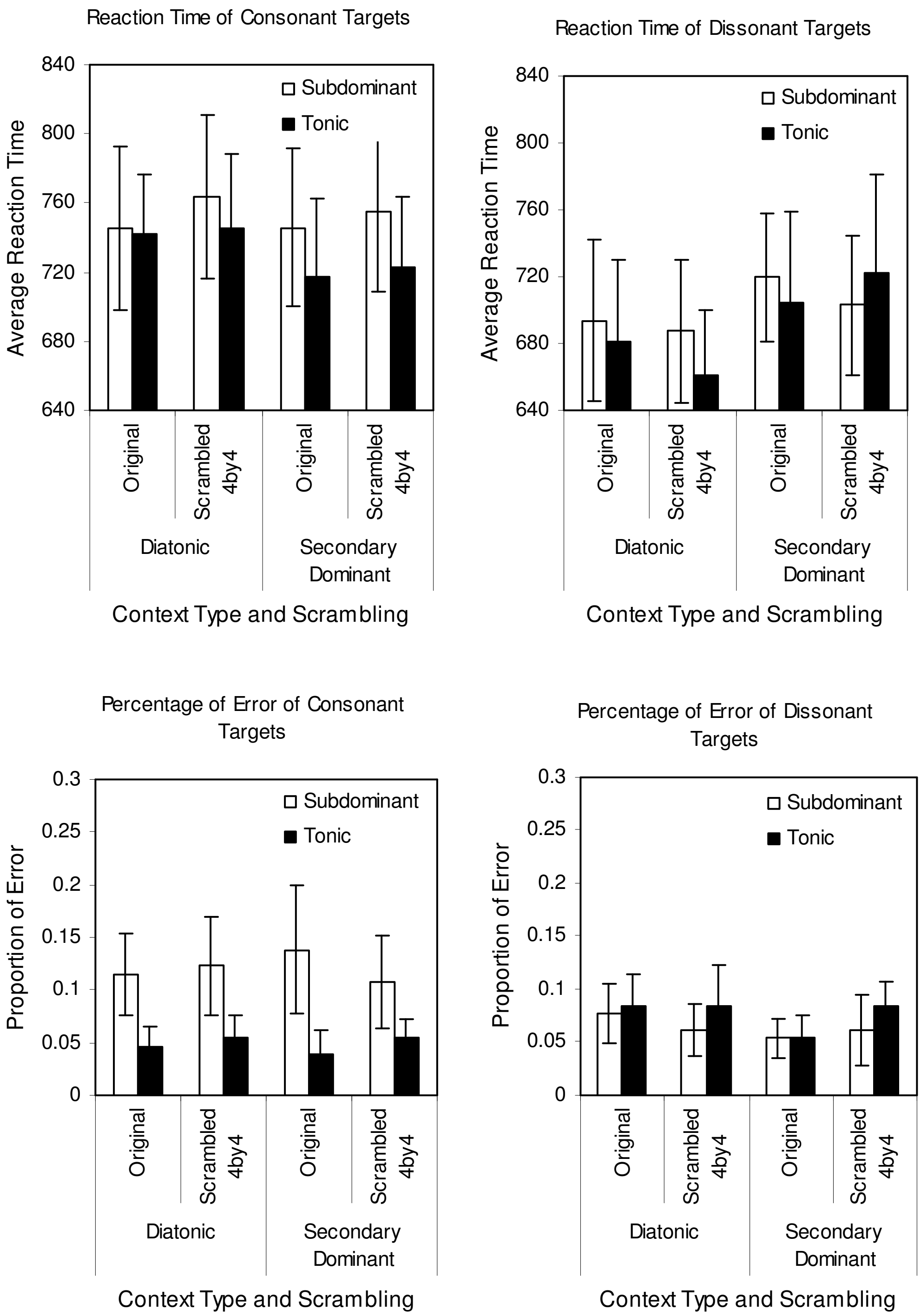


Figure 4.19 Average reaction time and percentage error of participants for consonant and dissonant targets for the conditions of Experiment 6.

CHAPTER 5

SELF-ORGANIZING MAP SIMULATIONS

In this chapter, the rationale behind studying self-organizing maps of chord schema is described, self organizing maps are introduced, and simulation results are presented.

The aim of artificial neural network simulations in this thesis was to develop self-organizing models of chord schema and to test their performance. This was not restricted by the results of the behavioral experiments of the thesis. The aim is to develop self-organizing chord schema models that were functionally equivalent to MUSACT (Bharucha, 1987). Previous studies with a similar aim (Gjerdingen, 1992; Leman, 1995; Leman and Carreras, 1997; Tillmann , Bharucha, Bigand, 2000) grounded the current attempt.

5.1 SELF-ORGANIZING MAPS

A self-organizing map is an artificial neural network. Artificial neural networks are computing devices inspired by the structure and function of the nervous system (McCulloch & Pitts, 1943). They are composed of units and links between units. Units implement simple functions, i.e. they map specific inputs to specific outputs. Inputs and outputs may be logical values or real numbers (Aleksander & Morton, 1995). Links carry output from a unit to another as an input. Links are weighted. Weight reflects connection strength between units. Artificial neural networks were both studied from the theory of computation (Minsky, 1967; Minsky & Papert, 1969) and the theory of statistics (Hertz, Krogh & Palmer, 1991) points of view. Artificial neural networks implement associative memories, functions of classification, feature extraction, pattern recognition, statistical analysis, generalization, and optimization etc. (Hertz, Krogh & Palmer, 1991).

A self-organizing map is a topologically correct feature map that develop with respect to input signals (Kohonen, 1982). It is a feature map because units of the network become sensitive to different features of the input signal after training. It is self-organizing because the training algorithm works in an unsupervised way; i.e. during training there is no teacher to depict features and corresponding signals. It is a map because network units have specific distance relation. It is a topologically correct feature map because nearby units on the map become sensitive to similar features. Self-organizing maps are designed in such a way that "... external signal activity alone ... is sufficient for enforcing mappings ... into the system." (Kohonen, 1982, p.59).

The training algorithm of a self-organizing map is quite simple. Training signals are presented to the map one by one in random order. For each signal the best matching unit, whose weight vector is closest to the signal vector, is detected (see Formula 5.1). Units that lie within the neighborhood of the best matching unit are subject to weight update. The weight update function rotates the weight vector towards the direction of the signal vector insignificantly (see Formula 5.2). The amount of rotation is a function of the learning rate (see Formula 5.3) and the distance between the best matching unit and the unit whose weight is updated (see Formula 5.4).

It is worth noting that there are two distance functions on a self-organizing map (distance on the map and distance on the feature space) whose domains are mutually exclusive. The distance function on the map operates on coordinates of units *on the map*. Units that lie within the neighborhood of the best matching unit are found with respect to the distance on the map. The distance function on the feature space operates on coordinates of weight and signal vectors *on the signal space*. The best matching unit and weight updates are calculated with reference to this function.

Four formulae (the best matching unit, weight update, learning rate and unit distance formulae) are sufficient to describe self-organizing maps. Formulae presented below are from Vesanto, Himberg, Alhoniemi, Parhankangas (2000). **i** is the input vector and $\mathbf{w_j}$ is the weight vector of j^{th} unit. $|a|$ refers to the length of vector **a**. Weight vector of best matching unit is marked with *.

Formula of best matching unit is

$$\left|w_{j*} - i\right| \leq \left|w_j - i\right| \qquad \forall j \tag{5.1}$$

Formula of weight update is

$$w_j(t+1) = w_j(t) + \alpha(t)h_{cj}(t)\left\langle i(t) - w_j(t)\right\rangle \tag{5.2}$$

where learning rate is

$$\alpha(t) = \alpha_0(0.005/\alpha_0)^{t/T} \tag{5.3}$$

and unit distance (neighborhood) is

$$h_{cj}(t) = e^{-d_{cj}^2/2\sigma_t^2} \tag{5.4}$$

The learning rate and the neighborhood gradually decrease. Training stops when the learning rate is zero. Kohonen (2000) suggested that better map organization is obtained by splitting the training into two phases: rough organization and fine tuning. In the rough organization phase, the neighborhood is large, the learning rate is high and training takes short. In the fine tuning phase, the neighborhood and the learning rate are small and training takes long.

Figure 5.1 demonstrates a topologically correct mapping obtained by a self-organizing map (from Kohonen, 2000). The self-organizing map was composed of 25X25 units, which were laid on a square surface. Feature space was two dimensional. Training data were randomly selected from points inside the square.

On the figure, intersection points represent coordinates of weight vectors. Coordinates of weight vectors were plotted for several epochs. It is clear from the figure that weight vectors were organized to approximate global features of the training signal; and neighbor units mapped the nearby coordinates.

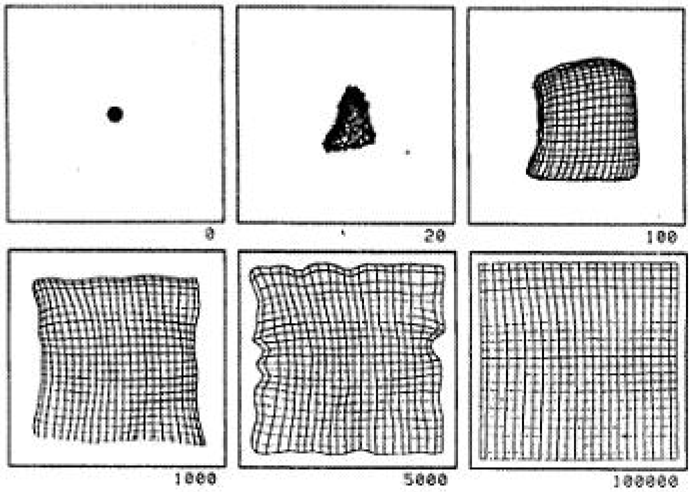


Figure 5.1 An example of development of topologically correct mapping, Kohonen (2000, fig. 3.6).

Self-organizing maps (Kohonen, 1982) produce topologically correct maps with neuron-like computations. Topological mappings were found in many places of the brain (Knudsen, du Lac, & Esterly, 1987 cited in Kohonen, 1982), which suggests that it has an important function for neural computation. Tonotopic mapping is an example of the topological mapping. In tonotopic maps, spatial arrangement of neurons reflects frequencies of sound waves. One of the tonotopic maps of the nervous system was located in right-temporal cortex (Fujioka, Ross, Okamoto, Hidehiko, Takeshima, Kakigi, & Pantev, 2003; Pantev, Hoke, Lehnertz, Lütkenhöner, Anogianakis, & Wittkowski, 1988; Zatorre, 1988; Pantev, Hoke, Lütkenhöner, Lehnertz, 1989).

5.2 SELF-ORGANIZING MODELS OF CHORD PERCEPTION

MUSACT (Bharucha, 1987) is a neural network model of tonal perception. (See Sections 1.3 for a detailed description of MUSACT.) It simulates a wide range of experimental results on perception of pitch, chord and tonality (Bharucha, 1987; Tillmann et.al. 2000). MUSACT (Bharucha, 1987) is composed of three groups of interconnected units; i.e. units of pitch categories (12 units), major and minor chords (12 units for each), and major tonalities (12 units). Every chord unit is connected to the units representing the three pitches contained in that chord. Every major tonality unit is connected to units representing the three major and three minor chords that are possible in that tonality. Units, connections and connection weights of MUSACT were hand-wired (Bharucha, 1987). However, in order to be a complete theory of tonal perception, connections and weights of MUSACT should be obtained by a self-organizing mechanism that adapts regularities in the musical environment (Tillmann et.al. 2000, p. 891).

Tillmann et.al. (2000) obtained the architecture of MUSACT by training two hierarchically ordered self-organizing maps. The first map was composed 6X6 (36) units located on a square surface. It was trained with chords to obtain representations of chord categories. Training data were composed of 24 12-dimensional vectors. Twenty four vectors represented 12 major and 12 minor chord categories. Chords were coded with either sparse or rich coding. In sparse coding, chords were represented by three pitches (pitches of the chord), whose weights were equal. Chords that share pitches were similar to each other in sparse coding. Sparse encoding represented the chord categories. In rich coding chords were represented by 12 pitches, whose weights were determined by the sum of subharmonics of pitches of chords (Parncutt, 1988). Rich encoding mirrored similarities between psycho-acoustical representations of chords (Leman, 1995; Terhadt, 1974; Parncutt, 1988), which approximated the circle-of fifths. At the end of training, 24 units of the first map represented 12 major and 12 minor chord categories. These units were located on the first map with respect to the similarities between input representations of different chord categories.

The second map was composed of 4X4 (16) units, which were fully connected to the units of the first map. The second map was trained with chord sequences to obtain representations of tonality categories, after the first map was fully trained. Simulations were carried with two chord corpora. In the first corpus there were 12 chord sequences each composed of six chords. Chords were three major and three minor chords that were possible in a major tonality. In the second corpus there were 120 chord sequences. Ten chord sequences were constructed randomly by referring to common chord progressions (Piston, 1978), statistical distributions of chords (Budge, 1943, cited in Tillmann et.al., 2000) and termination rules for chord sequences (Piston, 1978). These sequences were transposed to 12 major tonalities to guarantee identical distribution of chords in different tonalities. Pitches that composed chords of the sequences were presented to the first map one by one. Activation of the best matching unit on the first map was recorded and kept constant till the end of the sequence. The best matching unit was always one of the 24 chord categories. The second map was trained with activation patterns of the first map obtained at the end of sequences. After training, 12 units of the second map represented 12 major tonality categories. These units were located on the second map by forming a circle of fifths. At the end, hierarchically self-organized maps detecting chord and tonality categories were obtained.

These maps were far from being equal to MUSACT (Bharucha, 1987). MUSACT has bi-directional connections. Activation is repeatedly transmitted between layers in MUSACT; and after some time the transmission of activation reaches to a steady state (i.e. activation reverberates and reaches to the equilibrium state). Hierarchically self-organized maps of Tillmann et. al. (2000) had only bottom-up connections. In order to convert architecture of self-organized maps to the architecture of MUSACT (Bharucha, 1987), Tillmann et. al. (2000) pruned some links and scaled weights.

Pruning was necessary to rule out reverberation of activation between units that did not represent a category. At the pruning step, training stimuli were presented again to the network. For each stimulus, the unit with the highest activation was detected. Weights of links between input signals and best matching units were increased and remaining weights were decreased by some amount. At the end of pruning, links among the units were identical to MUSACT (Bharucha, 1987). Tillmann et. al. (2000) scaled weights of remaining links. Scaling of weights was necessary to ensure the termination of reverberation of activation after some time. The termination of reverberation depends on eigenvalues of connection matrices, which must be lower than 1 (footnote 7, Tillmann et. al. 2000). (See Searle (1967) for definition of eigenvalue.). In order to meet this constraint, weights of the network were scaled. (Tillmann et. al. (2000) did not give mathematical details of scaling.) The pruned and scaled neural networks, which were hierarchically self-organized according to the regularities of musical environment, were identical to MUSACT and simulated wide range of experimental results on perception of pitch, chord and tonality (Tillmann et. al. 2000). Among the simulation results, those about perception of chords are summarized here.

Tillmann et. al. (2000) gave the correlation between MUSACT (Bharucha, 1983) and self-organized chord schemata trained with different input representations and chord corpus. The lowest correlation with MUSACT (Bharucha, 1983) was obtained with the self-organized chord schema that was trained with the rich encoding and the big corpus ($r=0.97$, $p<0.01$, $N=24$). The highest correlation with MUSACT (Bharucha, 1983) was obtained with the self-organized chord schema that was trained with the sparse encoding and the small corpus ($r=0.99$, $p<0.01$, $N=24$). Tillmann et. al. (2000) showed that principles of chord perception, i.e. contextual identity, contextual distance, and contextual asymmetry (Krumhansl, 1990) were simulated by the models. In addition, their models simulated the facilitative responses given to the target chord when the prime closer to the target on the circle of fifths compared to when it was distant (the chord priming effect) (Bharucha &

Stoeckig, 1986); the fact that chord priming is due to the distance on the circle of fifths but not to the acoustical similarity between the prime and the target (Tekman & Bharucha, 1998); and the facilitative responses given to the target chord when it is tonic compared to when it is subdominant (the global priming effect) (Bigand & Pineau, 1997);

Although simulations were quite successful, steps taken on the way to obtain MUSACT (Bharucha, 1987) depended on many questionable assumptions. According to the self-organization scenario, first chord categories then tonality categories should be learned. This hypothesis was not supported by empirical findings in Tillmann et. al. (2000). Similarly, pruning and scaling suggested various stages in the time-course of learning chord and tonality categories. This is also a hypothesis waiting to be proven. In addition, it is very rare to find category exemplars of chords detached from musical context, as it was suggested by the training stimuli of the first map. Most importantly, activation and output functions implemented by units of MUSACT (Bharucha, 1987; Tillmann et. al. 2000, Formula 1) and units of self-organizing maps (Kohonen, 1982; Tillmann et. al. 2000, Formula 2) were different. Activation of a unit on MUSACT was the addition of previous activation with the activation received (Tillmann et. al. 2000, Formula 1). On the other hand, activation of a unit on a self-organization map depends only on the current input (Tillmann et. al. 2000, formula 1; Kohonen, 1982, formula 1). Output of a unit on MUSACT was the difference between signal that unit holds and receives (Bharucha, 1987). On the other hand, output of a unit on self-organizing map was a linear function of input (i.e. input and output was identical) (Tillmann et. al. 2000). Tillmann et. al. (2000) did not describe and rationalize these important changes. Lastly, Tillmann et. al. (2000) utilized chord sequences of major tonalities only. It was critical to represent categories of minor tonalities as well (Tillmann et. al., 2000, p.901).

Leman (1995) and Leman and Carreras (1997) trained self-organizing maps (Kohonen, 1982) to model chord and tonality perception. Their maps were composed of 20X20 (400) units and 100X100 (10000) units located on a torus[1]. Input signal was firing patterns of auditory nerves obtained by converting acoustical signals to auditory nerve patterns. These calculations were based on models of perception of pitches and chords (Parncutt, 1988; Terhadt, 1982; Terhardt, Stoll & Seewan, 1982a+b; Van Immerseel & Martens, 1992). In some simulations temporal relations between neural representations were converted into activation patterns with the help of a short-term memory model. Short-term memory model integrated activations of recent history into the current activation pattern (Leman, 1995, p.100). Leman (1995) utilized two training sets. The first training set consisted of 115 chord categories. (These were 12 major, 12 major, 12 diminished, 4 augmented, 12 major seventh, 12 minor seventh, 12 dominant seventh, 12 half-diminished seventh, 12 augmented-seventh, 12 minor with major seventh, 3 diminished seventh chords.) Acoustical representations of chord categories were converted to neural representations. The second training set consisted of 48 chord sequences. Two chord sequences were composed; and they were transposed to 12 tonalities to guarantee identical distribution of chords in different tonalities. Each sequence was divided into 72 acoustical windows. Acoustical representations were converted to neural patterns, upon which effects of short-term memory were added. Leman and Carreras (1997) utilized real performances of Well-Tempered Clavier[2] as training set. Training stimuli were divided into acoustical windows of three seconds. Acoustical representations were converted to neural representations, in which effects of short-term memory were added. Training with different sets of data gave similar results (see Leman, 1995; Leman & Carreras, 1997). After training, some units represented chord categories and some others represented tonality categories. Tonality category

---

[1] Torus is a surface having the shape of a doughnut (Pickett, 2003).

[2] Well-Tempered clavier is one of famous pieces of Johann Sebastian Bach. It is composed for solo keyboard.

units were located in a way reminding the circle of fifths. Location and activation and of chord category units reflected the organization of the circle of fifths. On the other hand, Leman (1995) and Leman and Carreras (1997) did not present simulations of empirical data on perception of chords with their models.

There were shortcomings of Leman (1995) and Leman and Carreras (1997). First, training samples were based on time intervals measured in seconds. This was not an efficient metric for finding chord boundaries. Duration of chords is measured with musical meter but not with seconds (Lerdahl, 2001; Lerdhal & Jackendoff, 1983). Musical time is measured with beats (Lerdahl & Jackendoff, 1983). Rhythm of chords cannot be captured with fixed time intervals (see Sheh & Ellis (2003); Yoshioka, Kitahara, Komatani, Ogata & Okuno (2004) for difficulties of chord boundary detection). Therefore, acoustical windows utilized in Leman (1995) and Leman and Carreras (1997) were not perceptually plausible. Second, neural representations utilized by Leman (1995) and Leman and Carreras (1997) occur early on the neural pathway of auditory perception to possibly function in association of chords. Psycho-acoustic theories of pitch perception (Parncutt, 1988; Terhadt, 1982; Terhardt, Stoll & Seewan, 1982a+b; Van Immerseel & Martens, 1992) model firing rates of auditory nerves (Cedolin & Delgutte, 2005; Moore, 1990). Auditory nerves transmit signals from cochlea to higher brain areas, at where more abstract feature detectors of pitch were located (Pantev et.al. 1988; Pantev et al. 1989). Therefore, it may be more appropriate to utilize more abstract features of pitch as input for the self-organizing mechanism.

Gjerdingen (1992) employed the ART 3 architecture (Carpenter & Grossberg, 1990) to train a self-organizing network sensitive to temporal patterns of chords. A short-term memory was employed to convert temporal patterns of chords into activation patterns of input. However, input representations did not " ... adequately represent the human perception of harmony." (Gjerdingen, 1992, p. 563). This model was not designed to capture the perception of relations between chords[3].

5.3 SIMULATIONS WITH SELF-ORGANIZING MAPS

In this thesis, neural network simulations were conducted to obtain self-organizing models of chord perception. Models were compared with MUSACT (Bharucha, 1987). Previous self-organizing models and training signals (Leman, 1995; Leman & Carreras, 1997; Tillmann et.al., 2000) provided a base for simulations.

The aim was to obtain self-organizing models functionally, but not structurally, equivalent to MUSACT (Bharucha, 1987). Tillmann et.al. (2000) obtained chord perception models that are structurally equivalent to MUSACT; but they depended on many questionable assumptions (see Section 5.2). Leman (1995) and Leman and Carreras (1997) showed the possibility of obtaining chord schema functionally equivalent to MUSACT without deforming the architecture of self-organizing maps. On the other hand, they did not compare the performance of their networks with experimental results on the perception of chords; and training signals of Leman (1995) and Leman and Carreras (1997) were psychologically unrealistic. Signals represented activation patterns of auditory nerves. More abstract features of pitches may be a better input signal to simulate association between chords (see Section 5.2). Training samples of were Leman (1995) and Leman and Carreras (1997) based on time intervals measured in seconds. Musical time is measured with beats, not with seconds.

Two simulations were conducted in this thesis. In the first simulation, training signals were chord categories; in the second simulation they were chord sequences

---

[3] It was designed to show the benefits of implementing a masking field in an ART3 architecture (see Gjerdingen, 1992 for details).

(see Sections 5.3.1 and 5.3.2). The structure of the self-organizing networks was kept constant between simulations. The self-organizing map was composed of 24X36 (864) units located on a torus. Activation of a unit reflected the cosine of the angle (correlation) between input vector and weight vector of the unit.

Before training, weights of units were randomized initially. Training was divided into two phases, rough organization and fine tuning. For all simulations, average quantization error (average distance between weight vector of best matching unit and corresponding input vector) was less than 0.1 unless noted otherwise. Simulations repeated many times in order to be sure that average quantization error did not decrease/increase considerably.

Self-organizing map simulations were run with SOM Toolbox 2.0 (Vesanto et.al., 2000), which is a MATLAB® library developed by Laboratory of Information and Computer Science in Helsinki University of Technology. Training data were assembled with Midi Toolbox (Eerola & Toiviainen, 2004), which is also a MATLAB® library developed by Department of Music Laboratory in University of Jyvaskyla.

After the training, performance of chord schemata was evaluated by simulating important experimental results on chord perception (Bhaucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986; Bigand & Pineau, 1997; Tekman & Bharucha, 1998; Krumhansl, 1990). These results were the corner stones of understanding perception of relations between chords; and they were simulated successfully with the self-organized models of MUSACT (Tillmann et.al. 2000).

Bharucha and Stoeckig (1986) demonstrated the facilitative performance on the target chord when the prime was one step away from the target on the circle of fifths compared to when it was six steps away. This study was the first demonstration of the priming paradigm and its relation with the circle of fifths. In order to simulate this result correctly, target unit activation before it was presented as input must be higher after the prime one step away compared to the six-steps-away prime. The self-organized model of MUSACT (Tillmann et.al., 2000) simulated this result successfully (see Figure 5.2) [4].

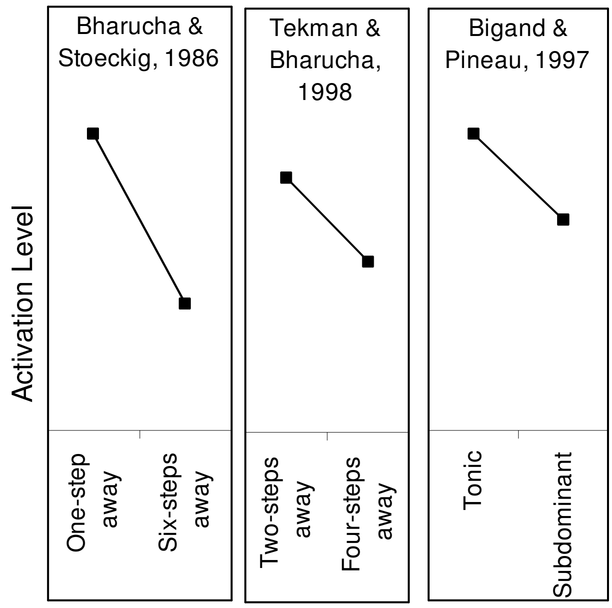


Figure 5.2 Activation levels of target units of the self-organizing chord schema of Tillmann et.al. (2000) after stimuli of Bharucha & Stoeckig (1986); Tekman & Bharucha (1998); and Bigand & Pineau (1997). Adopted from Tillmann et.al. (2000).The activation levels were obtained with the model that was trained with the sparse encoding and the big chord corpus.

Tekman and Bharucha (1998) demonstrated the facilitative performance on a target chord when the prime was two steps away from the target on the circle of fifths compared to when it was four-steps away. Target and four-steps-away prime share a pitch. Target and two-steps-away prime do not share a pitch. This study observed the

[4] For brevity, Tillmann et.al. (2000) reported only the simulation results obtained with the self-organized model of MUSACT that was trained with the sparse encoding and the big chord corpus.

fact that priming was not due to sharing pitches (i.e. not due to the acoustical similarity), but due to the distance on the circle of fifths. In order to simulate this result correctly, target unit activation before it was presented as input must be higher after the two-steps-away penultimate compared to four-steps-away penultimate. The self-organized model of MUSACT (Tillmann et.al., 2000) simulated this result successfully (see Figure 5.2).

Bigand and Pineau (1997) demonstrated the facilitative performance on a target chord when it was tonic compared to when it was subdominant of the chord sequence. This study observed the fact that harmonic function of a chord affects priming of the chord. In order to simulate this result correctly, activation of a tonic target before it was presented as input must be higher compared to activation of a subdominant target. The self-organized model of MUSACT (Tillmann et.al., 2000) simulated this result successfully (see Figure 5.2).

Krumhansl (1990) observed perceived stability of chords after major and minor musical contexts. Subjects listened to chord sequences in major and minor tonalities followed by a chord. They rated how well the last chord sounded after the sequence. Ratings were given in Table 7.3 of Krumhansl (1990). To simulate results, activation of the last chord before it was presented as input was collected; and activations were correlated with Table 7.3 of Krumhansl (1990). The correlation between stability ratings of chords and targets unit activations obtained from MUSACT[5] (Bharucha, 1987) is high and significant ($r=0.56$, $p<.01$, $N=24$ for the major context; $r=0.51$, $p<.05$, $N=24$ for the minor context). MUSACT simulated this result successfully.

---

[5] Tilllmann et.al. (2000) did not report this correlation. In this study, MUSACT (1987) was implemented with the code presented in Appendix A. See Appendix E for the testing stimuli, and target units.

Bhaucha & Krumhansl (1983) observed the perceived relation between chords. Subjects listened to a musical context in a major tonality followed by two chords. They rated how well the last two chords sounded together. Ratings of diatonic pairs were given in Table 8.2 of Krumhansl (1990). To simulate results, activation of the last chord before it was presented as input was collected; and activations were correlated with Table 8.2 of Krumhansl (1990). The correlation between the perceived relation between chords and targets unit activations obtained from MUSACT (Bharucha, 1987) is high and significant ($r=0.97$, $p<.01$, $N=36$)[6]. MUSACT simulated this result successfully.

5.3.1 TRAINING WITH CHORD CATEGORIES

In this simulation, self-organizing chord schemata were trained with chord categories. Chord categories were represented by vectors. Chord category vectors were obtained by the addition of the pitch category vectors that compose the chord. Pitch category vectors were constructed to reflect the perceived relations between pitch categories. Effects of different types of relations between pitch categories on the performance of self-organizing chord schemata were investigated.

Many geometric models were introduced to describe perceived relation between pitches (Lounget-Higgings & Steedman, 1971; Shepard, 1982; Chew 2001; Krumhansl, 2005). Chroma circle (Figure 2.1) captures cyclic relation of pitch categories (Shepard, 1964). Circle of fifths captures relation between pitches after a tonal context (Krumhansl & Shepard, 1979). Shepard (1982) suggested a torodial representation of pitches to capture features of chroma circle and circle of fifths together. Lounget-Higgings and Steedman (1971) located pitches on a lattice (see Figure 5.3). The lattice is two dimensional. On the horizontal axis pitches perfect

---

[6] Tilllmann et.al. (2000) did not report this correlation. In this study, MUSACT (1987) was implemented with the code presented in Appendix A. See Appendix E for the testing stimuli, and target units.

fifth apart were located nearby, reminding the circle of fifths. On the vertical axis pitches major third apart were located nearby[7].

A#	E#	B#	Fx	Cx	Gx
F#	C#	G#	D#	A#	E#
D	A	E	B	F#	C#
Bb	F	C	G	D	A
Gb	Db	Ab	Eb	Bb	F
Ebb	Bbb	Fb	Cb	Gb	Dbb

Figure 5.3 Lattice of pitches utilized by Euler (1739), Lounget-Higgings and Steedman (1971) and Chew (2001).

In this study, representations of pitch categories were designed to reflect the perceived relations between pitches. Pitch categories were represented with vectors; and the distance between vectors mirrored the similarity between pitch categories. Distance between vectors **a** and **b** is $\Sigma^n = sqrt((a_i-b_i)^2)$, where sqrt is the square root function, n is the number of elements of vectors, and a_i is the i^{th} element of the vector **a**.

[7] This representation was first suggested as a model of intonation by Euler (1739, cited in Lerdahl, 2001). Recently, Chew (2001) suggested a way to describe this relationship on a torus.

Five types of pitch category relations were studied: "orthogonal", "chroma", "circle of fifths", "chroma + circle of fifths", "major3rd + circle of fifths". In "orthogonal" representation distance between pitch categories was equal. Figure 5.4 presents "orthogonal" representation of some pitch categories. Figure 5.5 presents the distance between the pitch category "C" and other pitch categories in "orthogonal" representation.

C	1	0	0	0	0	0	0	0	0	0	0	0
C#(Db)	0	1	0	0	0	0	0	0	0	0	0	0
D	0	0	1	0	0	0	0	0	0	0	0	0
D#(Eb)	0	0	0	1	0	0	0	0	0	0	0	0
E	0	0	0	0	1	0	0	0	0	0	0	0
F	0	0	0	0	0	1	0	0	0	0	0	0
F#(Gb)	0	0	0	0	0	0	1	0	0	0	0	0

Figure 5.4 Some of the pitch category vectors of "orthogonal" representation.

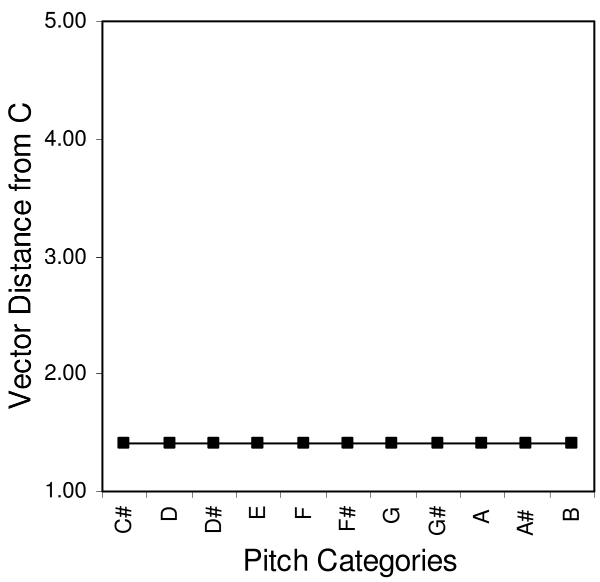


Figure 5.5 Distance between the pitch category "C" and other pitch categories in "orthogonal" representation.

In "chroma" representation distance between pitch categories approximated the chroma circle (Figure 2.1). Figure 5.6 presents "chroma" representation of some pitch categories. Figure 5.7 presents the distance between pith category "C" and other pitch categories in "chroma" representation.

C	1	1	1	1	1	1	0	0	0	0	0	0
C#(Db)	0	1	1	1	1	1	1	0	0	0	0	0
D	0	0	1	1	1	1	1	1	0	0	0	0
D#(Eb)	0	0	0	1	1	1	1	1	1	0	0	0
E	0	0	0	0	1	1	1	1	1	1	0	0
F	0	0	0	0	0	1	1	1	1	1	1	0
F#(Gb)	0	0	0	0	0	0	1	1	1	1	1	1

Figure 5.6 Some of the pitch category vectors of "chroma" representation.

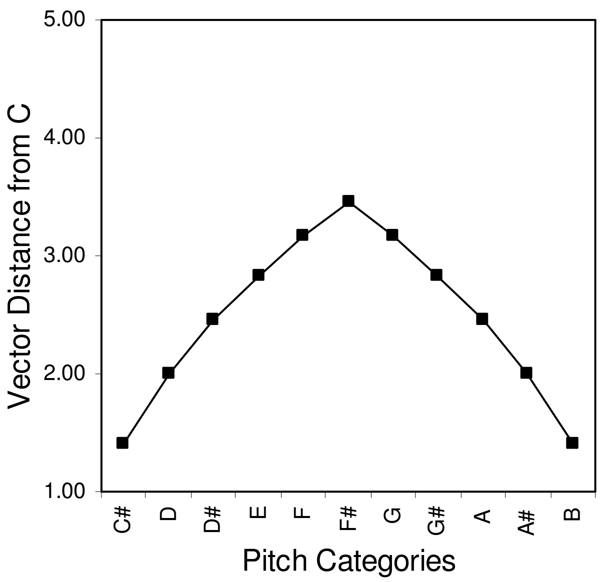


Figure 5.7 Distance between the pitch category "C" and other pitch categories in "chroma" representation.

In "circle of fifths" representation distance between pitch categories approximated the circle of fifths (Figure 2.2). Figure 5.8 presents "circle of fifths" representation of some pitch categories. Figure 5.9 presents the distance between pitch category "C" and other pitch categories in "circle of fifths" representation.

C	1	1	1	1	1	1	0	0	0	0	0	0
C#(Db)	1	0	0	0	0	0	0	1	1	1	1	1
D	0	0	1	1	1	1	1	1	0	0	0	0
D#(Eb)	1	1	1	0	0	0	0	0	0	1	1	1
E	0	0	0	0	1	1	1	1	1	1	0	0
F	1	1	1	1	1	0	0	0	0	0	0	1
F#(Gb)	0	0	0	0	0	0	1	1	1	1	1	1

Figure 5.8 Some of the pitch category vectors of "circle of fifths" representation.

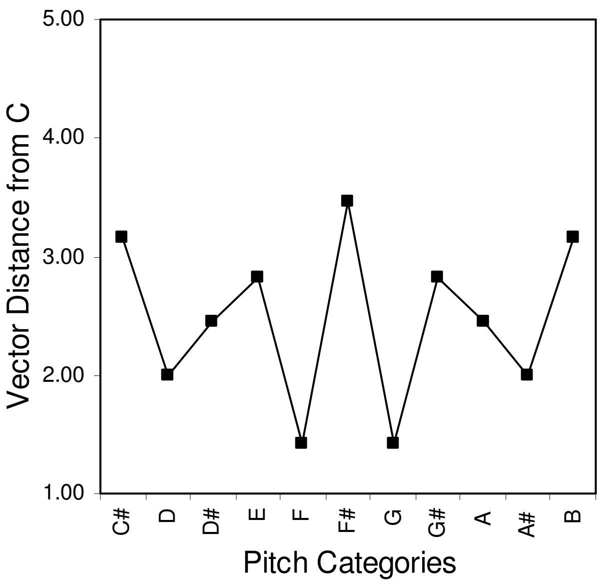


Figure 5.9 Distance between the pitch category "C" and other pitch categories in "circle of fifths" representation.

In "chroma + circle of fifths" representation distance between pitch categories approximated chroma circle and circle of fifths at the same time. This can be visualized by imagining a double helix wrapped around a torus (see Shepard, 1982, fig. 5). Figure 5.10 presents "chroma + circle of fifths" representation of some pitch categories. Figure 5.11 presents the distance between pitch category "C" and other pitch categories in "chroma + circle of fifths" representation.

C	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0
C#(Db)	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1
D	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0
D#(Eb)	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1
E	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0
F	0	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	0	0	0	1
F#(Gb)	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1

Figure 5.10 Some of the pitch category vectors of "chroma + circle of fifths" representation.

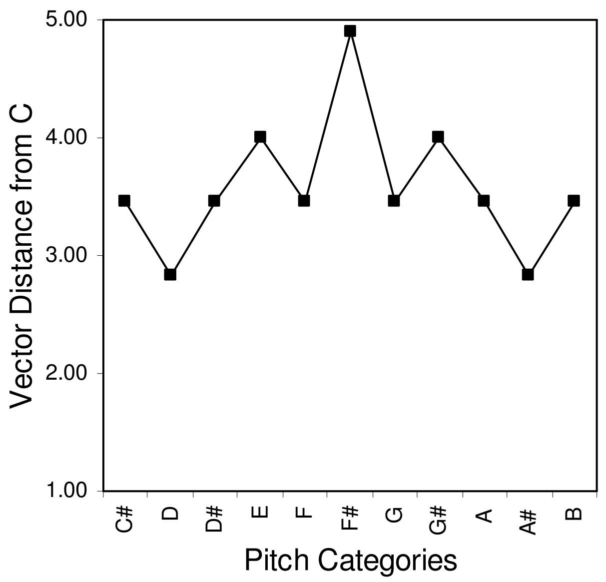


Figure 5.11 Distance between the pitch category "C" and other pitch categories in "chroma + circle of fifths" representation.

In "major3rd + circle of fifths" representation the distance between pitch categories approximated distances on the lattice of pitches (Figure 5.2). Figure 5.12 presents "major3rd + circle of fifths" representation of some pitch categories. Figure 5.13 presents the distance between pitch category "C" and other pitch categories in "major3rd + circle of fifths" representation.

C	1	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0
C#(Db)	1	0	0	1	1	0	0	0	0	0	0	1	1	1	1	1
D	0	0	1	1	0	0	1	1	1	1	1	1	0	0	0	0
D#(Eb)	0	1	1	0	1	1	1	0	0	0	0	0	0	1	1	1
E	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0
F	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1
F#(Gb)	0	0	1	1	0	0	0	0	0	0	1	1	1	1	1	1

Figure 5.12 Some of the pitch category vectors of "major3rd + circle of fifths" representation.

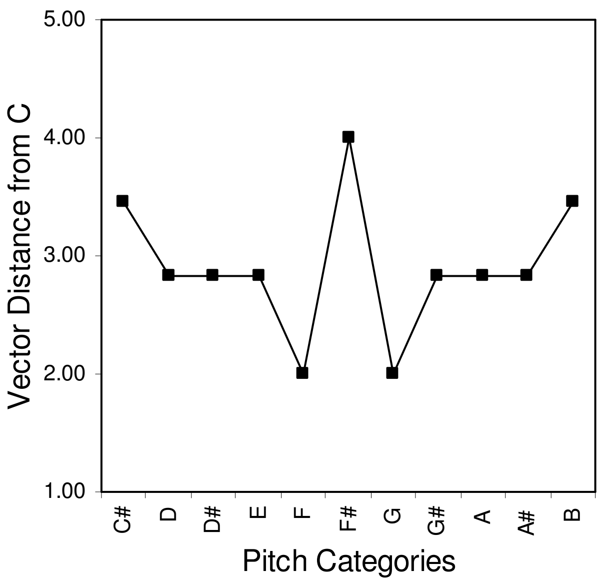


Figure 5.13 Distance between the pitch category "C" and other pitch categories in "major3rd + circle of fifths" representation.

Self-organizing maps (Kohonen, 1982) composed of 24X36 (864) units located on a torus were trained with chord category vectors. Chord category vectors were obtained by the addition of three pitch category vectors that represent pitches of the chord. Training data consisted of 24 vectors, 12 major and 12 minor chord categories. The same network architecture was trained with five different data sets derived from five different pitch category representations, i.e. "orthogonal", "chroma", "circle of fifths", "chroma + circle of fifths", "major3rd + circle of fifths".

Training of networks was done with the self-organizing map toolbox (Vesanto et.al., 2000) (see above for description of training procedure). This toolbox was developed to offer the correct implementation of the training algorithms of self-organizing maps (Kohonen, 1982). After training, 12 major and 12 minor chord categories were labeled on the map. Target units were located with reference to these labels.

In the test session, activation of target chords was read before they were presented as input. (see Appendix E for test stimuli and target units.) Short-term memory model of Toiviainen and Krumhansl (2003) was utilized to capture the temporal relation between input sequences.

According to this short-term memory model, accent (importance) of a pitch is a function of duration of the pitch, and accent gradually decreases as time passes. This is calculated with the following equation (Toiviainen & Krumhansl, 2003, eq. 2).

$$\overset{o}{p}_k = -\frac{p_k}{\tau_p} + \sum_{i=1}^{n_k} \delta\left(t - t_{ki}^{n}\right) - p_k \sum_{i=1}^{n_k} \delta\left(t - t_{ki}^{f}\right) \quad (5.5)$$

In this equation $\overset{o}{p}_k$ is the time derivative of kth pitch category p_k; δ is the Dirac delta function (unit impulse function); τ_p is time constant set to 0.5 seconds. Onset and offset times of tones having kth pitch are denoted by t_{ki}^{n} and t_{ki}^{f} respectively.

n_k represents total number of times the k[th] pitch occurs. With equation 5.5 pitch accents "saturates at about 1 s after tone onset approximating the durational accent as a function of tone duration (Parncutt 1994)." (Toiviainen & Krumhansl, 2003, p. 11).

Durational accents were integrated into memory, which was subject to a decay (Toiviainen & Krumhansl, 2003, eq. 3)

$$\overset{o}{m} = \frac{p - m}{\tau_m} \tag{5.6}$$

where τ_m is set to 3 seconds which estimated capacity of auditory sensory memory (Toiviainen & Krumhansl, 2003, p. 11). These functions were implemented in Midi Toolbox (Eerola & Toiviainen, 2004).

Simulation results were presented in Figures 5.14-5.18 and Table 5.1. In order to simulate the result of Bharucha and Stoeckig (1986) correctly, activation of the target unit must be higher after the input that is one-step-away from the target, compared to the input that is six-step-away. Self-organized chord schemata trained with different pitch category representations simulated the result successfully (Figure 5.14).

In order to simulate the result of Tekman and Bharucha (1998) correctly, activation of the target unit must be higher after the input that is two-steps-away from the target, compared to the input that is four-step-away. Only the self-organized chord schemata trained with "circle of fifths" and "chroma + circle of fifths" pitch category representations simulated the result successfully (Figure 5.15). Other networks were failed to simulate this study.

In order to simulate the result of Bigand and Pineau (1997) correctly, activation of the tonic target unit must be higher, compared to the subdominant target unit. Self-organized chord schemata trained with different pitch category representations simulated the result successfully (Figure 5.16).

Correlations between stability ratings of chords (Table 7.3, Krumhansl, 1990) and target unit activations for self-organizing chord schemata that are trained with different pitch category representations are presented in Figure 5.17. Correlations were high and significant. The highest correlation was obtained with the self-organizing chord schema trained with "orthogonal" representation ($r=0.8$, $p<.01$, $N=24$ for the major context; $r=0.76$, $p<.01$, $N=24$ for the minor context). The lowest correlation was obtained with the self-organizing chord schema trained with "circle of fifths" representation ($r=0.67$, $p<.01$, $N=24$ for the major context; $r=0.64$, $p<.01$, $N=24$ for the minor context). But all of the correlations were higher than the correlation calculated with target unit activations of MUSACT ($r=0.56$, $p<.01$, $N=24$ for the major context; $r=0.51$, $p<.05$, $N=24$ for the minor context) (not presented in Figure 5.17).

Correlations between the perceived relation between chords (Table 8.2, Krumhansl, 1990) and target unit activations for self-organizing chord schemata that are trained with different pitch category representations are presented in Figure 5.18. Correlations were high and significant. The highest correlation was obtained with the self-organizing chord schema trained with "circle of fifths" representation ($r=0.86$, $p<.01$, $N=36$). The lowest correlation was obtained with the self-organizing chord schema trained with "chroma" representation ($r=0.35$, $p<.05$, $N=36$). But all of the correlations were lower than the correlation calculated with target unit activations of MUSACT ($r=0.97$, $p<.05$, $N=36$) (not presented in Figure 5.18).

Correlations between self-organizing models were high (Table 5.1). The highest correlation was obtained between the self-organized chord schemata trained with "circle of fifths" and "chroma + circle of fifths" pitch category representations ($r=1$, $p<.01$, $N=96$). Correlations between MUSACT and self-organizing models were not high, but they were significant (Table 5.1). The highest correlation with MUSACT was obtained with the self-organized chord schemata trained with "circle of fifths" and "chroma + circle of fifths" pitch category representations ($r=0.43$, $p<.01$, $N=96$).

As a result, self-organizing models of chord perception trained with different pitch representations gave promising results for modeling of chord perception studies with self-organizing architectures dissimilar to MUSACT (Bharucha, 1987). On the other hand, to simulate all experimental results on perception of relation between chords correctly, and to achieve the highest correlation with MUSACT (Bharucha, 1987), it was necessary to train the self-organizing architecture with "circle of fifths" or "chroma + circle of fifths" pitch category representations.

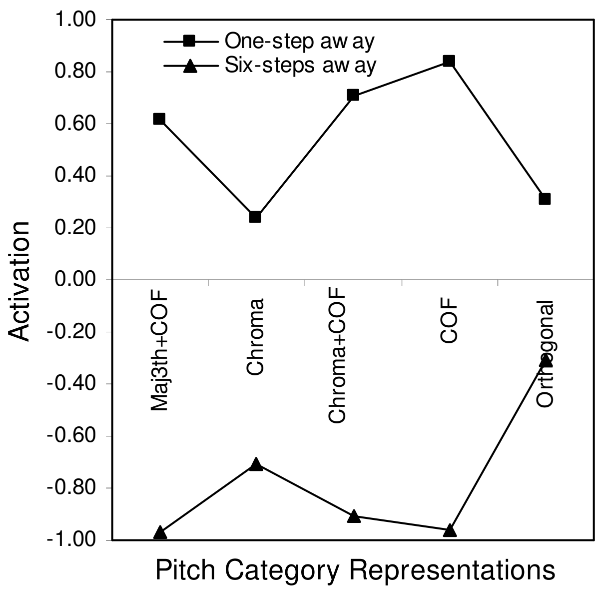


Figure 5.14 Activation of target units after stimuli of Bharucha & Stoeckig (1986) on models trained with different pitch category representations.

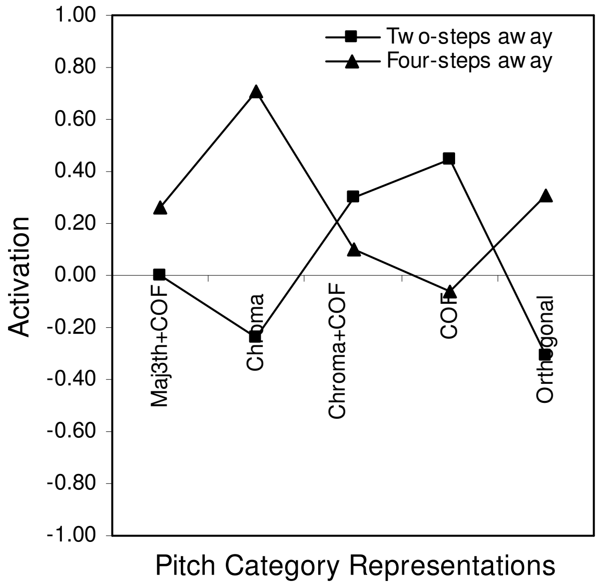


Figure 5.15 Activation of target units after stimuli of Tekman & Bharucha (1998) on models trained with different pitch category representations.

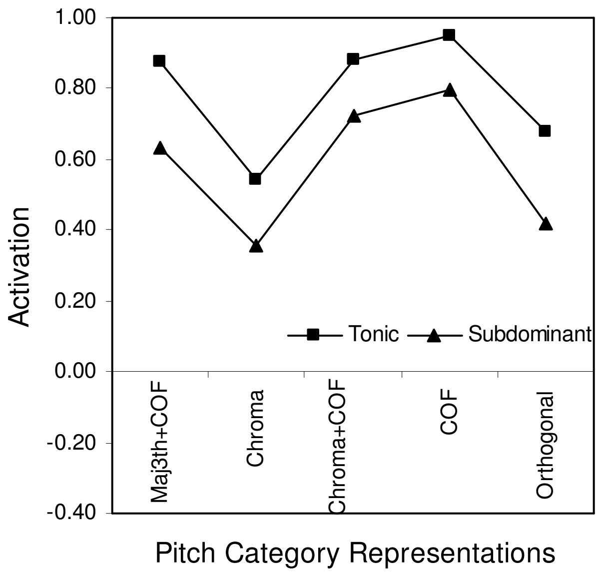


Figure 5.16 Activation of target units after stimuli of Bigand & Pineau (1997) on models trained with different pitch category representations.

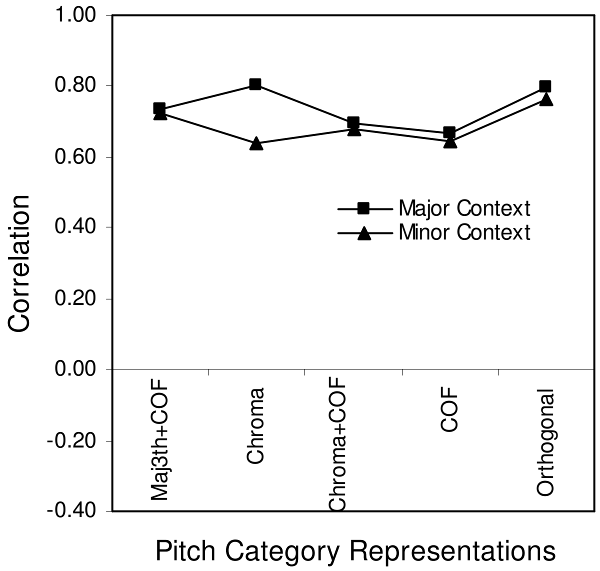


Figure 5.17 Correlations between stability ratings of chords (Krumhansl. 1990, table 7.3) and targets unit activations. (N=24, p<.01 for all correlations.)

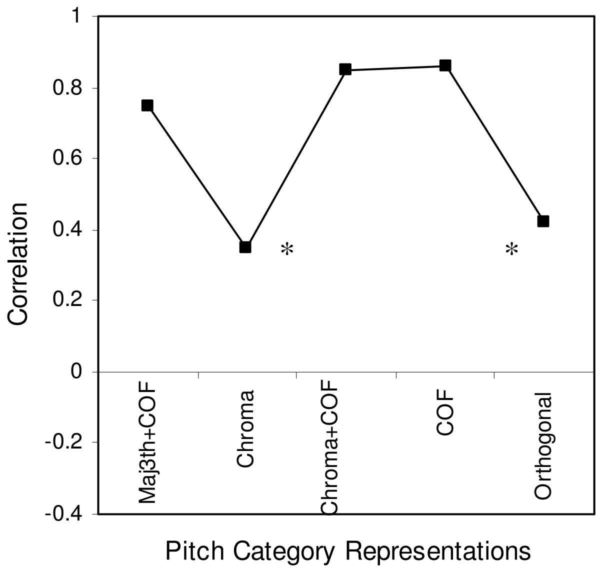


Figure 5.18 Correlations between ratings of relation between diatonic chords (Krumhansl. 1990, table 8.2) and targets unit activations. (N=36, * p<.05, p<.01 for all other correlations.)

Table 5.1 Correlations between activation of target units of self-organizing models that were trained with various pitch category representations (Maj3rd+COF, Chroma, Chroma, COF, Orthogonal) and MUSACT (Bharucha, 1987). (N=96, p<.01 for all correlations.) for simulations of Section 5.3.1.

	Maj3rd+COF	Chroma	Chroma+COF	COF	Orthogonal
Chroma	0.82				
Chroma+COF	0.98	0.72			
COF	0.96	0.65	1.00		
Orthogonal	0.92	0.90	0.84	0.79	
MUSACT	0.42	0.31	0.43	0.43	0.37

5.3.2 TRAINING WITH SIMPLE CHORD SEQUENCES

In this simulation, effects of a short-term memory model, which can differentiate what is playing now from what was played in the past, on the self-organization of chord schemata were investigated.

Training self-organizing networks with chord categories is unrealistic. Category exemplars of chords are not found independently in the musical environment. In addition, musical pieces heavily depend on temporal relations between chords. Therefore, samples of music would function better as training stimuli. But this approach leads to another difficulty. Temporal relations between chords must be represented in the input signal. Though there are neural network structures to capture temporal dependencies without explicitly encoding time in the input signal (see for example Elman, 1990), self-organizing maps are not of this kind.

Temporal relation between chords has been represented in the training set by using short-term memory models of pitch. The most common approach was to integrate incoming signal into the short-term memory buffer, which was subject to decay (Bharucha, 1987; Bharucha & Todd, 1990; Leman, 1995; Toiviainen & Krumhansl, 2003). However, this approach has a deficiency. It was impossible to differentiate current and past signals by observing contents of the short-term memory. To remedy the problem, Krumhansl, Louhivuori, Toiviainen, Järvinen, and Eerola (1999) suggested a diffusion memory model of short-term memory. In this model there were several short-term buffers ordered in time. Content of a buffer diffused into the next at some rate. Events of present and past could be differentiated accurately. In addition, time resolution of past events was worse compared to recent events, which is an important feature of short-term memory of pitches. Krumhansl et al. (1999) designed a diffusion memory based on six buffers. Their model successfully simulated effects of learning melodies on melodic expectation.

The approach taken here was similar to that of Krumhansl et al. (1999). In this thesis, there were two short-term memory buffers. The first one represented pitch content across a sampling window. The second one represented the integration of pitches starting from the next window (see Figure 5.8). The second buffer was identical to the short-term memory model of Toiviainen and Krumhansl (2003) (see Formulae 5.5 and 5.6); but here it follows one-step behind. In other words, a memoryless model and a model of integration was concatenated to obtain the short-term memory model of this simulation.

The training set of this simulation was composed of musical pieces. Leman (1995) and Leman and Carreras (1997) also trained self-organizing maps with musical pieces. Their sampling windows were located with reference to seconds. This is not realistic given the fact that musical lengths are measured with beats, not with seconds (Lerdahl & Jackendoff, 1983). In this simulation, sampling windows were located with reference to beats of musical pieces.

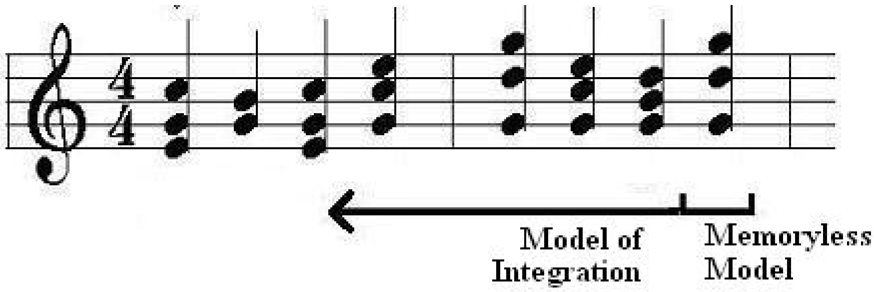


Figure 5.19 Schematic description of the short-term memory model utilized in simulations of Section 5.3.2.

The same self-organizing map architecture of Section 5.3.1 was used in this simulation as well. Training data were derived from 24 chord sequences (see Appendix F)[8]. Each sequence was divided into 30 windows that were two quarter beats long. The abovementioned short-term memory model was applied to each window. Training data were composed of 720 (24X30) signals vectors. The same network architecture was trained with five different data derived from five different pitch category representations ("orthogonal", "chroma", "circle of fifths", "chroma + circle of fifths", "major3rd + circle of fifths"). By this way effects of perceived relation between pitch categories on the performance self-organization of chord schemata were investigated.

After the training and before the test session, units representing 12 major and 12 minor chord categories were labeled on the map by presenting category exemplars of 12 major and 12 minor chords. Target units of the test session were located with reference to these labels. Test session was identical to the previous session.

---

[8] The same training set was used by Leman (1995, fig. 8.7) as well.

Simulation results are presented in Figures 5.20-5.24 and Table 5.2[9]. Results are almost identical to the result of Section 5.3.1. Training with chord sequences did not change the performance of self-organizing chord schemata significantly. Correlations between activations of MUSACT and self-organizing models were still significant, but they dropped by a small amount (Table 5.2) compared to the correlations of Section 5.3.1 (Table 5.1).

It may be the case labeling the map with category exemplars of chords was inappropriate given that the map was trained on data including information about past events. To test this hypothesis, chord category units were labeled with the last window of the chord sequences. Chord sequences always ended with the tonic. Therefore, the last windows represented 12 major and 12 minor chord categories situated in a musical context. However, labeling units with chord representations that include musical context did not change simulation results.

As a result, enriching the training set by including information about the musical context did not improve the simulation success of self-organizing chord-schemata; and slightly decrease the correlation with MUSACT (Bharucha, 1983). The same result was obtained by Tillmann et.al. (2000). In Tillmann et.al. (2000) performance of different chord schemata trained with different training sets was similar; and Tillmann et.al. (2000) observed the lowest correlation with MUSACT (Bharucha, 1983) with the self-organized chord schema that was trained with the rich encoding and the big corpus.

---

[9] It was not possible to compare results of this simulation with results of Leman (1995, chapter. 8), which were obtained with the same training set, because Leman (1995, chapter. 8) focused particularly on similarities between tonality categories.

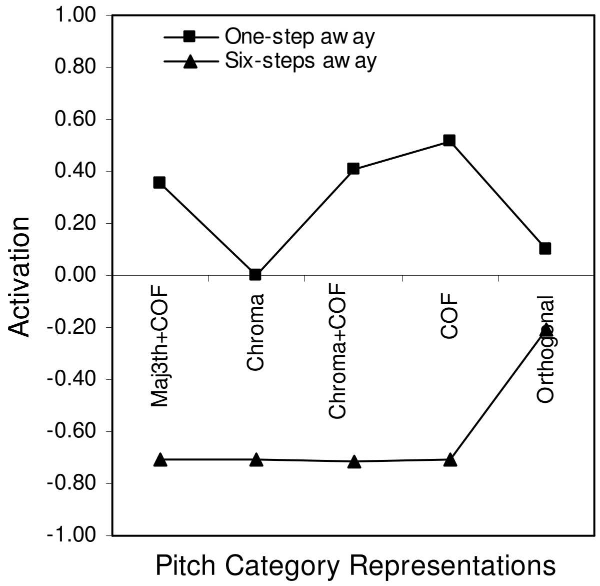


Figure 5.20 Activation of target units after stimuli of Bharucha & Stoeckig (1986) on models trained with simple chord sequences.

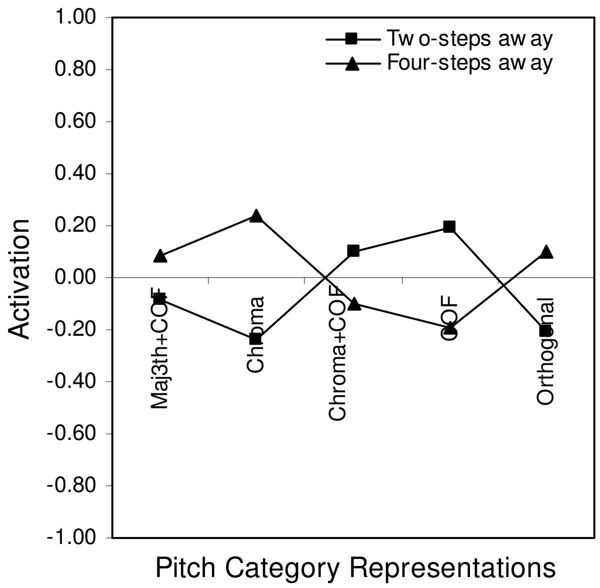


Figure 5.21 Activation of target units after stimuli of Tekman & Bharucha (1998) on models trained with simple chord sequences.

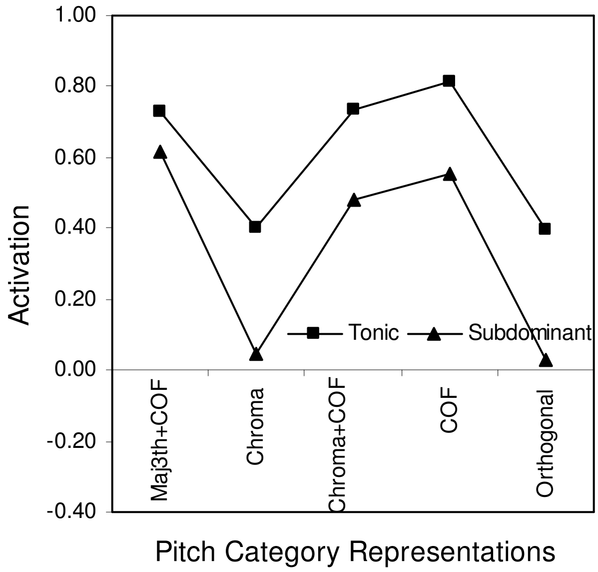


Figure 5.22 Activation of target units after stimuli of Bigand & Pineau (1997) on models trained with different simple chord sequences.

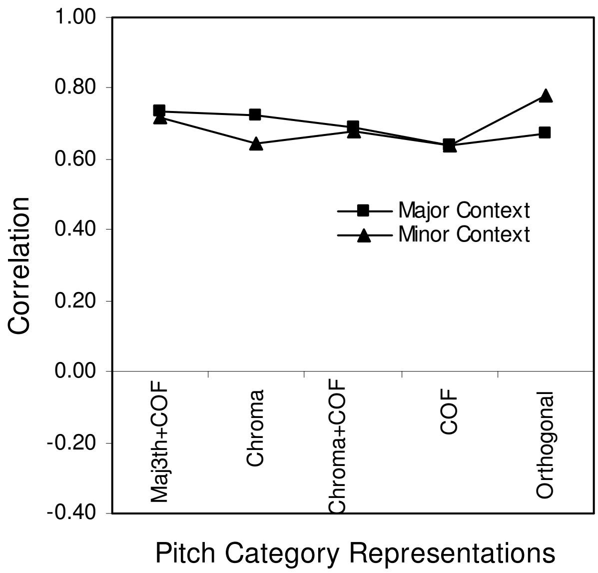


Figure 5.23 Correlations between stability ratings of chords (Krumhansl. 1990, table 7.3) and targets unit activations. (N=24, p<.01 for all correlations.)

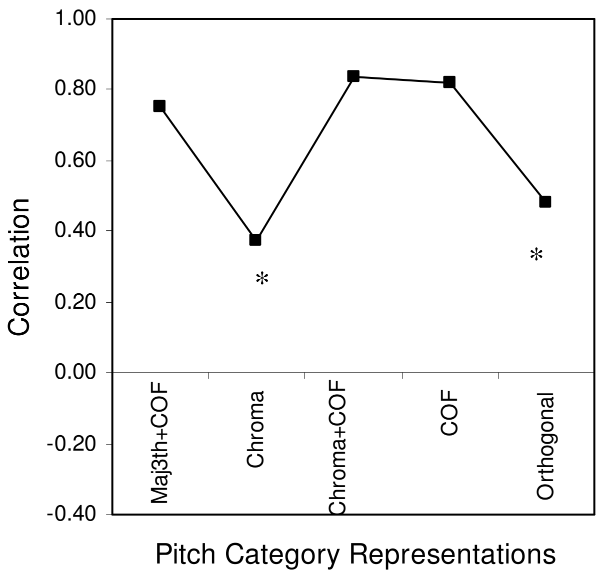


Figure 5.24 Correlations between ratings of relation between diatonic chords (Krumhansl. 1990, table 8.2) and targets unit activations. (N=36, * p<.05, p<.01 for all other correlations.)

Table 5.2 Correlations between activation of target units of self-organizing models that were trained with various pitch category representations (Maj3rd+COF, Chroma, Chroma, COF, Orthogonal) and MUSACT (Bharucha, 1987). (N=96, * p<.05, p<.01 for all other correlations.) for simulations of Section 5.3.2.

	Maj3rd+COF	**Chroma**	**Chroma+COF**	**COF**	**Orthogonal**
Chroma	0.73				
Chroma+COF	0.96	0.67			
COF	0.94	0.60	0.99		
Orthogonal	0.84	0.82	0.83	0.79	
MUSACT	0.36	0.23*	0.36	0.36	0.30

CHAPTER 6

SUMMARY AND DISCUSSION

In this thesis, the role of non-diatonic chords in perception harmony was investigated. Perception of chords is one of the most studied topics of music perception (see Krumhansl, 1990 & Tillmann, Bharucha & Bigand, 2000 for reviews). Studies have shown that chords have different perceptual status after musical context; and their perceptual characteristics are determined by the tonality of the musical context (see Krumhansl, 1990 for a review). After musical context, diatonic chords are perceived as more stable compared to non-diatonic chords; and there is a hierarchy of stability between diatonic chords. According to this hierarchy tonic is the most stable chord, followed by the dominant, subdominant and mediant, followed by the other degrees of the tonality. There is a similarity between music theoretical description of importance of chords and music psychological finding of stability of chords. In addition, it has been shown that perceptual relation between chords is based on stability of chords (see Krumhansl, 1990 for review). The connection between stability and perceptual relatedness has been described with principles of chord perception: contextual identity, contextual distance and contextual asymmetry (Bharucha & Krumhansl, 1983). Contextual identity states that tokens of a stable chord are perceived as more similar compared to tokens of an unstable chord. Contextual distance states that stable chords are perceived as more related compared to unstable chords. Contextual asymmetry states that the last chord contributes more to the perceptual relatedness of a chord pair. Bharucha (1987) proposed that implicit knowledge about pitches, chords and tonalities (i.e. tonal schema) govern these principles; and he presented a neural network implementation of tonal schema, which is called MUSACT. Tillmann, Bharucha and Bigand (2000) showed that MUSACT may self-organize as a result of tonal regularities in the musical environment. Another model of tonal schema is Tonal Pitch Space (Lerdahl, 2001); and it is a symbolic description of tonal aspects of the musical mind.

Theories of harmony give formal descriptions of the relations between chords (Piston, 1978); and among them, harmonic relations between non-diatonic and diatonic chords have interesting features. For example, the Neapolitan (a non-diatonic chord) is related with the dominant (a diatonic chord); and the Neapolitan chord may precede the dominant. Another example, every diatonic chord is associated with a major and a diminished secondary dominant chords; and secondary dominants must precede, and must not succeed, their diatonic associates[1]. Music theory relates specific types of non-diatonic (unstable) chords with specific types of diatonic (stable) chords; and defines a sequential relation between them. Music psychology, on the other hand, relies only on stability to describe perceptual relation between chords (Bharucha & Krumhansl, 1983). Empirical investigation of the relation between diatonic (stable) and non-diatonic (unstable) chords may reveal important aspects of chord perception. In this thesis, the role of the Neapolitan and secondary dominant chords in the perception of relation between chords was investigated.

The perceptual relation between chords has been investigated with memory tasks (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha & Castelleno, 1982), similarity ratings (Bharucha & Krumhansl, 1983; Krumhansl, 1990; Krumhansl, Bharucha & Castelleno, 1982; Krumhansl, Bharucha & Kessler, 1982) and priming studies (Bharucha, 1987; Bharucha & Stoeckig, 1986, 1987; Bigand & Pineau, 1997; Bigand, et. al. 1999; Tekman & Bharucha, 1992, 1998). In this thesis chord priming paradigm was used. In chord priming, the effect of a musical context on the processing of a target chord is investigated (Bharucha & Stoeckig, 1986), by asking participants to make a binary decision on a perceptual quality of the target chord.

[1] For a detailed description of these chord types see Section 3.1.

Because of this simple task demand, priming paradigm is suitable to collect data from both musicians and non-musicians. Facilitated performance on the target chord indicates expectation towards the target chord. Expectation for an incoming chord after musical context reflects the perceptual association between the musical context and the incoming chord, which is governed by tonal schema (Bharucha, 1987). Tonal schema is implicit knowledge about pitches, chords and tonalities that develops as a result of tonal regularities in the musical environment (Tillmann, et.al., 2000). Chord priming is an indirect measure of implicit knowledge about chords (Bharucha, 1987).

Chord priming has been observed under many and different experimental settings (Bharucha and Stoeckig 1986, 1987; Bigand & Pineau, 1997; Bigand, Madurell, Tillmann & Pineau 1999; Bigand, Poulin, Tillmann, & D'Adamo, 2003; Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001; Bigand, Tillmann, Poulin-Charronnat & Manderlier, 2006; Justus & Bharucha, 2002; Tekman & Bharucha, 1992, 1998; Tillmann & Bigand, 1999; Tillmann, Bigand & Pineau, 1998; Tillmann, Janata, Birk, & Bharucha, 2003); and the typical finding has been that primes that are closer to the target on the circle of fifths facilitates the processing of the target chord compared to distant primes. Circle of fifths (Figure 2.2) is a music theoretical figure to describe relations between chords, on which dominant and subdominant chords of a tonality are placed next to the tonic of the tonality. Tonic, subdominant, and dominant chords are stable chords. Therefore, chord priming experiments have been in agreement with the proposed connection between chordal relation and stability. Circle of fifths is an emergent property of MUSACT (Bharucha, 1987): activation level of major chord and major tonality units decreases as the distance from the input chord increases on the circle of fifths. Many significant results of the chord priming has been simulated with MUSACT (Tillmann et.al., 2000).

In this thesis, the role of the Neapolitan and secondary dominant chords on the priming of chords was investigated. The first three experiments of this thesis investigated the priming of the dominant with the Neapolitan chord. Music theoretically, the Neapolitan and the dominant are related chords; and the Neapolitan may precede the dominant chord. On the other hand, they are the most distant chords on the circle of fifths. Priming the dominant chord with the Neapolitan chord is contrary to explanations of chord priming. However, Atalay (2002) observed facilitative performance for the dominant target when the penultimate chord was the Neapolitan chord. Experiment 1 investigated whether this was simply due to the distance on the circle of fifths. In Experiment 1, dominant and subdominant target chords were preceded by penultimate chords either three or six steps away from the targets. The subdominant and the six-steps-away penultimate are not related music theoretically. Penultimate that is six steps away from the dominant target is the Neapolitan chord. Results of Experiment 1 showed that facilitated performance due to the six-steps-away penultimate was observed only from dominant targets. This was true for both musicians and non-musicians. As a result, there was a unique association between the Neapolitan chord and the dominant; and facilitated processing of the dominant target was not due to the distance on the circle of fifths. Predictions of MUSACT (Bharucha, 1983) and Tonal Pitch Space (Lerdahl, 2001) were compared with results. MUSACT (Bharucha, 1983) predicted that penultimate chords that were closer to the target on the circle of fifths would facilitate responses irrespective of the target type. Contrarily, Tonal Pitch Space theory (Lerdahl, 2001) predicted that penultimate chords that were distant to the target on the circle of fifths would facilitate responses irrespective of the target type. Predictions of MUSACT (Bharucha, 1983) were correct for subdominant targets; and predictions of Tonal Pitch Space (Lerdahl, 2001) were correct for dominant targets.

Experiment 1 utilized the piano timbre. Stimuli composed with the piano timbre contain melodic factors (i.e. pitch height) that might effect tonal expectation (Bharucha, 1996). It might be the case that results of Experiment 1 were due to melodic factors but not the association between the Neapolitan and dominant chords per se. In addition, Tonal Pitch Space (Lerdahl, 2001) incorporates melodic factors to calculate expectation between chords; but MUSACT (Bharucha, 1983) does not. Therefore, it was unfair to compare predictions of Tonal Pitch Space (Lerdahl, 2001) and MUSACT (Bharucha, 1983). To test effects of melodic factors on the results, stimuli composed with Shepard (1964) tones used in Experiment 2. Shepard tones emphasize class pitch class at the expense of pitch height, and create the illusion of the circularity of pitch (Shepard, 1964). Chord sequences composed of Shepard tones might reduce effects of pitch height on tonal expectation. In Experiment 2, dominant and subdominant target chords were preceded by penultimate chords either two or six steps away from the targets. Results of musically educated and naive participants were different. Musically educated participants gave facilitative responses to the dominant targets after the six-steps-away penultimate, and they gave facilitative responses the subdominant targets after the two-steps-away penultimate. Musically naive participants, on the other than, gave facilitative responses for both the subdominant and the dominant targets after the two-steps-away penultimate. Predictions of MUSACT (Bharucha, 1983) and Tonal Pitch Space (Lerdahl, 2001) were compared with the results. MUSACT (Bharucha, 1983) predicted that penultimate chords that were closer to the target on the circle of fifths would facilitate responses irrespective of the target type. Contrarily, Tonal Pitch Space theory (Lerdahl, 2001) predicted that penultimate chords that were distant to the target on the circle of fifths would facilitate responses irrespective of the target type. Predictions of MUSACT (Bharucha, 1983) were correct for responses of musically educated participant for subdominant targets, and responses of musically naive participant for subdominant and dominant targets. Predictions of Tonal Pitch Space (Lerdahl, 2001) were correct only for responses of musically educated participant for dominant targets.

When melodic factors that are effective on tonal expectation were reduced, only musically educated participants showed an expectation for the dominant chord after the Neapolitan chord; responses of musically naive participants were based on the distance on the circle of fifths. This picture is different from the findings of

Experiment 1, where melodic factors were effective. In Experiment 1 expectation towards the dominant chord after the Neapolitan was observed both from musicians and non-musicians. Therefore, it is reasonable to conclude that only musicians associated the Neapolitan chord and the dominant chord; and responses of non-musicians for the dominant targets in Experiment 1 were governed mainly by melodic factors. Musicians exposed to musical stimuli extensively. Therefore, musicians have chance to acquire detailed statistical knowledge about the relation between chords.

Experiment 1 was conducted in Turkey with Turkish non-musicians. Experiment 2 was conducted in Finland with European non-musicians. The population from which non-musicians of Experiment 1 came reported high rates of familiarity and positive evaluation of Western and Turkish musical styles (Tekman & Hortaçsu, 2002). Tonal structures of Turkish and Western musical styles are dissimilar. It may be the case that responses of Turkish non-musicians in Experiment 1 were governed by tonal schema of Turkish music; therefore, non-musicians of Experiment 1 and Experiment 2 might not come from the same population. This problem did not exist for Turkish musicians of Experiment 1, who were students of conservatories with a curriculum of Western tonal music.

In order to understand effects of melodic factors on the responses of non-musicians better, Experiment 3, which utilized Shepard (1964) tones, was conducted with Turkish non-musicians. It also investigated effects of listening to scales and chord sequences on the observation of chord priming. In Experiment 3, dominant and subdominant target chords were preceded by penultimate chords either three or six steps away from the targets. Penultimate chords were preceded either by chord sequences, or scales. Chord priming effect was not observed in Experiment 3 from Turkish non-musicians. It might be the case that the task demand of Experiment 3, which was timbre discrimination task, was too easy to reveal any chord priming effect. Indeed, responses of participants were very fast (average reaction time was below 750 msec. for all of the conditions of the experiment); and percentage of errors decreased with practice.

Experiment 3 did not help explain the difference between the results of Experiments 1 and 2. In Experiments 1 and 2, the perceptual association between the Neapolitan and the dominant was observed in musicians. Responses of musicians were governed by the distance of circle of fifths for subdominant targets, which showed that the association between the Neapolitan and dominant is an exception. On the other hand, non-musicians perceived the relation between the Neapolitan and the dominant in Experiment 1, but not in Experiment 2. In Experiment 1, melodic factors were operating, and Turkish non-musicians perceived the investigated relation. In Experiment 1, melodic factors were diminished and European non-musicians did not perceive the relation. It may be the case that cultural factors had a role on the results in non-musicians. Turkish music is a melodic musical culture, and responses of Turkish non-musicians were governed by their knowledge on Turkish music. Indeed, the Kurdi maqam of Turkish music includes the base note of the Neapolitan chord. In order to understand effects of melodic and cultural factors on the results, Experiment 1 may be repeated with Western non-musicians and Experiment 2 with Turkish non-musicians.

As an alternative hypothesis, it may be claimed that the perceptual association between the Neapolitan and the dominant is governed by the structure and the function of the auditory processing system, but not by learned knowledge of chords. This hypothesis has difficulties in explaining the data, given the fact that the Neapolitan and the dominant chords of a musical context are the subdominant and the-six-steps-away penultimate chords of another musical context. The function of a chord is relative to the context, and the same chord has a different function in a different context. In addition, this hypothesis failed to explain the observed difference between musicians and non-musicians. As another hypothesis, it may be claimed that the knowledge of the relation between the Neapolitan and the dominant has already been encoded in the brain after the birth; and this knowledge comes into play only if the mind/brain encounters with the critical stimuli. Because non-musicians rarely encounter with the Neapolitan chord, their knowledge was not activated, and thus, they were not able to perceive the relation between the Neapolitan chord and the dominant. This hypothesis can be tested by observing the perception of non-musicians, and then playing musical pieces that include the Neapolitan chord to them, and test them again to see if their perception has changed.

As a future study, it may be interesting to investigate the time course of priming of the Neapolitan chord. It may be the case that cognitive processes responsible for the priming of the Neapolitan chord and for the priming of diatonic chords are different. Different sources of chord priming may reveal themselves in the time course of the priming effect. Indeed, when the prime is single chord, the chord priming is first governed by the acoustical similarity between prime and target, and then by the distance on the circle of fifths (Tekman & Bharucha, 1992). Similarly, the priming characteristics of the Neapolitan chord may change within time. Chord priming is one of the methods to investigate implicit knowledge about chords. The other methods are recognition memory tasks (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha & Castelleno, 1982) and similarity ratings (Bharucha & Krumhansl, 1983; Krumhansl, 1990; Krumhansl, Bharucha & Castelleno, 1982; Krumhansl, Bharucha & Kessler, 1982). It may be interesting to study the perceptual relation between the Neapolitan chord and the dominant with these methods as well. It may be better to conduct future studies about the Neapolitan chord in countries where both non-musicians and musicians unquestionably develop tonal schema of Western music. By this way, differences between non-musicians and musicians will definitely be interpreted as reflections of the effects of musical expertise on chord priming.

The last three experiments of this thesis investigated the role of secondary dominant chords in perception of harmony with non-musicians. Secondary dominant is a non-diatonic chord. Music theoretically, every diatonic chord is associated with one major and one diminished secondary dominant chords; and secondary dominants must precede, and must not succeed, their diatonic associates (Piston, 1978). Experiments 4, 5 and 6 investigated the perception of sequential relation between secondary dominant chords and their diatonic associates from non-musicians by observing its effects on expectations towards tonic and subdominant targets after chord sequences (i.e. by observing global chord priming). The idea was to scramble the order of chords in sequences; and to observe the effect of scrambling on the global priming of chords. Secondary dominant chords of the scrambled sequences violated the sequential relation between themselves and their diatonic associates. Experiments 4, 5 and 6 investigated whether this violation disturbs the priming of chords. Chord sequences of Experiments 4, 5 and 6 were selected from Bach chorales (Greentree, 2005). Half of the chord sequences were composed of diatonic chords; and half of them include secondary dominant chords. Secondary dominant chords were utilized in chord sequences in their original form. In Experiments 4, 5 and 6, scrambled chord sequences were contrasted with original versions; and effects of scrambling on global chord priming was observed. Predictions of MUSACT (Bharucha, 1987) and Tonal Pitch Space theory (Lerdahl, 2001) were compared with the results.

In Experiment 4, sequences were played with Shepard (1964) tones; and they were scrambled with 2by2 scrambling algorithm[2]. Global priming of chord was observed from non-musicians in Experiment 4. Disruption of the global chord priming due to the violation of the sequential relation between the secondary dominant chords and their diatonic associates was not observed. However, global chord priming was disturbed by the use of secondary dominant chords, both in their original or in scrambled form. Shepard (1964) tones emphasize pitch class at the expense of pitch height. Pitch height is an important factor in harmony (Piston, 1978). Experiment 5 was designed to observe the effects of pitch height on the perception of the relation

[2] 2by2 scrambling algorithm maps the sequence "1,2,3,4,5,6,7,8" to "2,1,4,3,6,5,7,8".

between secondary dominants and diatonic associates. Stimuli of Experiment 5, which were same as the Experiment 4, were played with the piano timbre. Participants of Experiment 5 were non-musicians. Global chord priming was not observed in Experiment 5, though the percentage of error data suggested it. In Experiment 6, the perception of the relation between secondary dominant and diatonic associates was investigated by scrambling sequences with 4by4 algorithm[3]. Stimuli of Experiment 6 were same as the stimuli of Experiment 5, except they were scrambled with the 4by4 algorithm. Participants of Experiment 6 were non-musicians. In Experiment 6, global chord priming was observed. However, disruption of the global chord priming due to the violation of the sequential relation between the secondary dominant chords and their diatonic associates was not observed.

Predictions of MUSACT (Bharucha, 1987) and Tonal Pitch Space theory (Lerdahl, 2001) were not in a complete agreement with the results of Experiments 4, 5 and 6. According to Tonal Pitch Space theory (Lerdahl, 2001), global chord priming should have only been observed with the piano timbre and not with the Shepard (1964) tones. This was contrary to the results; global chord priming was observed from non-musicians with Shepard tones in Experiment 4, and with the piano timbre in Experiment 6; and it was suggested in Experiment 5, which utilized the piano timbre too. According to MUSACT (Bharucha, 1987), global chord priming should have been observed both with Shepard (1964) tones and the piano timbre; but the expectation towards tonic and subdominant chords should have been decreased due to the secondary dominant chords. Global chord priming was observed from non-musicians in Experiments 4 and 6 definitely, and in Experiment 5 suggestively; however disruption of expectation towards tonic and subdominant targets due to the secondary dominant chords was not observed from non-musicians in any of the experiments.

---

[3] 4by4 scrambling algorithm maps the sequence "1,2,3,4,5,6,7,8" to "4,1,5,2,6,3,7,8"

As a result, the perception of sequential relations between the secondary dominant chords and their diatonic associates was not observed in non-musicians in Experiments 4, 5 and 6. Turkish non-musicians perceived the difference between the perceptual status of tonic and subdominant targets. However, this perception was diminished when non-diatonic chords were present in chord sequences, and if the sequences were played with the Shepard tones (Shepard, 1964). This may be a result of the fact that it is hard for Turkish non-musicians to perceive music without clear melodies, since the \Turkish musical culture is mainly melodic. This hypothesis needs further investigation.

It may be interesting to investigate the perception of musicians for the secondary dominant chords; and compare it with non-musicians. It may be interesting to investigate the relation between the secondary dominants and their diatonic associates with recognition memory tasks (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha & Castelleno, 1982) and similarity ratings (Bharucha & Krumhansl, 1983; Krumhansl, 1990; Krumhansl, Bharucha & Castelleno, 1982; Krumhansl, Bharucha & Kessler, 1982). Lastly, it may be better to conduct future studies about the secondary dominant chords in Western countries, for the reasons given above.

In addition to behavioral experiments, self-organizing models of chord schema were studied in this thesis. MUSACT (Bharucha, 1987) is a hand-wired model of tonal schema; and it simulates many important results of chord priming. Tillmann et. al. (2000) showed that MUSACT (Bharucha, 1987) may be developed by training self-organizing maps (Kohonen, 1982; Rumelhart & Zipser, 1985) with chord sequences that are typical to Western tonal music. In order to develop network structures identical to MUSACT (Bharucha, 1987), Tillmann et. al. (2000) pruned and scaled weights of the self-organized networks. However, Tillmann et. al. (2000) did not support pruning and scaling steps with strong empirical results, which made them questionable. In this thesis, self-organizing models of chord schema that was not based on pruning and scaling of weights were investigated. In the first set of simulations, self-organizing models (Kohonen, 2001) were trained with chord categories. Chord categories were composed of pitch categories. Effects of different pitch category representations on the self-organization of tonal schema were investigated. Pitch category representations were based on descriptions of relations between pitch categories. Five types of pitch category representations were studied: "orthogonal", "chroma", "circle of fifths", "chroma + circle of fifths", "major3rd + circle of fifths". In "orthogonal" representation similarity between pitch categories was equal. In "chroma" representation similarity between pitch categories approximated the chroma circle. In "circle of fifths" representation similarity between pitch categories approximated the circle of fifths. In "chroma + circle of fifths" representation similarity between pitch categories approximated chroma circle and circle of fifths at the same time. In "major3rd + circle of fifths" representation similarity between pitch categories approximated distances on the lattice of pitches[4]. Self-organizing maps (Kohonen, 2001) were trained with 12 major and 12 minor chord categories that were composed of pitch categories.

After training, important studies on the perception of chords were simulated, namely Bhaucha and Krumhansl (1983), Krumhansl (1990), Bharucha & Stoeckig, (1986), Tekman and Bharucha (1998), and Bigand and Pineau, (1997). Bhaucha and Krumhansl (1983) quantified the perceptual relation between chords after musical context; Krumhansl (1990) observed stability of chords after musical context; Bharucha and Stoeckig, (1986) observed the chord priming effect; Tekman and Bharucha (1998) showed that chord priming is due to the distance on the circle of fifths; and Bigand and Pineau, (1997) observed the global chord priming effect. Results showed that networks trained with "circle of fifths" and "chroma + circle of fifths" representations were the most successful tonal schema that simulated all of the studies; and they were significantly correlated with MUSACT (Bharucha, 1983).

[4] For a detailed description of these representations see Section 5.3.1.

As a result, self-organizing chord schemata were developed without pruning and scaling of weights.

In the second set of simulations, self-organizing models (Kohonen, 2001) were trained with simple chord sequences. Sequences were divided into windows. Short term memory model of pitch, which can differential past events from present events, was applied to each window. The short term memory model represents temporal relations between musical events. In the end, there were 720 short term memory images; and self-organizing networks were trained with these images. Effects of pitch category representations on the self-organization of networks were investigated in these simulations as well. Simulation results were similar to the first set. Networks trained with "circle of fifths" and "chroma + circle of fifths" representations were again successfully simulated all of the studies; and they were significantly correlated with MUSACT (Bharucha, 1983).

Behavioral experiments of this thesis were not simulated with the self-organizing models because training sets did not include the Neapolitan and secondary dominant chords. It may be interesting to train the networks with a big corpus of Western tonal music; and observe their performance for the Neapolitan and secondary dominant chords. Unfortunately such a study requires knowledge and resource for high performance computing, which are beyond the scope of this thesis work. Self-organization algorithms depend on computationally expensive operations, such as vector product, vector length, vector distance. As the number of units on the self-organizing map and the amount of data in the training set increases the number of operations necessary to converge the map into the final step increases. There are over 200 Bach chorales, which include hundreds of thousands of notes. Therefore, training self-organizing maps with the complete set of Bach chorales is the beyond the capacities of desktop computers, and conventional programming.

REFERENCES

Acker B.E., Pastore, R.E. (1996). Perceptual integrality of major chord components. *Perception and Psychophysics, 58*, 748-761.

Acker B.E., Pastore, R.E., Michael, D.H. (1995). Within-category discrimination of musical chords: Perceptual magnet or anchor? *Perception and Psychophysics, 57*, 863-874.

Aleksander I., & Morton H. (1990). *An introduction to neural computing*. London: Chapman & Hall.

Atalay, N.B. (2002). *Chord priming beyond association*. Unpublished masters thesis, Middle East Technical University: Ankara, Turkey.

Beer, R.D. (1995). Computational and dynamic languages for autonomous agents. In Robert F. Port & Timothy van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition*. Cambridge: A Bradford book the MIT Press.

Bernstein, L. (1976). *The unanswered question*. Cambridge, Mass.: Harvard University Press.

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

Bharucha, J.J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5*, 1-30.

Bharucha, J.J. (1992). Tonality and learnability. In M.R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication* (213-223). Washington, DC: American Psychological Association.

Bharucha, J. J. (1996). Melodic anchoring. *Music Perception, 13*, 383–400.

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

Bharucha, J.J., & Mencl, W.E. (1996). Two issues in auditory cognition: Self-organization of categories and pitch-invariant pattern recognition. *Psychological Science, 7*, 142-149.

Bharucha, J.J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 403-410.

Bharucha, J.J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra. *Perception & Psychophysics, 41*, 519-524.

Bharucha, J.J., & Todd, P. (1989). Modeling the perception of tonal structure with neural nets. *Computer Music Journal, 13*, 44-53.

Bharucha, J.J., Curtis, M., & Paroo, K. (2006). Varieties of musical experience. *Cognition, 100*, 131-172.

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception and & Psychophysics, 59*, 1098-1107.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 184-197.

Bigand, E., Poulin, B., Tillmann, B., & D'Adamo, D. (2003). Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 159–171.

Bigand, E., Tillmann, B., Poulin, B., & D'Adamo, D. A., & Madurell, F. (2001) The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81*, B11-B20.

Bigand, E., Tillmann, B., Poulin-Charronnat, B., & Manderlier, D. (2005) Repetition priming: Is music special?. *Quarterly Journal of Experimental Psychology, 85A*, 1347-1375.

Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge: MIT Press.

Budge, H. (1943). *A study of chord frequencies (Contributions to Education No.882)*. New York: Columbia University Teacher Collage.

Burns, E.M., & Campbell, S.L. (1994). Frequency and frequency-ratio resolution by possessors of absolute and relative pitch: examples of categorical perception. *Journal of Acoustical Society of America , 96*, 2704-2719.

Burns, E.M., & Ward, W.D. (1978). Categorical perception - phenomenon or epiphenomenon: evidence from experiments in the perception of melodic musical intervals. *Journal of Acoustical Society of America , 63*, 456-468.

Carpenter, G.A. & Grossberg, S. (1990). ART 3: Hierarchical search using chemical transmitters in self- organizing pattern recognition architectures. Neural Networks, 3, 129-152.

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

Cedolin, L. & Delgutte, B. (2005). Pitch of complex tones: Rate-place and interspike-interval representations in the auditory nerve. *Journal of Neurophysiology, 94*, 347–362.

Chew, E. (2001). Modeling Tonality: Applications to Music Cognition. In J.D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society, CogSci2001* (206-211). Mahwah, NJ/London: Lawrence Erlbaum Assoc. Pub.

Clarke, E.F. (1999). Rhythm and timing in music. In: D. Deutsch (Ed.), *The Psychology of Music (2nd Edition)*. San Diego: Academic Press.

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

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

Deliège, I. (1987). Grouping conditions in listening to music: An Approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception 4*, 325-360.

Deliege, I. & Sloboda, J. (1997). *Perception and cognition of music*. Mahwah NJ: Lawrence Erlbaum Associates.

Deliège, I., Melen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real time listening to a piece of music. *Music Perception, 14*, 117-160.

Deutsch, D. (1982). The influence of melodic context on pitch recognition judgment. *Perception & Psychophysics, 31*, 407-410.

Deutsch, D. (1991). The tritone paradox - an influence of language on music perception. *Music Perception, 8*, 335-347.

Deutsch, D. (Ed.) (1999). *The psychology of music (2nd Ed.)*. San Diego: Academic Press.

Deutsch, D. (1999). The processing of pitch combinations. In D. Deutsch (Ed.) *The Psychology of Music (2nd Ed.)*. San Diego: Academic Press.

Deutsch, D. (2006). The enigma of absolute pitch. *Acoustics Today, 2*, 11-19.

Deutsch, D. & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*, 503-522.

Dowling, W. and D. Harwood. (1986). *Music cognition*. New York: Academic Press.

Eerola, T. & Toiviainen, P. (2004). MIR in Matlab: The Midi Toolbox. In *Proceedings of 5th International Conference on Music Information Retrieval (ISMIR 2004)* (22-27). Barcelona: Universitat Pompeu Fabra.

Elman, J. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.

Euler, L. (1739/1968). *Tentamen novae theoria musicae*. St. Petersburg. New York: Broude.

Fodor, J.A. & Pylyshyn, Z.W. (1988): Connectionism and cognitive architecture. *Cognition, 28*, 3-71.

Fraisse, P. (1982). Rhythm and tempo. In: D. Deutsch (Ed.), *The Psychology of Music*. New York: Academic Press, 149-180.

Fujioka T., Ross B., Okamoto H., Takeshima Y., Kakigi R., & Pantev C. (2003) Tonotopic representation of missing fundamental complex sounds in the human auditory cortex. *European Journal of Neuroscience, 18*, 432-40.

Gjerdingen, R.O. (1992) Learning syntactically significant temporal patterns of chords: A masking field embedded in an ART3 architecture, *Neural Networks, 5*, 551-564.

Greentree, M. (2005). *The Bach Chorales*. Retrieved February 24, 2005, from http://www.jsbchorales.net/

Harwood, D.L. (1976). Universals in music: A perspective from cognitive psychology, *Ethnomusicology, 3*, 521-533.

Helmholtz, H. (1954). *On the Sensations of Tone (second English ed.)*. Dover.

Hertz, J., Krogh, A., Palmer, R. G. (1991*). Introduction to the Theory of Neural Computation (Santa Fe Institute Studies in the Sciences of Complexity. Vol.1).* Boston: Addison-Wesley Longman Publishing Co., Inc.

Horton, T. (2001). The compositionality of tonal structures: A generative approach to the notion of musical meaning'*, Musicæ Scientiæ 5*, 131–156 .

Horton, T. (2002). Some data that falsify spreading-activation accounts of global context effects in tonal music. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition*.

Janata, P., Birk, J.L., Tillmann, B., Bharucha, J.J. (2003). Online Detection of Tonal Pop-Out in Modulating Contexts. *Music Perception, 20*, 283-305.

Justus, T.C. & Bharucha, J.J. (2001). Modularity in musical processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1000-1011.

Justus, T.C. & Bharucha, J.J. (2002). Music perception and cognition. In S. Yantis & H. Passler (Eds.), *Steven's Hanbook of Ecperimental Psychology, Volume 1: Sensation and Perception (Third Edition)*. New York: Wiley, 2002.

Kessler, E. J., Hansen, C., Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and the West. *Music Perception, 2*, 131-165.

Knudsen, E.I., du Lac, S. and Esterly, S.D. (1987). Computational maps in the brain. *Annual Review of Neuroscience, 10*, 41-65.

Koelsch, S., Gunter, T., Friederici, & A.D., Schröger E. (2000). Brain indices of music processing: 'Non-musicians' are musical. *Journal of Cognitive Neuroscience, 12*, 520-541.

Koelsch, S., Friederici, & Gunter, T. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysiology, 39*, 38-48.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*, 59-69.

Kohonen, T. (2001). *Self-organizing maps (3th Ed.)*. Springer: Berlin.

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

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Krumhansl, C.L. (1991). Music psychology: Tpnal structures in perception and memory. *Annual Review of Psychology, 42*, 277-303.

Krumhansl, C.L. (1992). Internal representations for music perception and performance. In M.R. Jones and S. Holleran (Eds.), *Cognitive bases of musical communication*. Washington, DC: American Psychological Association.

Krumhansl C.L. (1996). A perceptual analysis of Mozart's piano sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception, 13*, 401-432.

Krumhansl, C.L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159-179.

Krumhansl, C.L. (2005). The cognition of tonality – as we know it today. *Journal of New Music Research, 33*, 253-268.

Krumhansl, C.L. & Keil, F.C. (1982). Acquisition of the hierarchy tonal functions in music. *Memory and Cognition, 11*, 325-334.

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

Krumhansl, C.L. & Schmuckler, M.A. (1986). The Petrouska chord. *Music Perception, 4*, 153-184.

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

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

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

Krumhansl, C.L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception, 17*, 151-196.

Krumhansl, C.L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition, 75*, 1-46.

Leman, M. (1997). *Music, gestalt and computing*. Berlin: Springer.

Leman, M. & Carreras, (1997). Schema and Gestalt: Testing the hypothesis of psychoneural isomorphism by computer simulation. In M. Leman (Ed.), *Music, gestalt and computing*. Berlin: Springer.

Lerdahl L. (2001). *Tonal Pitch Space*. New York: Oxford University Press.

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

Lerdahl, F., & Krumhansl, C.L. (2001). *Modeling tonal tension and attraction in chromatic contexts: Chopin*. (Presentation given at Kingston, Ontario: Society for Music Perception and Cognition.)

Lerdahl, F., Krumhansl, C.L. Hanon, E., & Fineberg, J. (2000). *Modeling tonal tension and attraction*. (Presentation given at Toronto: Musical Intersections.)

Longuet-Higgins, H.C. & Steedman, M.J. (1987). On interpreting Bach. In H.C. Longuet-Higgins (Ed.), *Mental processes: Studies in Cognitive Science* (82–104). British Psychological Society/MIT Press, London, England and Cambridge, Mass.

Luria, A., Tsvetkova, L., & Futer, J. (1965). Aphasia in a composer. *Journal of Neurological Sciences, 2*, 288-292.

Maess, B., Koelsch, S., Gunter, T.C. & Friederici, A.D. (2001). Musical Syntax is processed in Broca's area: an MEG study. *Nature Neuroscience, 4*, 540-545.

McCulloch, W. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics, 7*, 115 - 133.

McDermott, J. & Hauser, M. (2005) The origins of music: Innateness, development, and evolution. *Music Perception, 23*, 29-59.

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

Minsky, M. (1967). *Finite and infinite machines (Automatic computation)*. Englewood Cliffs, N.J.: Prentice-Hall.

Minsky, M. & Papert, S. (1969). *Perceptrons*. Cambridge: MIT Press.

Moore, B.C.J. (Ed.) (1995). *Hearing*. New York: Academic Press.

Moore, B.C.J. (2003). *An Introduction to the Psychology of Hearing (5th ed.)*. San Diego: Academic Press.

Nam, U. (1998). Pitch distributions in Korean court music: Evidence consistent with tonal hierarchies. *Music Perception, 16*, 243-247.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.

Oram, N., & Cuddy, L.L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research, 57*, 103–118.

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

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

Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G., Wittkowski, W. (1988). Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephaphy in Clinical Neurophysiology, 69*, 160-170.

Pantev, C., Hoke, M., Lütkenhöner, B., Lehnertz, K. (1989). Tonotopic organization of the auditory cortex: pitch versus frequency representation. *Science, 246*, 486-488.

Parncutt, R. (1988). Revision of Terhardt's psychoacoustical model of the root(s) of a musical chord. *Music Perception, 6*, 65-94.

Parncutt, R. (1994). Template-matching models of musical pitch and rhythm perception. *Journal of New Music Research, 23*, 145-168.

Patel, A.D. (1998). Syntactic processing in language and music: Different cognitive operations, similar neural resources? *Music Perception, 16*, 27-42.

Patel A.D. (2003). Language, music, syntax and the brain. *Nature Neuroscience 6*, 674-681.

Patel, A.D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Annals of the New York Academy of Sciences, 1060*, 59-70.

Patel, A.D., Gibson, E., Ratner, J., Besson, M. & Holcomb, P.J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*, 717-733.

Peretz I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology, 10*, 21–56.

Pickett, J.P. (Ed.) (2003). *The American Heritage® Dictionary of the English Language (Fourth Edition)*. Boston: Houghton Mifflin Company.

Pinker, S. (1997). *How the mind works (1st ed.)*. New York: Norton.

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

Plomp R., & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America,. 38*, 548-560.

Randel. D. M. (Ed.) (2003). *The Harvard music dictionary*. Cambridge, MA: Belknap Press of Harvard University Press.

Rosch, E. (1973). Natural categories. *Cognitive Psychology 4*, 328-350.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192-233.

Rumelhart, D.E. & McClelland, J.L. (Eds.) (1986). *Parallel distributed processing, Vol. I: Foundations*. Cambridge, MA: MIT Press.

Rumelhart, D.E., Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science, 9*, 75-112.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). Learning internal representations by error propagation. In . D.E. Rumelhart & J.L. McClelland (Eds) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations* (318-362) Cambridge, MA: MIT Press.

Schmuckler, M.A. (1989). Expectation in music: Investigations of melodic and harmonic processes. *Music Perception, 7*, 109-150.

Schmuckler, M.A., & Boltz, M. G. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception & Psychophysics, 56*, 313-325.

Schwartz, D.A., Howe, C.Q., & Purves, D. (2003). The statistical structure of human speech sounds predicts musical universals. *The Journal of Neuroscience, 23*, 7160-7168

Searle, S.R. (1982). *Matrix algebra useful for statistics*. New York: John Wiley & Sons.

Sheh A., & Ellis D. (2003). *Chord segmentation and recognition using EM-Trained Hidden Markov Models*. (4th International Symposium on Music Information Retrieval.)

Shepard, R.N. (1964). Circularity in judgment of relative pitch. *Journal of the Acoustical Society of America, 36*, 2346–2353.

Shepard, R.N. (1982). Geometric approximations to musical pitch. *Psychological Review, 89*, 305-333.

Simpson, G.B., Peterson, R.R., Casteel, M.A., & Burgges, C. (1989). Lexical and sentence context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 88-97.

Sloboda J.A. (1985). *The musical mind: the cognitive psychology of music*. Oxford: Clarendon Press.

Sloboda, J. (2004). *Exploring the musical mind*. Oxford: Oxford University Press.

Speer, J.R. & Meeks, P.U. (1985). School children's perception of pitch in music. *Psychomusicology, 5*, 49-56.

Tekman, H. G., & Bharucha, J. J. (1992). Time course of chord priming. *Perception & Psychophysics, 51*, 33-39.

Tekman, H. G., & Bharucha, J. J. (1998). Implicit knowledge versus psychoacoustic similarity. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 252-260.

Tekman, H. G., & Hortaçsu, N. (2002). Aspects of stylistic knowledge: What are different styles like and why do we listen to them? *Psychology of Music, 30*, 28-47.

Temperley, D. (2000). Meter and grouping in African music: A view from music theory. *Ethnomusicology, 44*, 65-96.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of Acoustical Society of America, 55*, 1061-1069.

Terhardt, E., Ward, W.D. (1982). Recognition of musical key: Exploratory study. *Journal of the Acoustical Society of America, 72*, 26-33.

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

Terhardt, E., Stoll, G., & Seewan, M. (1982b). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America, 71*, 679-687.

Tillmann, B., & Bharucha, J.J. (2002). Effect of harmonic relatedness on the detection of temporal asynchronies. *Perception and Psychophysics, 64*, 640-649.

Tillmann, B. & Bigand, E. (1998). Influence of global structure on musical target detection and recognition. *International Journal of Psychology, 33*, 107-122.

Tillmann, B., & Bigand, E. (2001). Global context effect in normal and scrambled musical sequences. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1185-1196.

Tillmann, B., Bharucha, J.J., & Bigand, E. (2000). Implicit learning of tonality. *Psychological Review, 4*, 885-913.

Tillmann, B., Bigand, E., & Madurell, F. (1998). Influence of global and local structures on the solution of musical puzzles. *Psychological Research, 61*, 157-174.

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of local and global context on harmonic expectancy. *Music Perception, 16*, 99-118.

Tillmann, B., Bigand, E, Escoffier, N., & Lalitte, P. (2006) Influence of harmonic context on musical timbre processing. *European Journal of Experimental Psychology, 18*, 343–358.

Tillmann B., Janata P., Birk J. & Bharucha J.J. (2003). The costs and benefits of tonal centers for chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 470-482.

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A.D. & vonCramon, D.Y. (2006) Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *NeuroImage, 31*, 1771-1782.

Todd P. & Loy G. (Eds.) (1991). *Music and Connectionism.* Cambridge, MA: MIT Press.

Toiviainen, P. & Krumhansl, C.L. (2003). Measuring and modeling real-time responses to music: the dynamics of tonality induction. *Perception, 32*, 741-766.

Trainor, L.J. & Trehub, S.E. (1992). A comparision of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 394-402.

Trainor, L.J. & Trehub, S.E. (1993). Musical context effects in infants and adults: Key distance. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 615-626.

Trainor, L.J. & Trehub, S.E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception & Psychophysics, 56*, 125-132.

Trehub, S.E. (2000). Human processing predispositions and musical universals. In N.L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins of Music*. Cambridge, MA: MIT Press.

Trehub, S.E., Cohen, A.J., Thorpe, L.A., & Morrongiello, B.A. (1986). Development of the perception of musical relations: Semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 295-301.

Van Immerseel L. & Martens, J. P. (1992). Pitch and voiced/unvoiced determination with an auditory model, *Journal of the Acoustical Society of America, 91*, 3511-3526.

Vaissière, J. (2005). Perception of intonation. In David B. Pisoni & Robert E. Remez (Eds.) *The handbook of speech perception*. Malden: Blackwell Publishing.

Vesanto, J., Himberg, J., Alhoniemi, E., & Parhankangas, J. (2000). *SOM Toolbox for Matlab 5*. Technical Report A57, Helsinki University of Technology.

Ward, W.D., & Burns, E.M. (1982). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press.

Yoshioka, T., Kitahara, T., Komatani, K., Ogata, T., Okuno, H.G. (2004). *Automatic chord transcription with concurrent recognition of chord symbols and boundaries*. (Presentation given at Fabra, Spain: 5th International Symposium on Music Information Retrieval.)

Zatorre, R.J. (1988). Pitch perception of complex tones and human temporal lobe function. *Journal of Acoustical Society of America , 84*, 566-572.

Zatorre, R.J., & Beckett, C. (1989). Multiple coding strategies in the retention of musical tones by possessors of absolute pitch. *Memory & Cognition, 17*, 582-589.

APPENDICES

APPENDIX A: MUSACT CODE

Appendix A presents code that implements MUSACT (Bharuca, 1987) (from Bigand et.al. 1999).

```
%MUSACT Matlab(1997) Implementation of Bharucha's (1987) Model
%in the Appendix of Bigand, Madurell, Tillmann and Pineau (1999)
%
%tones a ais b c cis d dis e f fis g gis
%chords Fis Cis Gis Dis Ais F C G D A E B dis ais f c g d a e b fis cis gis
%keys Fis Cis Gis Dis Ais F C G D A E B

ntones=12;nchords=24;nkeys=12;

W=[0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 1 1 0 0 1 0 0;1 0 0 1 1 0 0 0  0 0 0 0  1 1 0 0 1 0 0 0   0
0 0 0;0 0 0 0     0 0 0 1  0 0 1 1  0 0 0 0  0 0 0 1  1 0 0 1;0 0 1 0   0 1 1 0  0 0 0 0  0 0 1 1  0
0 1 0    0 0 0 0;1 1 0 0   0 0 0 0  0 1 0 0  0 1 0 0  0 0 0 0  0 1 1 0;0 0 0 0   1 0 0 1  1 0 0 0  0
0 0 0    1 1 0 0  1 0 0 0;0 0 1 1   0 0 0 0  0 0 0 1  1 0 0 1  0 0 0 0  0 0 0 1;0 0 0 0   0 0 1 0  0
1 1 0    0 0 0 0  0 0 1 1  0 0 1 0;0 1 0 0 1 1 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0;1 0 0 0       0
0 0 0    1 0 0 1  1 0 0 0  0 0 0 0  1 1 0 0;0 0 0 1   0 0 1 1  0 0 0 0  0 0 0 1  1 0 0 1  0 0 0 0;0 1
1 0      0 0 0 0  0 0 1 0  0 0 1 0  0 0 0 0  0 0 1 1];

Z=[1 1 0 0 0 0 0 0 0 0 0 1; 1 1 1 0 0 0 0 0 0 0 0 0;0 1 1 1 0 0 0 0  0 0 0 0;0 0 1 1   1 0 0 0  0
0 0 0;0 0 0 1     1 1 0 0  0 0 0 0;0 0 0 0   1 1 1 0  0 0 0 0;0 0 0 0   0 1 1 1  0 0 0 0; 0 0 0 0  0
0 1 1    1 0 0 0;0 0 0 0   0 0 0 1  1 1 0 0;0 0 0 0   0 0 0 0  1 1 1 0;0 0 0 0   0 0 0 0  0 1 1 1;1 0
0 0      0 0 0 0  0 0 1 1;1 1 0 0   0 0 0 0  0 0 0 1;1 1 1 0   0 0 0 0  0 0 0 0;0 1 1 1   0 0 0 0  0
0 0 0;0 0 1 1     1 0 0 0  0 0 0 0;0 0 0 1   1 1 0 0  0 0 0 0;0 0 0 0   1 1 1 0  0 0 0 0;0 0 0 0 0 1
1 1      0 0 0 0;0 0 0 0   0 0 1 1  1 0 0 0;0 0 0 0   0 0 0 1  1 1 0 0;0 0 0 0   0 0 0 0  1 1 1 0;0 0
0 0      0 0 0 0  0 1 1 1;1 0 0 0   0 0 0 0  0 0 1 1];


%weights w1=tones to chords; w2= major chords to keys; w3= minor chords to keys
```

```
w1=.0122;w2=.244;w3=.22;
W=W*w1; Z(1:(nchords/2),:)=Z(1:(nchords/2),:)*w2;
Z((nchords/2+1):nchords,:)=Z((nchords/2+1):nchords,:)*w3;

%initializations(t=tones; ch= chords; k= keys)
t_p=zeros(1, ntones);ch_p=zeros(1, nchords);k_p=zeros(1, nkeys);
t=zeros(1, ntones);ch=zeros(1, nchords);k=zeros(1, nkeys);
delta_t=zeros(1, ntones);delta_ch=zeros(1, nchords);delta_k=zeros(1, nkeys);

%input pattern(mat_activ) example of C major chord
mat_activ=[0 0 0 1 0 0 0 1 0 0 1 0];
[nc,nn]=size(mat_activ);
ti=1;decay=0.4;
counter=0;hdelta=1;
for kk=1:nc;
t=t_p+mat_activ(kk,:);
%reverberation phase
        while any ([delta_t';delta_ch';delta_k';hdelta]>.005)
                counter=counter+1;hdelta=0;
                %compute phasic activation
                delta_t=abs(t-t_p);delta_ch=abs(ch-ch_p);delta_k=abs(k-k_p);
                %actual activations becomes previous ones
                t_p=t;ch_p=ch;k_p=k;
                %compute activation
                t=t_p+delta_ch*W';
                ch=ch_p+delta_t*W+delta_k*Z';
                k=k_p+delta_ch*Z;
        end
%decay
        hdelta=1;
        t_p=t_p*((1-decay)^ti);
        ch_p=ch_p*((1-decay)^ti);k_p=k_p*((1-decay)^ti);
        ch=ch*((1-decay)^ti);k=k*((1-decay)^ti);
end
```

APPENDIX B: STIMULI OF EXPERIMENT 1

Appendix B presents example stimuli for the conditions of Experiment 1 (see Section 4.1). In A dominant target (last chord) follows three-steps-away penultimate. In B dominant target (last chord) follows six-steps-away penultimate. In C subdominant target (last chord) follows three-steps-away penultimate. In D subdominant target (last chord) follows six-steps-away penultimate.

APPENDIX C: LIST OF CHORALES

Appendix C presents list of Bach chorales utilized in Experiments 4, 5 and 6. Numbers in parenthesis represent start and end measures of excerpts employed as stimulus. In order to meet experimental needs subdominant chord appended to the end of chorales that marked with star (*). Scores of chorales were obtained from Greentree (2005).

BWV 3.6 (5-6)
BWV 32.6 (9-10)
BWV 70.11 (3-6)
BWV 72.6 (1-2)
BWV 250 (11-13)
BWV 255 (1-2)
BWV 255 (6-8)
BWV 335 (1-2)
BWV 392 (9-11)
BWV 402 (12-14)*
BWV 404 (2-4)
BWV 410 (9-11)
BWV 401 (2-3)
BWV 411 (16-17)*
BWV 413 (6-8)*
BWV 414 (3-5)
BWV 414 (13-15)*
BWV 415 (2-4)
BWV 417 (8-10)*
BWV 422 (1-3)
BWV 426 (1-2)
BWV 426 (14-16)
BWV 427 (12-14)
BWV 428 (11-13)
BWV 428 (13-15)
BWV 429 (2-4)
BWV 429 (11-13)
BWV 430 (2-3)
BWV 430 (6-8)
BWV 430 (11-13)
BWV 431 (1-3)*
BWV 431 (6-8)*
BWV 432 (1-3)
BWV 432 (5-7)
BWV 433 (6-8)*
BWV 436 (4-6)
BWV 436 (19-20)*
BWV 438 (3-5)
BWV 438 (6-8)

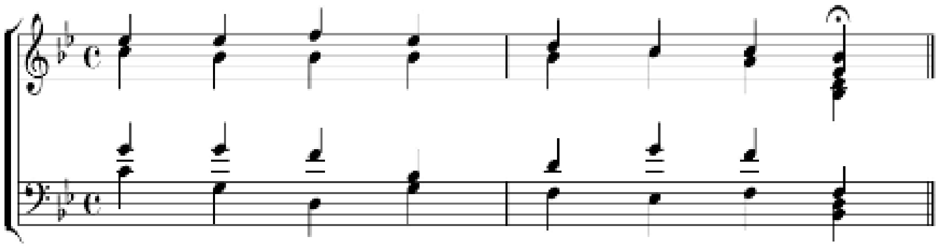
APPENDIX D: STIMULI OF EXPERIMENTS 4, 5, 6

Appendix D presents original (not-scrambled) chorale excerpts and their scores utilized in Experiments 4, 5 and 6. Chords shaded in gray are non-diatonic chords.

TONIC ENDING SEQUENCES DIATONIC

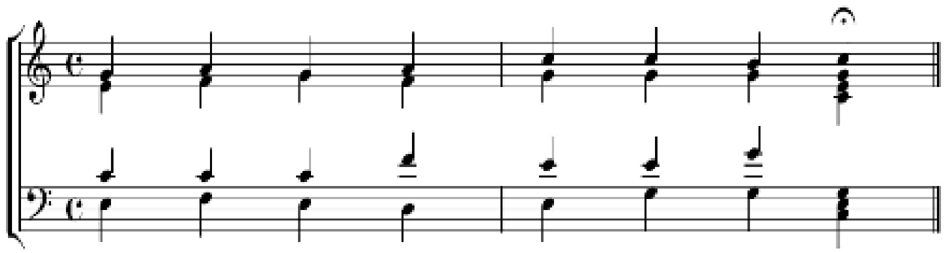
Cm	Eb	Bb	Eb	Bb	Cm	F	Bb
Gm	F	F	Bb	Bb	Cm	F	Bb
C	F	G	Dm	C	C	G	C
Am	C	F	C	C	C	G	C
D	G	A	G	D	A	A	D
D	A	G	D	A	D	A	D
Dm	Am	Bb	F	Dm	Gm	C	F
C	F	Bb	C	F	Gm	C	F
Em	C	C	G	G	D	D	G
D	Em	Am	F#dim	G	Am	D	G

Based on BWV 427 (measures 12-14)

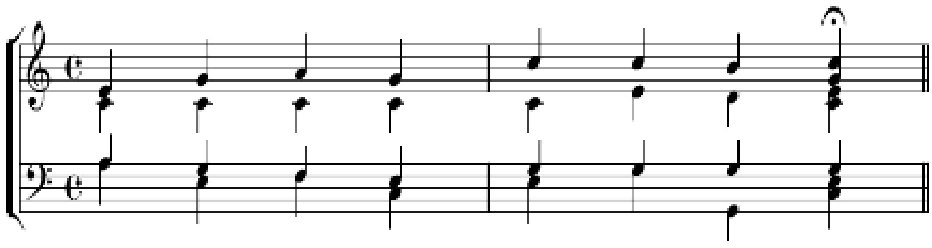


Based on BWV 429 (measures 2-4)

Based on BWV 422 (measures 1-3)



Based on Stimuli of D Bigand & Pineau (1997)



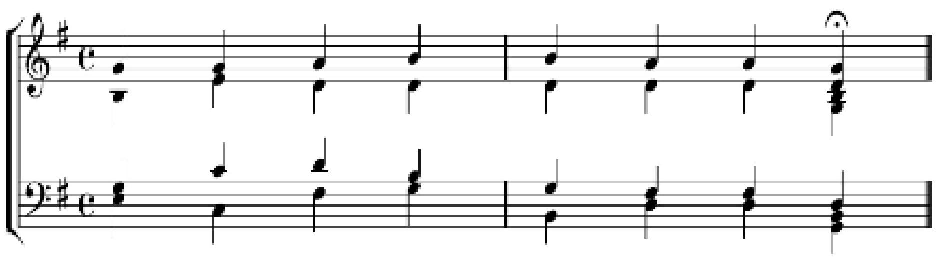
Based on BWV 431 (measures 1-3)

Based on Stimuli A of Bigand & Pineau (1997)

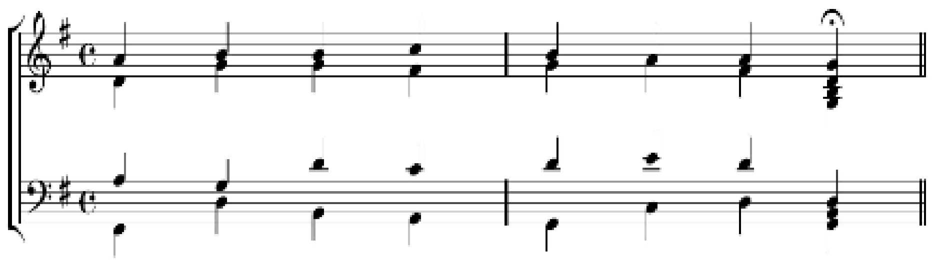
Based on BWV 438 (measures 6-8)

Based on Stimuli B of Bigand & Pineau (1997)

Based on BWV 414 (measures 3-5)



Based on BWV 428 (measures 13-15)

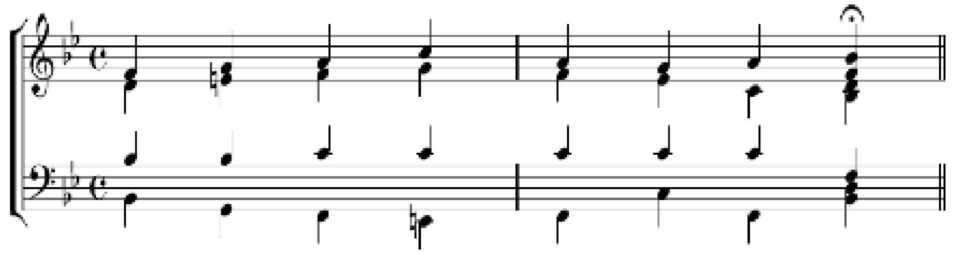


TONIC ENDING SEQUENCES NON-DIATONIC

G	Cm	D	Gm	Bb	Cm	F	Bb
Bb	Edim	F	C	F	Cm	F	Bb
C	G	D	G	C7	F	G	C
C	D	G#dim	Am	C	C	G	C
D	D	D	A#dim	Bm	G#dim	A	D
D	G	G	E	E	A	A	D
A	Dm	A	Dm	F	G	C	F
F7	Bb	C	F	Dm	G	C	F
G	B	Em	G#dim	Am	G	D	G
G	Em	A	D	G	A	D	G

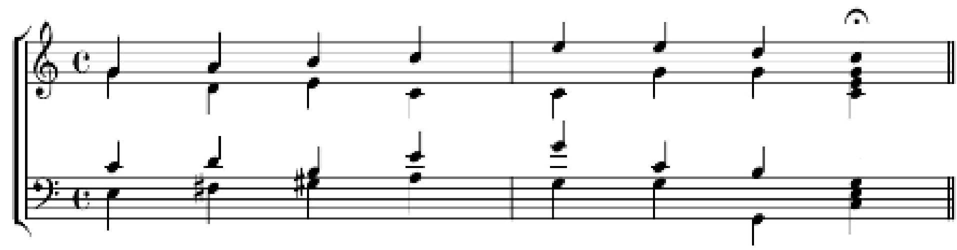
Based on BWV 410 (measures 9-11)

BWV 426 (measures 14-16)

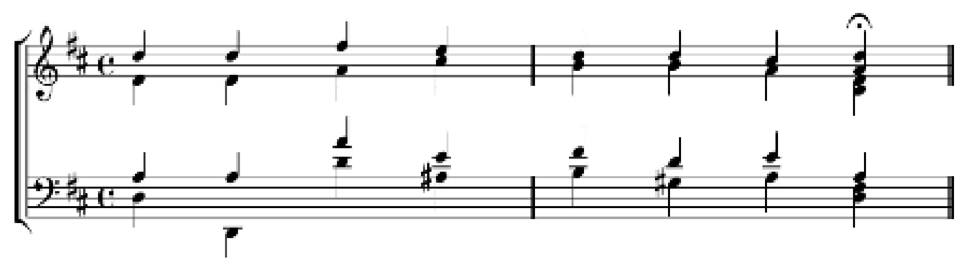


Based on BWV 255 (measures 1-2)

Based on BWV 255 (measures 6-8)



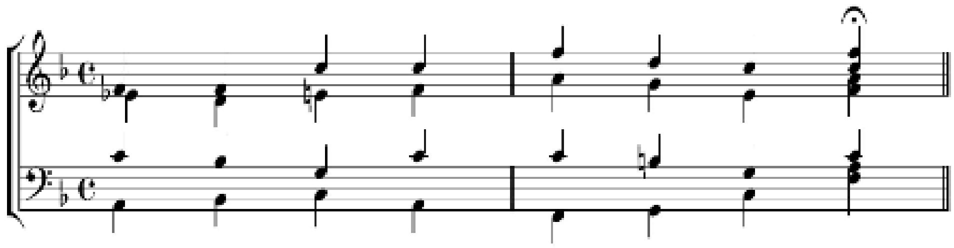
Based on BWV 415 (measures 2-4)



Based on BWV 367 (measures 7-9)

Based on BWV 404 (measures 2-4)

BWV 401 (measures 2-3)



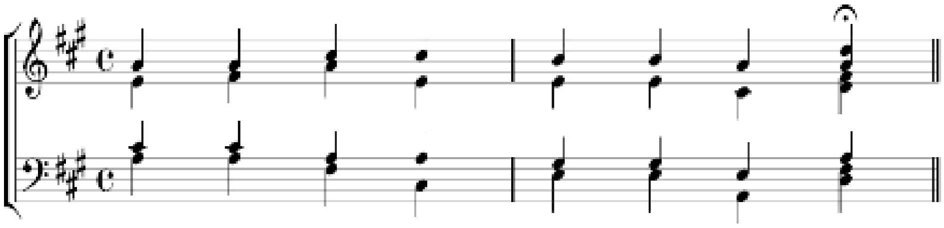
BWV 3.6 (measures 5-6)

BWV 428 (measures 11-13)

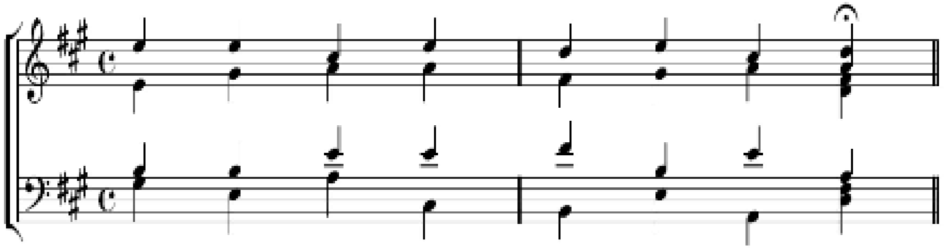
SUBDOMINANT ENDING SEQUENCES DIATONIC

A	F#m	F#m	A	E	E	A	D
E	E	A	A	Bm	E	A	D
C	Em	F	F	Dm	G	C	F
C	G	F	C	Dm	G	C	F
Bm	Bm	Em	Em	D	A	D	G
A	D	A	Bm	D	A	D	G
Bb	Bb	F	Bb	Gm	C	F	Bb
F	C	Bb	C	F	C	F	Bb
G	D	D	G	Am	D	G	C
Em	D	G	D	G	D	G	C

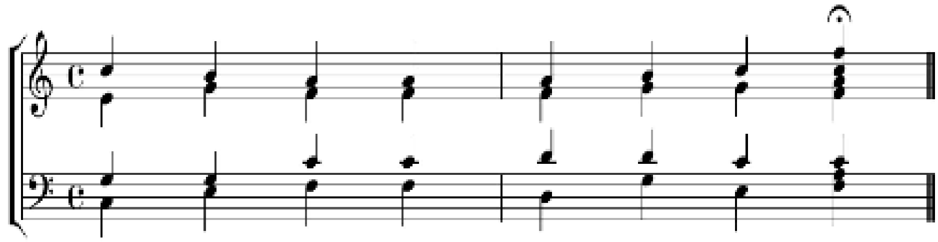
Based on BWV 414 (measures 13-15)



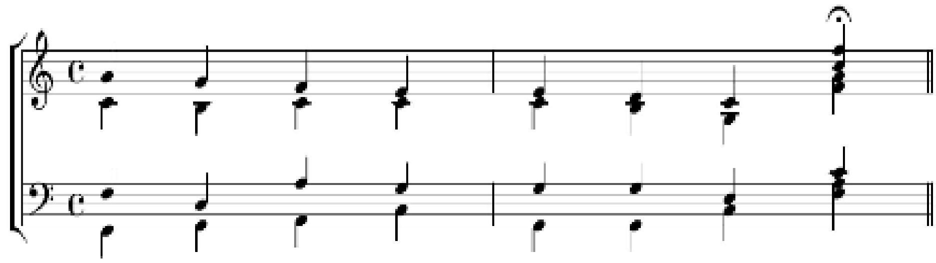
Based on BWV 430 (measures 7-8)



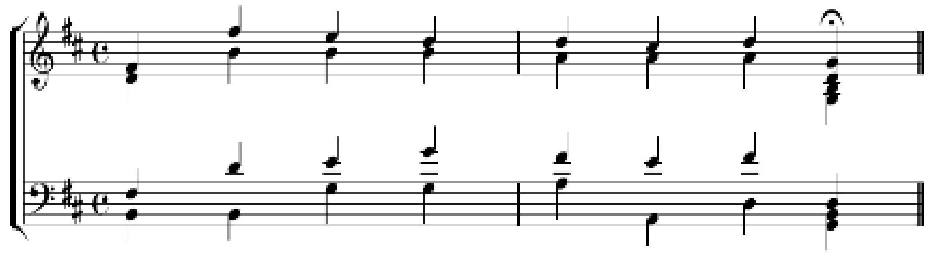
Based on BWV 426 (measures 1-2)



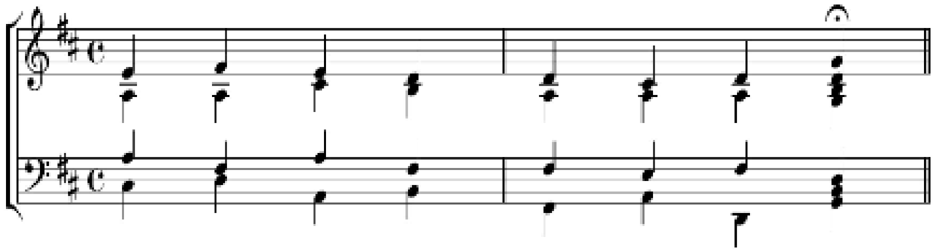
Based on BWV 436 (measures 5-6)



Based on BWV 417 (measures 8-10)



Based on BWV 433 (measures 6-8)



Based on BWV 413 (measures 6-8)

Based on BWV 438 (measures 3-5)

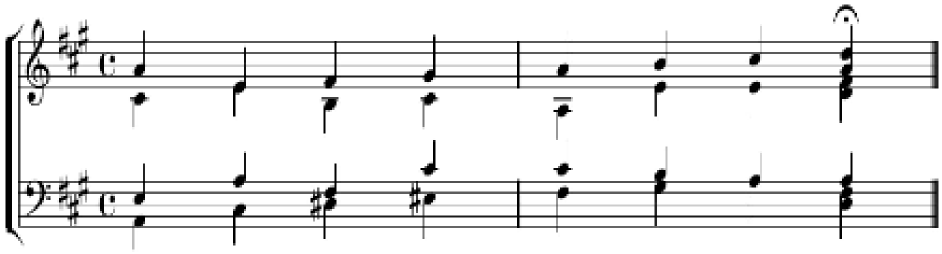
Based on BWV 411 (measures 16-17)

Based on BWV 432 (measures 1-3)

SUBDOMINANT ENDING SEQUENCES NON-DIATONIC

A	A	B	C#	F#m	E	A	D
F#m	B	E	A	B	E	A	D
D	G	C	D	G	C	C	F
C	C	D	G	C7	F	C	F
A	E	A	D	E	A	D	G
A	Em	F#	Bm	E	A	D	G
Bb	G	C	F	G	C	F	Bb
E	Am	Bb	F	G	C	F	Bb
A	D	G	G	A	D	G	C
A	D	G	Em	A	D	G	C

Based on BWV 430 (measures 2-3)



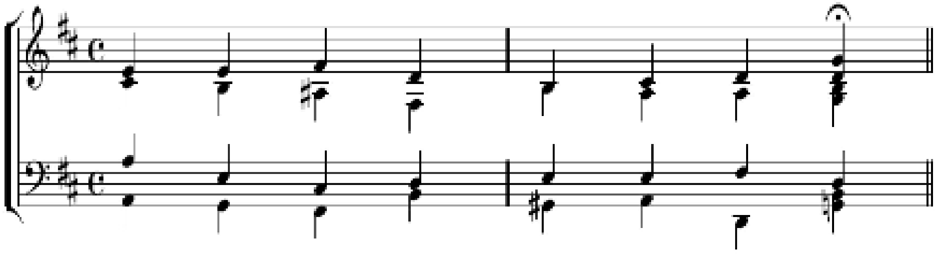
Based on BWV 429 (measures 11-13)

Based on BWV 70.11 (measures 3-6)

Based on BWV 402 (measures 12-14)

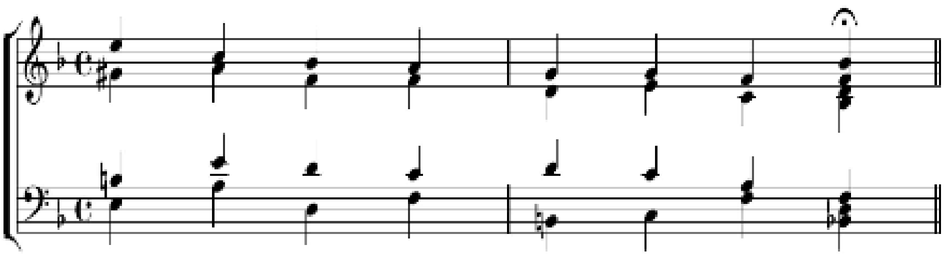
Based on BWV 377 (measures 3-5)

Based on BWV 362 (measures 2-4)

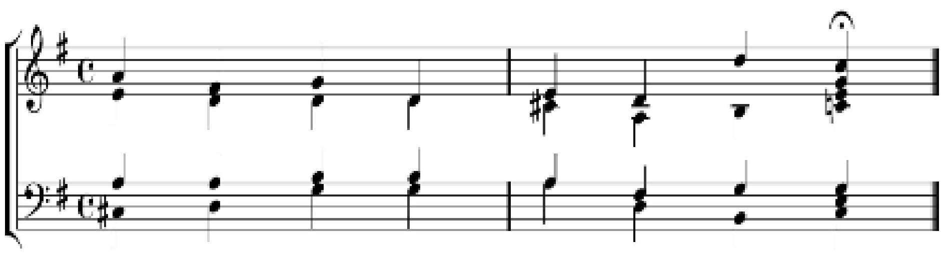


Based on BWV 431 (measures 6-8)

Based on BWV 436 (measures 19-20)



Based on BWV 250 (measures 11-13)



Based on BWV 432 (measures 5-7)

APPENDIX E: TEST SEQUENCES AND TARGET UNITS OF SIMULATIONS

Appendix E presents testing stimuli and target units of self-organizing chord schema simulations.

Study Reference	Label	Testing Stimuli	Target Unit(s)
Bharucha & Stoeckig (1986)	One-step-away	G	C
	Six-steps-away	F#	C
Tekman & Bharucha (1992)	Two-steps-away	D	C
	Four-steps-away	E	C
Bigand & Pineau (1996)	Tonic	D-G-G-D-Em-C-D	G
		G-D-D-G-C-Em-D	G
		D-D-Em-G-C-G-D	G
		C-F-Bb-C-F-Bb-C	F
	Subdominant	D-G-A-D-Em-A-D	G
		G-D-D-A-A-Em-D	G
		D-D-Em-G-A-A-D	G
		C-Dm-G-Am-Dm-G-C	F
Krumhansl (1990)	Major	C-F-Bdim-Em-Am-Dm-G-C	12 major and 12 minor chords
	Minor	Cm-Fm-Bb-Eb-Ab-Ddim-G-C	12 major and 12 minor chords
Bharucha & Krumhansl (1983)		F-G-C-C	Dm,Em,F,G,Am
		F-G-C-Dm	C,Em,F,G,Am
		F-G-C-Em	C,Dm,F,G,Am
		F-G-C-F	C,Dm,Em,G,Am
		F-G-C-G	C,Dm,Em,F,Am
		F-G-C-Am	C,Dm,Em,F,G

APPENDIX F: SIMPLE CHORD SEQUENCES

Appendix F presents training sequences utilized by simulations in Section 5.3.2 (Leman 1995, fig. 8.7). 24 training sequences of Section 5.3.2 were obtained by transposing A and B in 12 tonalities.

A

B