Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Social Sciences 138

A Dynamic View of Melodic Organization and Performance

Perception of Structure and Emotional Expression in Music

BY

ERIK LINDSTRÖM

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2004

Dissertation presented at Uppsala University to be publicly examined in Sal X, Universitetshuset, Uppsala, Monday, May 24, 2004 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Lindström, E. 2004. A Dynamic View of Melodic Organization and Performance. Perception of Structure and Emotional Expression in Music. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 138. 51 pp. Uppsala. ISBN 91-554-5953-6

Psychology of music has shown renewed interest in how music expresses emotion to listeners. However, there is an obvious lack of research on how interactions between musical factors such as harmony, rhythm, melodic contour, loudness, and articulation may affect perceived emotion. From previous literature on music analysis and music cognition there is evidence that tonality may be activated and affected by rhythm and melody. These ideas generated hypotheses regarding melodic organization and performance, for instance, (a) certain notes in a melodic structure have expressive potentials due to their place in the key/chord, (b) these notes could be activated by accents in the melodic structure and/or in live music performance. In Study I, a simple tune was systematically manipulated with regard to harmonic progression, rhythm and melodic contour. Listener ratings of the resulting versions showed that perceived structure (instability, complexity, tension) and emotion (sadness, anger, expressivity) could be partly interpreted as resulting from accent structures and stress on certain notes. In Study II, musicians were asked to perform some of the above-mentioned versions so as to express happiness, sadness, tenderness and anger. The performers used loudness and articulation to compensate for lack of adequate inherent expression in melodies. They also highlighted certain notes of relevance for the emotional meaning by means of stress in articulation, loudness and timing. In Study III, simple three-note sequences were manipulated with regard to melodic, metric and rhythmic accents as well as (computer-) performed accents (loudness, articulation and timing) on certain target notes. Listening tests showed that accent on a tense note enhanced perceived anger. A note essential for the identity of major mode affected perception of happiness, whereas a note essential for minor mode affected perception of sadness. The results in this thesis have implications for a dynamic view of melodic organization and performance.

Keywords: accents, emotion, expression, musical structure, tonal structure

Erik Lindström, Department of Psychology, Box 1225, Uppsala University, SE-75142 Uppsala, Sweden

ISSN 0282-7492
ISBN 91-554-5953-6
urn:nbn:se:uu:diva-4242 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4242)

To Anette, Linus and Elin

List of Papers

The present thesis is based on the following studies, which will be referred to in the text by their Roman numerals:

I Lindström, E. (2004). Impact of melodic organization on perceived structure and emotional expression in music. (In Progress)

II Lindström, E. (2004). Interplay between musical structure and performance in emotional expression of music. (Manuscript)

III Lindström, E. (2003). The contribution of immanent and performed accents to emotional expression in short tone sequences. *Journal of New Music Research, 32*, (3) 269-280.

Contents

Introduction

1.1 The perspective

The work of this thesis started one day when I was faced with two melodies that were very similar with regard to pitch, tonality, harmony, melodic contour and rhythm. They looked so similar but were still sounding so different and I asked myself, why?

Maybe this experience seems a bit trivial, but in fact it became an important impetus for this thesis. The general background is that researchers in psychology of music have shown renewed interest in how music expresses emotion to listeners (see 1.2 Emotion in music and 1.3 Earlier empirical studies). The research has basically focused on the effects of *single* musical factors such as tonality, harmony, melodic contour and rhythm ignoring possible effects of the interplay between different musical factors. Drawing on previous literature from music analysis and music cognition I realized that there were certain empirical findings and ideas, especially regarding the relationships between basic musical concepts such as tonality, melody and rhythm that were relevant for my research. These ideas led me to develop some hypotheses regarding melodic organization and performance (see 2. A dynamic view of …). In short, I am arguing that certain notes in a melodic structure have expressive potentials due to their place in the tonal hierarchy and their implied harmonic function (see 2.1 Music perception and expressive notes), and furthermore that these notes have to be activated by accents in the melodic structure (emphasis by rhythm, melody, meter) and/or in live music performance (emphasis achieved by the performer, see 2.2 The activation of expressive notes).

The thesis contains three papers (see 3. Empirical studies). In Study I, I used systematic manipulations of a simple tune to begin to explore the hypotheses hinted at above. Study II focused on the possible interplay between factors in the melodic structure and in the live performance of such structures. The implications of these studies and further hypotheses led to Study III in which I applied systematic variation of short tone sequences regarded both as melodic structures and as (quasi/computer-) performed. The results in this thesis have implications for the perception of emotional as well as structural characteristics of tonal music.

1.2 Emotion in music

Researchers in the psychology of music have conducted several studies on emotional expression in music. Recent reviews (Gabrielsson & Lindström, 2001; Gabrielsson & Juslin, 2003) cover about 75 empirical studies, the results of which indicate systematic relationships between various musical factors - tempo, mode, loudness, pitch, intervals, melody, harmony, tonality, rhythm, timbre, timing, articulation, amplitude envelopes - and emotions such as happiness, sadness, anger, fear, tension, relaxation, solemnity, tenderness and many others. For example, happiness is typically associated with fast tempo, major mode, rather high pitch and loudness, consonant harmony, staccato articulation (rather than legato), bright timbre, rapid tone onsets and endings, whereas expressions of sadness show about the opposite relationships: slow tempo, minor mode, relatively low pitch and loudness, elements of dissonance, legato articulation, dark timbre, and slow tone onsets and decays.

A recent review of empirical studies on both music and speech (Juslin & Laukka, 2003) reveals that there also are obvious similarities in the factors behind emotional expression in music and speech. For instance, expression of anger has (a) fast speech rate / fast tempo, (b) high voice intensity / high sound level, (c) F0-contours up / pitch contours up, and (d) microstructural irregularity in speech and music, respectively.

It should be noted that these studies examine perception of emotion, not induction of emotion. That is, the focus is on listeners' recognition of specific emotions in music, not on their own emotional reactions (see further discussion in 1.3 Earlier empirical studies). The present thesis likewise focuses on perceived emotion, that is, to what extent emotions are represented in melodic structure and performance. Questions about representation and communication of emotions in music have often been discussed. A common idea (Langer, 1957; see also in Dowling & Harwood, 1986) is that there is a resemblance (isomorphism) between the structure of feelings and the structure of music, music can be considered as an icon of feelings. Music aestheticians as Kivy (1980) and Davies (1994) both argue that "natural" expressiveness in a musical piece is rooted in surface resemblance between aspects of the music - especially its dynamic character - and aspects of human behavior. According to them the resemblance between music and emotion stems from analogies between aspects in the music (e.g., melody, rhythm, meter, and tempo) and aspects of human gestures and actions (e.g., human bearings, gaits, and movements) that are indicative of emotions. An important question is to what extent these relationships may be "intrinsic" or "acquired". Researchers focusing on the origin of non-verbal communication of emotions argue that we are genetically predisposed to perceive emotions (e.g., Tomkins, 1962; Izard, 1972; Ekman, 1973; Clynes, 1977), especially a limited set of basic emotions from which more complex emotions may be derived. These issues have been recently discussed with regard to music by Juslin (1997a; 1997b; 1998; 2001) and Peretz (2001). Peretz, Gagnon and Bouchard (1998) found that a brain-injured listener easily recognized happy and sad expressions from tempo and mode, even though the brain damage hindered adequate processing of temporal and pitch information in cognitive tasks. The listener's judgments of emotion appeared to be immediately arising from the operations of a (spared) emotionally - relevant processing subsystem. Important works have also been achieved to explain and clarify the usage and the origin of the expressive code (Juslin, 1997a; 1997b; 1998; 2001) as well as the similarity between music and speech (Juslin and Laukka, 2003).

The question regarding intrinsic or acquired emotions is outside the scope of this thesis. However, it is obvious that some expressive qualities in the music only make sense to listeners familiar with certain cultures, different practices, and different musical styles. Factors as tonality, mode and harmony are rather specific for different cultures, musical styles and epochs, whereas other factors such as tempo, pitch height, loudness, timbre and articulation should be relatively more independent of culture. The interest in this thesis is to examine perception of emotional expression in western tonal music, which is of course an obvious limitation. However, the focus on the impact of musical structure on emotional expression may contribute to an understanding of fundamental expressive effects. Even if the tonal systems differ in different parts of the world, this does not exclude the possibility of more general outcomes, for instance, experiences of musical tensions and resolutions resulting from the musical structure.

The psychology of music has also been influenced by music theory and musicology. In fact several music analysts such as Heinrich Schenker, Leonard Meyer, and Rudolph Reti have adopted psychological approaches to their analyses in so far that they were concerned with the experience of musical sound rather than with the sound itself (Cook, 1994). Music psychologists have often been interested in testing music theory in experimental settings. Perception and cognition of musical pitch, mode, and harmony have been investigated in numerous studies; see Krumhansl (1990) for a review. However, emotion in music has usually not been much addressed in music theory and adjoining works on music cognition. Adherents of formalist approaches to music, for instance, Hanslick (1891), claimed that the contents of music were simply ”tonally moving forms” to be studied and enjoyed as such, whereas they showed less interest in music's possible references to extra-musical phenomena such as music's representation of situations, events and feelings.

The pioneering and often cited book by Meyer (1956) ”Emotion and meaning in music” is an exception. However, his theory on musical emotion refers to emotions arising in the listener as a consequence of inhibited expectations. That is, explicit connection between musical structure and specific emotions, such as happiness, sadness and anger, never became the topic for discussion, even though many of his analyses of musical structure are useful to the general knowledge about music and emotion. Critics have meant that Meyer talked about music in more or less the usual way that is in the vein of musicology and much of later writings seem open to the same criticism (Cook and Dibben, 2001).

In music theory, then, music is typically described in terms of structural qualities such as tension, stability/instability and complexity, whereas description in emotional terms is rare. An interesting question concerns the relationship between structural and emotional variables. For instance, Nielsen (1983) suggested that tension occupies a position between structural and emotional variables. On the one hand, perceived tension is dependent on various structural variables; on the other hand tension may mediate emotional expression. This thesis attempts to make use of certain concepts from music theory and analysis in investigating listeners' perception of emotion in examples of tonal music (see 2. A dynamic view of melodic organization and performance).

1.3 Earlier empirical studies

Empirical studies of emotion in music date back to the very beginning of the 20th century (Gabrielsson & Lindström, 2001; Gabrielsson & Juslin, 2003). Most studies have focused on listeners' perception/recognition of emotion in music (e.g., this is sad music); there are less studies concerning listener's own emotional reactions (e.g., this music makes me feel sad). The relationship between perceived emotion and induced emotion was discussed in Gabrielsson (2002).

Many studies (i.e., Gundlach, 1932; 1935; Watson, 1942; Wedin, 1969; 1972; Nielzén & Cesarec, 1982) have used selected pieces of music available in recordings and asked listeners to judge their emotional expression using choice among or rating of given descriptive terms, usually adjectives. The resulting expressions have then been related to various characteristics of the music. In contrast to this *post hoc* procedure, other studies (Hevner, 1935; 1936; 1937; Rigg, 1939; Scherer & Oshinsky, 1977; Juslin, 1997b) have used experimental manipulation of musical factors and asked listeners to judge the emotional expression of the different versions. By varying one factor at a time (e.g., mode) while keeping other factors constant, differences in listeners' judgments of various cases could be regarded as an effect of the manipulated factor. Both approaches have advantages and disadvantages. Using existing pieces of music in real performance provides ecological validity, but the conclusions concerning relationships between musical factors and emotional expression remain less definite than when experimental manipulation is applied.

The number of factors examined within the same study has varied. Usually several features have been considered, but in some studies the focus has been on a single or just a few features (see references in Gabrielsson & Lindström, 2001). In the former case the researcher may be able to study the relative importance of different factors within the study; the latter case may allow more intense study, and perhaps better understanding of the effects of a certain factor.

Other differences regard the choice of rating scales (judgment scales) and descriptive adjectives. In many studies the extraction of meaning is accomplished using multivariate statistical techniques (e.g., factor analysis, cluster analysis, multidimensional scaling) to find common traits underlying listeners' ratings of a large number of emotion adjectives. In other studies the focus is directed to only a limited number of emotions, for instance, so-called basic emotions, and the selection may be governed by some underlying theory (see references in Gabrielsson & Lindström, 2001).

A final distinction is between studies that examine perceived emotion as inherent in the musical score, that is, composed music, and studies in which perceived emotion is studied using different performances of the same piece of music. In the latter case the typical procedure is to have musicians perform monophonic pieces of music to express different emotions. By analyzing the physical characteristics of these performances it becomes possible to study the relationships between the physical and experiential (emotional) variables in much detail. This approach has become dominating since some ten years ago (e.g., Gabrielsson & Lindström, 1995; Gabrielsson & Juslin, 1996; Juslin, 1998).

The papers in the present thesis examined perceived emotion using systematic manipulations of the melodic structure of a simple tune/composition (Study I), different performances of selected versions of the same tune (Study II), and both systematic manipulations and (quasi/computer-) performances of short three-tone patterns (Study III).

1.4 Interplay between musical factors

When listening to music, we never perceive factors such as mode, harmony, pitch level, loudness and rhythm in isolation from each other, they appear in a continuous interplay. This does not necessarily exclude the possibility that separate musical factors may be additive in achieving perceived emotion, but it seems more probable that perceived emotion results from a complex pattern of both the effects of single musical factors and interactions between them.

There is, however, an obvious lack of research on such interactions (Gabrielsson & Lindström, 2001). Schellenberg, Krysciak and Campell (2000) recently found empirical evidence for pitch x rhythm interaction affecting emotional expression. Melodies representative of happy, sad, and scary expressions were selected. The listeners' task was to rate the expression of these melodies and of three additional versions of each example: (a) rhythm as in the original version but all tones set to the same pitch; (b) the tones (pitches) intact but all set to the same duration (rhythm neutralized); or (c) all tones of same duration and pitch. The effects of pitch and rhythm varied across melodies with the strongest influence from pitch. Whenever rhythm affected the ratings, it did so in interaction with pitch. However, the authors did not provide any interpretation of the interactions.

In the present thesis an attempt is made to interpret certain interactions as resulting from emphasis (accentuation) on important notes in the melodic structure. This is discussed within the framework of a dynamic view of melodic organization and performance described in the following.

2 A dynamic view of melodic organization and performance

The following sections will highlight certain perceptual mechanisms in music that are central for the hypotheses in this thesis. The first assumption regards the relative importance or function of certain notes for the perceived structure and the perceived emotion. I will argue that certain notes in the melodic structure are perceived as especially *expressive* as a function of implied harmony, tonality and position in the key. More specifically, certain notes could, in the specific tonal context, be perceived as tense and unstable, and in that way also carry expressive information of instability and emotions such as anger; or notes could be perceived as relaxed and stable, therefore implying stability and emotion such as tenderness. A supplementing assumption is that these notes may be activated - receive emphasis - by local musical features such as rhythm, melodic contour, loudness and articulation resulting in specific accent structures.

Of course, the idea of a dynamic view of tonal music is not new but stems from previous research. My contribution concerns some hitherto neglected aspects of such a view. A dynamic view of tonality stands in opposition to the traditional static view. The latter has been criticized because it describes the tonal hierarchy in fixed patterns (Butler, 1989). Tonality has, in other words, often been described as if it were completely unaffected by other musical features, such as melody, rhythm, loudness, etc. The assumptions addressed above refer to the activation (emphasis) of certain expressive notes as a function of the interplay between musical features such as melodic structure, mode, (implicit) harmony, rhythm, loudness and articulation. Due to this interplay these notes become especially important for the listener's perception of the musical structure and, by assumption, also for the perceived emotional expression.

2.1 Music perception and expressive notes

2.1.1 Tonal hierarchy

Tonality refers to a hierarchically organized system of pitch relationships (Cuddy & Badertscher, 1987; Krumhansl, 1990). The tonal hierarchy means that notes in the same key differ with regard to perceived stability. Taking C major as an example, the tonic, C, that is, the first scale degree (I), is the most stable note followed by the dominant, the fifth scale degree (V), that is, G, and the mediant (the third scale degree, III), that is E (Krumhansl 1979; 1990). Together these notes form the C major triad and correspond to the second, third, and fifth harmonics of the tonic, that is, the lowest distinct harmonics in the natural harmonic series (Sloboda, 1985). The notes in the remaining diatonic scale degrees B, D, F and A are less stable and tend to be pulled toward—resolved into—a neighboring stable note. The labels tonic, supertonic (the note an octave above the tonic), mediant, subdominant (the note below the dominant), and dominant suggest something about the functions of the different notes and their relationships (Povel & Egmond, 1993). Notes that are not included in the diatonic scale but are part of a complete chromatic scale (C sharp, D sharp, F sharp, G sharp and A sharp) are least stable. These conceptions from music theory have been confirmed in experiments in which listeners rated how well different notes, diatonic as well as chromatic, fitted into a key-defining context such as scales, tonic triads and chord cadences, in both major and minor mode (Krumhansl, 1990, p. 30). Krumhansl and Kessler (1982) derived tonal weights for each chromatic note in major and minor key contexts, defining the level of perceived stability of each note. In tonal music composers tend to use stable pitches at phrase endings and as the final note in melodies (Dowling & Harwood, 1986). The concept of tonal hierarchy and the effects of instability and stability, tension and resolution are central to the notions of emotion and meaning in music (Meyer 1956, also cited in Krumhansl, 1990), and is also a fundamental aspect in this thesis. However, as earlier argued, the expressive effects of different notes have to be considered in relation to the actual musical context, further discussed below.

2.1.2 Chord progressions

In western tonal music, single notes of a melody are usually perceived as related to an underlying harmonic structure, a chord, mainly based on tonal triads. Sequences of such events are called chord progressions or harmonic progressions. The probability of various harmonic progressions, implicit or not, is often listed in tables of usual ”root progressions” (Piston, cited in Meyer, 1956, p. 54). The root is the position of the tonic in each chord and is designated with roman numerals indicating the position by degree in the diatonic scale with beginning on the keynote. The I-V progression in C major thus means a C major chord followed by a G major chord, which is among the most usual progressions. The I-IV progression (C major – F major) is also usual and often followed by a V-I progression, together forming a full cadence. On the other hand the second (II) and third (III) roots are seldom combined with the tonic (I). The tonic (I), subdominant (IV) and dominant (V) chords together represent all notes in the diatonic scale (the white keys on the piano with the C major tonic as reference). The consequence is that the majority of melodic sequences could be harmonized or accompanied by these three chords. Since the overlap between the three chords is at minimum the progression will be perceived as clear and decisive. However, there are of course melodies that cannot be harmonized or "go well" with these three basic chords in an easy way but require other chord combinations. There are in fact a lot of possible implicit chord combinations in music. The attribution of chords is dependent both on various perceptual conditions such as perception of consonance and dissonance, and to historical conventions, different musical styles etc. However, even the perception/evaluation of consonance and dissonance has changed over time (Dowling & Harwood, 1986).

Meyer (1956) emphasized the importance of vague and ambiguous progressions for the aesthetic and emotional expression in music. However, the expressive potentials of implied harmony have been mainly neglected in empirical research so far (Gabrielsson & Lindström, 2001).

2.1.3 Key and mode

Different keys, or tonal centres, may also be arranged hierarchically. In tonal music a tune or a composition is usually related to a single key, for instance, C major, but may temporarily change (modulate) into another key. The perception of a specific key can be more or less distinct. In some cases it may be felt definite and stable, as when the harmonic progression mainly includes the chords underlying the tonic triad, in other cases ambiguous and changing when more unusual chord progressions come into play.

Different keys can be more or less distant to each other, usually explained with reference to the circle of fifths upwards—C, G, D, A and so on until one arrives at C again—or downward: C, F, B flat, E flat….C. Keys in neighboring positions in the circle are more related to each other than keys at larger distances, for instance, the keys of C and G are closer related than the keys of C and D. Alteration of key(s) may bring about feelings of instability and ambiguity, factors that according to Meyer (1956) largely determine the affective response of the listener (also cited in Krumhansl, 1990).

The emotional impact of the minor mode compared to the major mode has been the subject of much debate and puzzle among scholars. Meyer (1956) mentioned several factors of importance in this question. First, from a harmonic point of view, the minor mode is more ambiguous and less stable than the major mode. Ambiguity arises because there are more possible chord combinations in minor mode than in major mode, and consequently the probability of any particular harmonic progression is smaller. Moreover, the minor mode is potentially chromatic (thus less stable), and the tendency of the third toward the tonic is stronger in minor mode than in major mode since it is closer to the second scale degree and thus to the tonic. The minor third has a strong tendency for resolving to the tonic and a delay will be intensely felt (Meyer, 1956, p. 225).

The major and minor thirds are the notes that primarily distinguish between major and minor mode. It can be expected, therefore, that the position and emphasis on these notes also will be important with regard to the emotions usually associated with these respective modes, that is, major mode associated with happiness, minor mode with sadness.

From the discussion in this and the preceding sections (2.1.1, 2.1.2) it follows that single notes in the melodic structure may have certain specific characteristics depending on their general position in the tonal hierarchy, their relative stability or instability, their relation to underlying harmonic progressions etc. While these characteristics concern the perceived structure of a melodic sequence, it may be assumed that they also have an impact on the emotional expression. What I would suggest is that stable notes are important for representation of stability and relaxed emotions such as tenderness, whereas unstable notes are apt for representing instability and emotions such as anger. Furthermore, notes which are decisive for the identity of major and minor mode would be important for the emotional expression usually associated with these modes (major mode - happiness, minor mode - sadness).

2.2 The activation of expressive notes

From the preceding sections it is evident that perception of melodic structure is affected by the properties of the tonal hierarchy. However, although the static view of tonal hierarchy has been very convincing, it does not acknowledge the importance of dynamic activation. A great body of research has shown that tonal tendencies within a musical sequence could be activated locally by (a) the surface structure of a melody and (b) the performance of the melody. The following review will shed some light over possible accent structures in melodic organization and their assumed importance for emotional expression.

2.2.1 Accents in melodic organization

There is still little empirical research on how implied harmony is evoked in melodic sequences or in chords played arpeggio. However, it has been pointed out that musical features like rhythm and serial ordering of notes may have dynamic impact on the perception of tonal structure (Jansen & Povel, 1999). A similar argument is that "In short, both a rhythmic framework and criteria for pitch stability are needed if the listener is to hear events in a dominating - subordinating manner" (Lerdahl, 1988, p. 316). Furthermore, meter and rhythm may ”coincide ” with the tonal hierarchy, for instance, if the stable notes of the major triad occur at phrase endings or at other important positions in the melodic sequence, thus reinforcing the effects due to the note's position in the tonal hierarchy. The relative salience of different notes as exhibited in accent structures is therefore assumed to be important for the listener's perception of the musical structure. It affects perception of implied harmony and also, as addressed in this thesis, affects the perception of emotional expression in music.

The meaning of the term accent has varied considerably in music-theoretic literature (Friberg & Battel, 2002). Two main categories have been identified, immanent and performed accents (Parncutt, in press; Lerdahl & Jackendoff, 1983). The former is assumed to be apparent in the notated score, whereas the latter refers to accents under the control of the performer. A note is accented when it is marked for consciousness in some way. Jones (1987) says that "an accent is anything that is relatively attention-getting", and Drake and Palmer (1993) state that "an accent is an event that stands out and captures a listener's attention". There are several (immanent) accents in a melody. Metric accent refers to the comparatively strong first and third beats in quadruple meter. Melodic accent occurs at big leaps in pitch, or at the turn of the top note in the melodic contour. ”Valleys” in the melodic contour may also be perceived as accents, but are not as effective as peaks (Thomassen, 1982); an explanation may be that important phonemes in speech utterances tend to be relatively high in pitch (Parncutt, in press). The final note in a measure or phrase is also perceived as accentuated. Accents may, furthermore, arise at boundaries in rhythmic groups, or by means of harmonic progression (see Thomassen, 1982; Huron & Royal, 1996). Sometimes more than one kind of accent is active on the same note. This is often defined as so-called *joint accent structures* (Jones, 1987).

Other factors related to the surface level of the melodic structure, such as gestalt principles of proximity, good form and good continuation may also have an influence on the organization of harmonic materials. For instance, Deutsch (1984) and Deutsch & Feroe (1981) suggest that when presented with a sequence of notes, the listeners form low-level groupings on the basis of pitch proximity, and in the absence of other cues, assign greater prominence to the last note of each group. This means that the specific positions and orderings of notes included in a melody are crucial for the shaping of the underlying or implied harmonic function.

Other order principles were demonstrated in experiments by Bharucha (1984) and Laden (1994). They used the term ”anchoring of tones” which means that some notes, mainly those at unstable positions, became attached to a particular harmonic structure. The anchoring appears if a note approaches a following note by stepwise motion along the chromatic or diatonic scale, and means that the dissonant or unstable note is assimilated to the actual key. For instance, given a key context in C-major, the note C sharp will be anchored to C if the serial order is C sharp – C, but not in the reverse order, that is, C followed by C sharp.

Rhythm and meter are often closely related to each other (Graybill, 1989). The longest note in a rhythm pattern will usually be more salient to the listener. If it appears on the second (weak) beat in quadruple meter, this may be perceived as syncopation. "..the effect of syncopation is there as part of the basic primary beat" (Meyer, 1956, p 121). That is, the salience of the weak beat is unexpected and the syncope may be perceived as an accent. Expressive effects may also be achieved by delaying the process of resolution (Meyer, 1956, p 225).

Jones and Boltz (1989) emphasized the importance of temporal structure which guides attention toward or away from notes positioned at various places in the tonal hierarchy. Boltz (1989) found that listeners perceived the complete endings of melodies when resolutions (leading note to tonic) appeared "on time" according to expectations and the temporal structure. A conclusion is that notes in the melody and the hierarchy of temporal information are lawfully related to each other (Boltz, 1993).

Almost all contemporary music-theoretic analyses have more or less adopted ideas of expectation (Schmuckler, 1989); this concept appears in many areas of psychology, such as perceptual organization, learning theory, attention, memory, and linguistic processing (Schmuckler, 1990). In music, expectation means that an event (note, chord, rhythm, harmonic progression etc) implies that certain other events (notes, chords etc) will appear in the following. The best known psychological theory regarding musical expectation is Meyer's (1956) theory. He suggested that violation of expected patterns will evoke an emotional response in the listener. It should be noted, however, that this theory thus refers to emotional response, not to perceived emotion, and that the relationship between perceived and induced emotion may vary in different ways; (Gabrielsson, 2002).

Narmour (1990), following Meyer (1956), developed certain rules in his so-called implication - realization model. The model describes how melodic passages can be felt as more or less satisfying depending on the realization of implications. For instance, a stepwise progression, both upward and downward, implies a continuation in the same direction to the tonic (or dominant). Another rule states that a large interval at the beginning of a phrase will be followed by a smaller interval in opposite direction, that is, the gap of the large interval demands to be filled.

However, it is obvious how musical expectation or inhibition of something expected also may be perceived as an accentuation.

Generally, then, there are several factors besides the "static" tonal hierarchy that influence listeners' perception of tonality and melodic structure. These other variables may sometimes reinforce, sometimes contradict the sense of key or harmony to different degrees. Krumhansl (1990, p. 110) notes that "it is important to bear in mind that music is, in some cases, intentionally structured to place various structural properties in opposition, creating perceptual ambiguity and tension that is resolved only when the separate components finally come into correspondence”. The crucial topic for this thesis is to what extent these other variables also affect perception of emotional expression in music.

2.2.2 Accents in music performance

Measurements of performances in music have been undertaken since about 1900 and demonstrated many types of deviations from the nominal values notated in the musical score (see reviews in Gabrielsson, 1999, 2003). Deviations in timing, articulation, dynamics, and intonation are generally considered important to provide a genuine and live character to the music and have been described as *artistic deviation from the exact and rigid* (Seashore, 1938, p. 249), *systematic variations* (Bengtsson, Gabrielsson & Thorsén, 1969, Bengtsson & Gabrielsson, 1980), or *expressive transformations* (Clarke, 1985). The purposes of such deviations are discussed both in terms of their importance for the perception of the musical structure and for their impact on the emotional expression.

A common research finding is that deviations in music performance are made in order to emphasize the structure of the piece, for instance to clarify the meter, signal the end of phrases by slowing down, separate phrases by micropauses, resolve possible ambiguities etc. (Bengtsson, 1974; Clarke, 1985, 1988; Sundberg, Friberg & Frydén, 1991, De Poli, Roda & Vidolin, 1998). These procedures may involve explicit use of performed accents. There are several ways to make a note accentuated in music performance, for instance, increase its loudness and/or duration, anticipate or delay the onset of the tone, play it more legato or more staccato than the neighboring tones (Drake & Palmer, 1993; Friberg & Battel, 2002). Making a tone accentuated in performance may be a mean of clarifying the structure of the music to the listener, for instance to resolve possible ambiguity regarding the proper meter or simply reinforce the meter inherent in the score. Generally the accentuated tone will receive increased attention and thus affect the listener’s perception in various ways. Accents by pitch, intensity, and durations are in a similar way used in speech prosody to direct attention to focal words in a sentence and important phrases (Patel & Peretz, 1997). For instance, the meaning of ”Meet me IN FRONT OF the bank, I said” is quite different from ”Meet me in front of the BANK, I said”.

While the above examples concern performed accents to serve the perception of the musical structure, little, if anything, is known about the role of performed accents for the perception of emotional expression. There are several studies on how various performance characteristics can affect the motional-emotional character of a piece, for instance, the too “early second beat“ in Viennese waltzes (Bengtsson & Gabrielsson, 1983; Gabrielsson, 1986, 1988), and on how musicians use tempo, timing, articulation, intonation and timbre to express various emotions in performances of the same notated piece (Gabrielsson, 1994, 1995; Gabrielsson & Juslin, 1996; Gabrielsson & Lindström, 1995; Juslin, 1997a, 1997b, 1998). For instance, happiness and anger were performed fast and mostly with staccato articulation. Anger was performed louder than happiness but with abrupt tone onsets and harsh timbre. Performances of tenderness and sadness had legato articulation, slow and soft tone onsets, and soft timbre in tenderness. Expressionless versions were characterized by no or considerably smaller variation in the above-mentioned variables.

It seems probable that such performances involve the use of various accents in performance, but this question has not been addressed in the investigations so far. In analogy with my earlier reasoning on the importance of accents for both structural and emotional purposes, it is of course expected that if certain notes are especially marked in musicians' performances, they serve the same purposes—most probably in order to reinforce the effect(s) stemming from notated accents but perhaps sometimes to de-emphasize (counteract) such effects.

2.3 Aim of the thesis

The general aim of this thesis is to elaborate some aspects of a dynamic view of melodic organization and performance, which may have implications not only for the perception of musical structure but also, especially, for the perception of emotion in music.

As mentioned earlier, there is a renewed and increasing interest in music psychology for theory and research on emotion in music (Juslin & Sloboda, 2001), whereas such questions have received little attention in conventional music analysis as well as in related studies on music cognition. My idea is to put some focus on items in music analysis that could be essential for emotional expression in music. Music theorists usually claim that certain notes, in a tonal context, are perceived as more stable, or unstable, than others. This may imply that such notes also affect the emotional expression, in other words, mediate perception of emotion in music (cf. Nielsen, 1983), especially if they receive emphasis through accent structures, in the melodic structure itself or as a result of performance.

Faced with the idea that single notes have different stability in tonal contexts, I suggest that some of these notes could be called ”expressive notes” in the sense that the note carries information about emotional expression. In order that this information really comes through to the listener, these expressive notes have to be activated, that is, somehow emphasized, either through accent structures in the melodic structure itself or through such accent structures provided by a performer.

This then is the common theme for the three empirical studies described in the following. Study I applies systematic manipulations of a simple tune and listening tests to begin to explore the ideas about expressive notes. In Study II the focus is on the possible interplay between factors in the melodic structure and in the live performance of such structures. Study III, finally, in a way combines the approaches taken in the previous two studies by using systematic variation of short tone sequences as well as (quasi/computer) performance of the sequences.

3 Empirical studies

3.1 Study I

Impact of melodic organization on perceived structure and emotional expression in music

3.1.1 Aim and background

In Study I an attempt was made to elaborate some new ideas of melodic organization which may have implications for listeners' perception of both structure and emotional expression in music. The main background was a dynamic view of tonal music suggested by several authors (Boltz, 1993; Jones, 1987; Bharucha, 1984; Brown, 1988; Butler, 1989; Cuddy; Cohen & Mewhort, 1981; Deutsch & Feroe, 1981; Povel & Egmond, 1993). The main idea for this view is that perception of tonality is affected by melody and rhythm. Important questions for me were if certain notes in a melody may have special expressive characteristics and to what extent emphasis (stress) or suppression of these notes by melodic accent structure would affect listeners' perception of structure and, especially, emotional expression.

It was hypothesized that: (a) Emphasis on certain key/chord-relevant notes may affect listeners' perception of both structural and emotional properties. (b) Emphasis on a stable/relaxed note may promote perception of stability and tenderness. (c) Emphasis on an unstable/tense note may promote perception of instability and anger. (d) Such effects may be especially obvious when emphasis is achieved by so-called joint accent structures. Furthermore (e) emphasis on a note important in major mode may promote perception of happiness, whereas (f) emphasis on a note important in minor mode may promote perception of sadness. (g) Interactions between the manipulated factors may be possible to relate to the position and emphasis on key/chord-relevant notes.

3.1.2 Methods

Participants. Nineteen non-musicians (psychology students), 13 females, 6 males, aged 20 – 31 years, and 20 musicians, 10 females, 10 males, 21 - 39 years, volunteered to participate in the experiment.

Musical material. The musical material consisted of systematic manipulations of the well-known tune *Frère Jacques* regarding latent harmony, rhythm pattern, and melodic contours, all performed and generated "dead-pan" by a computer. Six melodies were generated that looked like the original tune concerning melodic lines and contours but differed in tonal progression. The intention was to create a variation extending from simple and stable harmony to complex and unstable harmony, across versions of both major and minor mode. These "prototypic" melodies were further systematically varied in melodic direction, melodic contour, and in rhythm, resulting in a total of 72 sequences. All sequences were programmed in a tempo of 120 bpm (beat = quarter note), and all notes were given the same loudness and articulation (the onset-offset duration was 75% of the inter-onset interval). The sequences were reproduced at a comfortable sound level using a DAT recorder, an amplifier and loudspeakers.

Procedure. The experiment was conducted in a lecture hall. The participants were instructed that they were going to listen to different melodies that all were variations of a famous tune and that their task was to rate each melody with regard to its expression on six bipolar scales, "stable – unstable", "simple – complex", "relaxed – tense", "happy – sad", "tender – angry", and "expressionless – expressive". The subjects were divided into several groups and listened to all 72 musical sequences in one of two randomized orders, distributed over two sessions with one-hour break between the sessions. Each session lasted for about 45 minutes, including a short break in the middle.

3.1.3 Results

Correlations. There were high correlations (redundancy) between the different judgment scales (r = .25 - .94) and between the two listener groups (r = .60 - .94). This allowed a reduction to analysis of only two scales, instability and sadness, judged by non-musicians.

ANOVA. In both scales there were significant main effects of progressions and rhythm patterns but not of melodic direction or contour. However, there were several interactions that usually could be linked to differences in accent structures; such differences were also found between versions which represented the largest contrasts within the same progression. For instance, analysis of the stable-unstable scale showed as many as five two-ways and one three-way interaction, and the analysis of the happy-sad scale revealed four two-way, one three-way, and one four-way interaction.

Versions with stress on stable notes in the respective tonal context (i.e., tonic C and/or major/minor thirds E/E-flat) were judged relatively more stable than versions in which accent structures implied stress on tonally unrelated notes (i.e., passing note D).

The results of the happy-sad judgments showed that the two higher order (three-way and four-way) interaction effects could be partly explained by emphasis on, or suppression of, certain notes resulting from the accent structure. Judgment of a version as less happy could be related to suppression of the major third (E), a note essential for major mode, whereas a happy version showed emphasis on the same note. Judgments of sadness seemed furthermore to be related to emphasis on notes important for defining the minor mode, such as A flat.

In accordance with the hypotheses ANOVA showed that emphasis of important notes in the respective key may affect perceived structure and emotion. This suggestion was more thoroughly examined by means of multiple regression analysis.

Multiple regression analysis. Multiple regression analyses were conducted in order to study the relative importance of the manipulated variables and furthermore two new variables reflecting variation in accent structure and the emphasis on stable and unstable tones, called *"Non Triad"* and *"Triad"* described below. The predictors were coded as follows: Mode, major = 1, minor = 2; Rhythm, even =1 (1:1 ratio), uneven = 2 (3:1 or 1:3 ratio); Melodic contour, original = 1, complex = 2; Melodic direction, original = 1, reversed = 2; Musical training, non-musicians = 1, musicians = 2; Harmony, simple = 1, moderately complex = 2, and complex = 3.

The new predictor variables were derived from a score matrix for each version that shows occurrence of accents due to meter (on the first and third beats of the measure), melody (upward jump and pivot notes), and rhythm/duration for each note in each version. Table 1 shows the scores for the version with the original melodic contour and even 1:1 rhythm ratio (that is, the original tune). The rightmost column "Sum of accent scores" shows levels of joint emphasis by accents for each specific note in this version. The analogous principles were used to allocate accent scores for all 72 melodic versions.

Table 1.

Allocation of scores for different types of accents on notes in the original version of Frère Jacques.

Measure	Tones	Metric first beat	Metric third beat	Jump	Pivot	Duration	Sum of accent scores
1-2	C	2	0	0	0	2	4
	D	0	0	0	0	2	2
	E	0	2	0	2	2	6
	C	0	0	0	0	2	2
3-4	E	2	0	0	0	2	4
	F	0	0	0	0	2	2
	G	0	2	0	2	4	8
5-6	G	2	0	1	0	1	4
	A	0	0	0	2	1	3
	G	0	0	0	0	1	1
	F	0	0	0	0	1	1
	E	0	2	0	0	2	4
	C	0	0	0	0	2	2
Total							43

The next step was to identify tense/unstable notes in each version, note the sum of accent scores for each of them, and then finally add all of them to obtain a total sum of accent scores on tense notes for each of the 72 versions. For instance, tense/unstable notes in the original tune were D, F, and A, that is, all notes outside the tonic C major triad (C, E and G).

After identifying the tense notes in each melodic version, their accent scores were added to provide a total sum of accent scores on tense notes in each of the 72 versions. This predictor, called *Non Triad*, was used together with the earlier mentioned predictors—mode, harmony, rhythm, melodic contour, melodic direction, musical training—in stepwise multiple regression analyses to examine its relative importance for each of the six judgment scales. The dependent variables were the mean ratings of the six judgment scales, separately for non musicians and musicians (n = 72 x 2 = 144).

The variance accounted for was high (R^2 = .70 to .87) for all scales, except for expressiveness (R^2 = .50). As could be expected, the highest beta weights usually occurred for (minor) mode, (complex) harmony and (uneven) rhythm.

However, the most interesting result with regard to the hypotheses was that the new predictor Non Triad, indicating structural accents on tense notes, contributed significantly (beta weights .21 to .34, p< .001) to perceived instability, complexity, tension, and anger. It was, however, not included in the regression model for sadness and expressiveness.

As a supplement and further check, corresponding regression analyses were conducted using the same predictor but replacing the Non Triad variable with a variable that reflected presence of accents on *stable* notes, that is, notes belonging to the triad of the respective harmonic function, called *Triad*. The sum of accent scores on stable notes for each version was simply obtained by subtracting the number of accent scores for tense notes from the total number of accent scores for the respective version. The results looked very similar to the results of the earlier analyses but with reversed sign in comparison with Non Triad.

In the preceding analysis, the Non Triad and Triad predictors played no or only little role for sadness and expressiveness. Therefore, further regression analyses were performed separately for minor and major mode versions, thus removing mode as predictor since (minor) mode was the by far most important predictor, especially for sadness (*Beta* = .91, $p<.001$). The results showed that when (minor) mode was no longer a predictor, the variable best reflecting the minor triad character, that is, Triad, showed up as the most important predictor of sadness (and Non Triad worked in the opposite direction). However, for versions in major mode, neither Triad nor Non Triad contributed to perceived sadness.

3.1.4 Conclusions

Taken all together, the findings from ANOVA and inspection of maximum contrasts in ratings showed both main effects and intricate interactions between the manipulated factors, which usually could be interpreted as related to accent structures and stress on certain notes, thus supporting hypothesis (a). Emphasis on stable notes by metric, melodic and rhythmic accents affected listeners' perception of stability and tenderness, whereas stress on unstable notes (passing tones and tones unrelated to implied harmony) affected perceived instability and anger; these results supported hypotheses (b) and (c). Emphasis on notes important for identification of the major mode (e.g., major third) increased perceived happiness, and emphasis on notes defining the minor mode increased perceived sadness, thus corroborating hypothesis (e and f). Interactions between manipulated factors could be related to the position and emphasis on key/chord-relevant notes, supporting hypothesis (g). Further evidence for these results was found in stepwise multiple regression analyses, where predictor variables representing accents on tense (non triadic) and relaxed (triadic) notes affected perceived instability, complexity, tension, sadness, anger and expressiveness. These predictors further indicated that the total sum of joint metric, melodic and rhythmic accents had effect in accordance with hypothesis (d).

Since various aspects of the melodic structure carry expressive potentials, an interesting question is if and to what extent live *performance* of melodic structures may be influenced by the emotional expression inherent in a given melody. This question was the object for Study II.

3.2 Study II

Interplay between musical structure and performance in emotional expression of music

3.2.1 Aim and background

In earlier empirical studies very little has been said about the *interplay* between, on the one hand, factors in the musical piece itself and, on the other hand, the performance of the piece, (Gabrielsson & Lindström, 2001). The aim of Study II was to examine the relationship between factors in the melodic structure and factors in the performance. The general question was if musicians would adjust for lack of inherent expression in melodic sequences by their performance and if they would perform accents and emphasize certain notes important for perception of emotion.

It was hypothesized that (a) The musicians would take the inherent expression in melodic structure into account and therefore somehow compensate for lack of inherent expression. For instance, lack of inherent happiness would be compensated for by a more explicit ”happy” performance. The analogue type of compensation was expected to apply for the other emotions as well. (b) A note assumed to be important for a certain emotional expression would be marked for consciousness in a performance intended to express the particular emotion (e.g., happiness) but not in a performance intending to express an opposite emotion (e.g., sadness). (c) Emphasis on stable notes in a given key/chord-context may be apt for perception of stability and tenderness. (d) Emphasis on notes considered unstable in a given key/chord-context may lead to perception of instability and anger. (e) Given that major and minor mode are associated with happiness and sadness, emphasis on notes that are important for these respective modes should contribute to perception of these emotions. For instance, accents on notes central in minor mode should contribute to the expression of sadness.

To investigate these hypotheses two experiments were conducted, the first examining music performances by two musicians and the second examining listeners’ ratings of the recorded performances.

3.2.2 Methods

Performers. Two professional male musicians (organ players), both about 30 years of age, volunteered to participate in the first experiment. Nineteen skilled musicians (12 females, 7 males, 19 - 30 years old) and seventeen psychology students (9 females, 8 males, 20 - 48 years old) volunteered to participate as listeners in the second experiment.

Musical material. The stimuli were a selection of sequences derived from the manipulated *Frère Jacques* versions which had been rated for their expression in Study I. Out of the 72 rated versions, 2 versions from different progressions were chosen, which had received the highest and next highest mean rating in happiness (called "happy", abbreviated as H in the following). The analogous procedure was applied regarding sadness, tenderness, and anger. In addition to each of these 8 highest rated versions, also the worst (lowest rated) version within the same progression and expression was selected for comparison (i.e. called less happy, abbreviated as LH in the following). There were thus in all 16 versions, labeled "happy“ (H) versus "less happy“ (LH), "sad“ (S) versus "less sad“ (LS), "angry" (A) versus "less angry" (LA), and "tender" (T) versus "less tender" (LT).

Procedure. The musicians in experiment 1 performed the different musical sequences on a digital piano (*Casio Celviano AP-20*) connected via midi to a touch-sensitive (velocity) synthesizer (*Korg Wavestation*) using a saxophone-like timbre (the same sound source as in Study I). In the listening experiment the recorded sequences for experiment 1 were reproduced at a comfortable sound level using the DAT recorder, a *Pioneer* amplifier and *AR 12* loudspeakers.

The performer was not given any preview of the musical sequences used in the experiment. The order of the sequences was randomized differently for each performer, however, the corresponding H and LH versions were performed immediately after each other (in free order) to increase the contrast between them. The H, LH, S and LS versions should be performed to sound as happy and as sad as possible, and the T, LT, A and LA versions should be performed to sound as tender and as angry as possible. The first performer started with the "happy-sad session“, the second performer started with the "tender- angry session“.

Experiment 2, the listening test, was conducted in a lecture hall separately for musicians and psychology students. The participants in each category were divided into three groups. Each group listened to all 64 performed sequences in one of three randomized orders, distributed over two sessions with one half-hour break between the sessions. The task was to rate each melody with regard to its expression on six bipolar scales stable-unstable, simple-complex, relaxed-tense happy-sad, tender-angry scales, expressionless-expressive. However, only the happy - sad and the tender - angry scales were treated in this study.

The design involved two modes of expression in melody (i.e. happy versus sad) referred to as *expression in melody*, two *positions of expression* in melody (i.e. S and LH versions in upper position versus H and LS versions in lower position), two modes of *performance intention* (happy versus sad, and tender versus angry, respectively), and furthermore two levels of *repetition*, replication between progressions (i.e., progression 1 versus 2, and progression 4 versus 6 regarding happy-sad). The same design was applied to the tender-angry versions.

3.2.3 Results

Experiment 1 corroborated the result of earlier studies and showed that musicians' expressive intentions affected their performances. Happy and angry performances were played faster, louder, less legato, and had sharper rhythm ratios than in sad and tender performances.

However, the musicians also seemed to use loudness level (velocity) to compensate for less adequate expression inherent in melody. Figure 1 shows means and 95 % confidence intervals of velocity values, that is, loudness used by the performers. For performer A, a very clear pattern in loudness appears in intended happy performances of the different versions (Figure 1, left panel, upper row). The versions are ordered along the X-axis such that a "happy version" (H) is followed by a "less happy" version (LH) - the first four cases - and a "less sad" (LS) version is followed by a "sad" (S) version, the last four cases. The pattern means that LH versions were performed with higher loudness than the corresponding H versions, and that S versions were performed with higher loudness than the corresponding LS versions. That is, versions that were "less happy" (LH) and versions that were "sad" (S) were performed with higher loudness than their respective counterparts (H and LS), presumably in order to increase the happy expression of the version that did not sound enough happy in themselves.

For performer A:s sad performances (lower row) a similar pattern appears. In order to perform a H version to sound sad and a LS version to sound sad, he seemed to decrease the intensity in comparison with the corresponding LH and S versions. Similar results were found for tender and angry versions.

In total, as many as 26 out of 32 cases demonstrated an adjustment (compensation) for the less expressive structure by means of velocity (loudness) manipulation in the performance. Regarding articulation the corresponding compensation occurred in 20 cases out of 32, that is, performing less happy or less angry melodies with increased staccato articulation and, conversely, with increased legato articulation in less sad and less tender melodies.

However, systematic patterns of compensation by means of tempo manipulation were not apparent.

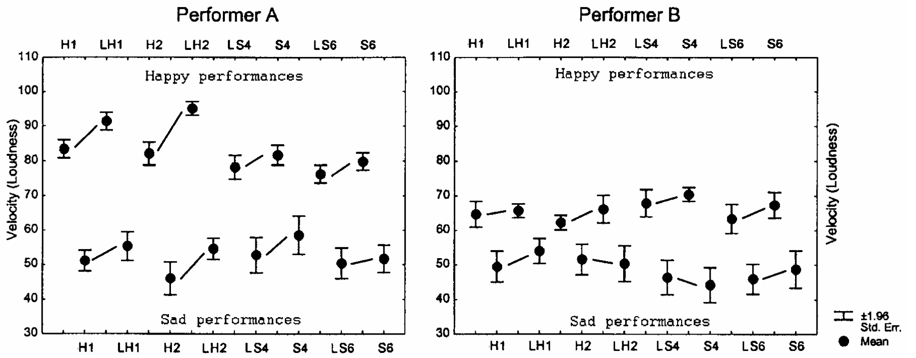


Figure 1. Means and confidence intervals for loudness (velocity) in happy (upper row) and sad (lower row) performances of all happy/sad versions, performer A (left panel) and B (right panel).

Thus, the musicians seemed sensitive to the implications of inherent expression in melodic structure and adjusted their performances at different expressive intentions, as assumed in hypothesis (a).

Performance and articulation of certain notes. The performed material was analyzed *note by note* regarding velocity (loudness), articulation and timing to see whether single notes were treated differently for different expressions. To facilitate comparison, performance data was transformed into z-values for the respective melodic version. The first step was to identify versions that showed large differences in performances, for instance, if there were large differences in z-scores between corresponding tones in intended happy versus intended sad performances. The next step was to see if this variation was intelligible and if both musicians seemed to apply the same systematic variations (patterns) around the same notes.

As an example, Appendix A shows the difference between happy performances (filled circles) and sad performances (empty boxes) of the H2 version, for both musicians regarding velocity (upper panels) and articulation (lower panels). In measures 3-4 tones 9 and 12, A (major third in an implied subdominant F chord) is stressed in relation to surrounding notes both in velocity (louder note, positive z-value) and articulation (more legato, positive z-value) in intended happy performance, but usually not in intended sad performance. The tonic C, tones 19 and 25 in measures 5-6, is also stressed in happy performance in both velocity and articulation for performer A, and in articulation for performer B. These notes (A and C) were already emphasized in the melodic structure by the longer duration (dotted quarter note) in this H2 version.

Timing may also be used for emphasizing certain notes in the performance. Appendix B shows the timing profiles for intended happy and sad performances of the S4 version (upper panel) and the LS4 version (lower panel). In the intended sad performance both performers had long duration on A flat (an important note in minor mode, tones 19 and 25 in the S4 version, and tones 20 and 26 in the LS4 version). In the S4 version A flat was also prolonged by the dotted quarter-note.

In a similar way (but not shown here), the stable tonic (G) of the implied dominant chord was stressed in tender performances either by duration or loudness. Conversely, B natural, a very tense note in C minor, was stressed by loudness in angry performances, but this was avoided in tender performances. These findings were further corroborated in statistics (mean, maximum and minimum values) of z-transformed difference values for timing, articulation and velocity (loudness) on selected target-notes, A and A flat in happy-sad performances, and G and B natural in tender-angry performances. These notes were previously discussed in Study I. The statistics revealed that A was clearly performed more legato and louder in happiness, whereas A flat was longer, but also less legato and less loud, in performances of sadness. The tone G was performed louder and somewhat longer in tenderness, and B-natural was stressed by loudness in short duration.

These results on contrasting expressive intentions (happy vs. sad and tender vs.angry) revealed that certain notes assumed to be important for expression of a certain emotion were stressed by increase in loudness, articulation or duration, but not in performances intended to express the opposite emotion. These results are in agreement with the hypotheses (b), (c), (d) and (e).

Experiment 2: Listening test. The performances by the two musicians were rated by musicians and non-musicians in listening experiments in order to see whether the expressive intentions by the performers were accurately communicated to the listeners. A four-way ANOVA was conducted on the musicians' judgments on each rating scale (happy-sad and tender-angry), separately for each performer (A and B). The results for psychology students were not shown but looked very similar to musicians' ratings. The results would also allow some conclusions regarding the ability of the musicians to compensate for a less adequate version by means of variables in the performance.

In the ratings of happiness and sadness it was fairly apparent that the expression inherent in the melodic structure affected listeners' judgments most, followed by performance intentions and position (good and less good versions). The ratings of tenderness and anger revealed that inherent expression and performance intention had about equal impact for Performer A, whereas performance intention had the strongest impact for performer B. The effects of position (good and less good versions) were blurred for both performers. Non-significant results regarding repetition revealed that there were no differences between versions of the same emotion (A, LA, T and LT versions) taken from different progressions. Differences in results between the two performers suggest that they differed in their expressive ability. Musicians' ability to adapt their performance to the musical structure may be an important aspect of emotional communication of music.

3.2.4 Conclusions

It seemed appropriate to study the interplay between melodic structure and performance for emotional expression in music. The fact that performers adjusted the average levels of loudness depending on the expressive content in melodic versions was important evidence that musicians took the musical structure and its emotional impact into account during their expressive performance. The second finding was that certain notes were stressed for certain emotional expressions. This means that musicians were not only sensitive to the general emotional impact of the piece, but also to the expressive content attached to different tones. The listening experiment showed, on the whole that listeners responded to the recordings in a way consistent with the idea that the expression inherent in the melodic structure was taken into account (compensated for) in the performance.

The outcome of this paper may have many implications. First of all, expression inherent in the melodic structure seemed to have been taken into account by the musicians. This means that studies on music and emotion should consider the interplay between “what to play” and “how to play”. The note by note inspection further corroborated findings in Study I in so far that musicians emphasized tones of assumed importance for the emotional expression. This means that a dynamic view could be applied to the perception of music, in which tonal tendencies are activated both by melodic structure but also by the musicians’ performance.

3.3 Study III

The contribution of immanent and performed accents to emotional expression in short tone sequences.

3.3.1 Aim and background

Study I showed that interactions between musical factors such as mode, harmonic progression, rhythm and contour could partly be related to the occurrence of immanent accents in the melodic structure and emphasis on notes assumed to be important for perception of structure and emotional expression (i.e., instability, tension). The results of Study II furthermore indicated that musicians seemed to emphasize these notes in their live performance. The aim of Study III was hence to more systematically investigate different kinds of accents on selected target notes in short tone sequences. A first experiment aimed at investigating immanent accents (melodic, metric, and rhythmic accents) and a second experiment at investigating performed accents (using loudness, articulation and timing).

The hypotheses were: (a) Generally, the notes of a melody are not all of equal importance for expression of emotion; certain notes are more important than others as suggested in the following hypotheses; (b) The implicit harmonic function of certain notes may be important for perceived emotional expression; (c) Tense or unstable notes may be more apt for expression of negative emotion such as anger, and (d) relaxed or stable notes more apt for expression of positive emotion such as tenderness. Furthermore, (e) happiness will be better represented by a note important in major mode, and (f) sadness by notes important in minor mode. Finally, (g) if notes assumed to be more important for a certain expression are marked for consciousness by either immanent or performed accents, this will increase listeners' perception of the expression in question.

3.3.2 Methods

Participants. In the first experiment, eleven non-musicians (psychology students) 7 females and 4 males, 20 -34 years old, volunteered to participate. In experiment 2, eighteen non-musicians (psychology students) 14 females, 4 males, 20-41 years old volunteered to participate.

Musical material. The musical stimuli were created in both C major and C minor mode contexts. Each sequence started with a major or minor chord in half notes, followed by three melody notes. Two settings of notes were selected to each context; C – F – A, and B natural – D – F in major mode; D – F – A flat, and E flat – G – B-natural in minor mode.

The melody notes were chosen in a systematic way. The last high-pitched note in the three note sequences was of special interest for emotional expressions and was called *the target-note*. The note A in progression 1 may be prominent for positive emotions like happiness, because it is the major third (the best note to signal major mode) in an implied subdominant. In progression 2 the target-note F should have positive connotations since it does not stick out and may be perceived as stable in the tonal context with its dominant-tonic progression. In progression 3, the target-note A-flat is a prominent signifier of minor mode, which is usually associated with sadness. In progression 4, finally, the target-note B-natural is unstable and tense because it does not fit in the minor tonic chord implied by the preceding notes (E-flat and G) and therefore probably induces a sense of negative emotion (anger).

Each three-note progression was manipulated to represent all possible permutations with regard to pitch: pitch contours 123, 132, 231, 213, 312, 321, in which 1= lowest and 3 = highest pitch. This manipulation results in three descending and three ascending melodic versions in which each of them contains straight, convex, and concave melodic contours. The target-note was always the highest pitch level (3). The versions were furthermore varied in rhythm: long-short-short (2:1:1 ratio), short-long-short (1:2:1 ratio) and short-short-long (1:1:2 ratio). The target-note in each progression appears sometimes on notes with metric and/or melodic accents. The two-measure sequences were repeated four times ending on the context chord in whole notes. This melodic material, reproduced by computer in deadpan performance, was investigated in its completeness in experiment 1.

In experiment 2 a sample of versions was selected, created out of the two melody versions 123 and 132 with the longer (half note) always as the first note. These versions were applied to three progressions: progressions 1, 3 and 4. The choice of just a few melody versions was made in order to enable rich variation of performance factors but still allow comparison between two melodies.

Half of the listeners were given melodic versions with the target notes lowered one octave to avoid confounding with the otherwise always high-pitched target note. It should be noted that this also means that the melodic contours were different between the two groups.

Performed accents (in Experiment 2) altered between notes in the middle position and the last position (phrase ending) of the three-note sequences. Since the target notes also altered between these positions, this means that a performed accent was supplied to the more interesting target note half of the times, in different positions and pitch levels.

The different types of performed accent were (a) accent by increased loudness, (b) increased loudness and legato, (c) increased loudness and staccato articulation, d) increased loudness, legato and delayed tone onset, and e) increased loudness, staccato articulation and delayed tone onset. That is, performed accents were made by increasing MIDI velocity (loudness) in combination with different articulations (legato, staccato) and delayed tone onset, which makes it possible to study the impact of so-called joint accents. That is, the target note could be in position with melodic accent (pivot note) and stressed in loudness, articulated by either staccato or legato, and furthermore stressed by delayed onset.

Procedure. The sequences were programmed ”dead-pan” in MIDI (using MIDI-sequencer) in a tempo of 120 bpm (beat = quarter-note) and all the tones were given the same loudness and articulation (except for the performed accents in Experiment 2). In both listening experiments the midi-sequences were played at comfortable sound level using a *Yamaha Disclavier* Grand Piano in a lecture hall. Accents (Experiment 2) were achieved by increasing loudness to velocity 75 (from 60), by setting the onset-offset duration to 50 % (instead of 90%) of the inter-onset interval (staccato) or to 125% (legatissimo). Onset delay meant 30 ms delay of the onset of the accented note.

The participants performed the experiments divided in smaller groups. In Experiment 1 the participants rated all 72 sequences on two bipolar scales, *happy-sad* and *tender-angry*, each comprising seven steps. However, in Experiment 2, the three progressions constituted blocks, the order of which was rotated between listening sessions (ABC, BCA and CAB) The musical sequences were furthermore randomized for each order of blocks. The listeners made their judgments on a single seven-step scale, one for each block. That is, progression 1 (block A) judgments were made for happiness, progression 3 (block B) for sadness, and progression 4 (block C) for anger.

3.3.3 Results

Experiment 1. An ANOVA was conducted for the 4 x 3 x 2 x 3 mixed factorial within subjects design, involving 4 levels of tonal progression, 3 levels of melodic contour (straight, convex, concave), 2 levels of melodic direction (ascending and descending), and 3 rhythm patterns (2:1:1, 1:2:1, and 1:1:2 ratios). The result was very similar to that in Study I. There were significant main effects of progression in both scales, $F(3,30) = 45.09$; $p <.001$, and $F(3,30) = 5.14$; $p <.01$, of rhythm on the tender – angry scale, $F(2,20)= 3.91$; $p <.05$, and of melodic direction on the happy – sad scale $F(1,10)= 6.18$; $p <.05$. All together main effects showed that progression 1 was perceived as the most happy, short-short-long rhythm patterns were perceived as the most angry rhythm pattern, and descending direction was perceived as more sad than ascending direction.

There were also a few significant two-way and three-way interactions in the happy – sad scale, and even a significant four-way interaction, $F(12,120)= 2.00$; $p <.05$, which shows how intricate the four variables may interact, even with very simple musical stimuli as in this experiment. Rather than inspecting higher order interactions it was more informative to apply contrast analysis as is in Study I. This comparison was restricted to the progression that was judged most representative for the respective emotion. (Tenderness and progression 2 were excluded because of weak and fuzzy effects in mean ratings).

Largest contrasts in progressions. The most interesting note for happiness in progression 1 (major mode) was A (as discussed earlier). The results showed that this note was emphasized by melodic, rhythmic and metric accents in versions judged as happy (GH) (Figure 2, left), whereas in versions rated ”less good as happy” (LGH) this principle seemed not to be present (Figure 2, right). Confidence intervals (95%) indicated but small overlaps between GH versions and LGH versions.

In GH versions 2a and 2b, A receives metric accent. In versions 2b and 2c A obtains melodic accent as pivot note (on the turn), following after an upward jump from C. In version 2c, there is, furthermore, rhythmic accent and syncopation. In LGH versions (2d-f), A appears on a metrically unaccented beat in 2d and 2f (however, it carries melodic accents in 2d (pivot) and 2f (jump). In 2e, however, A receives metric accent, but its effect may be reduced by the strong rhythmic accent and syncopation on the following F.

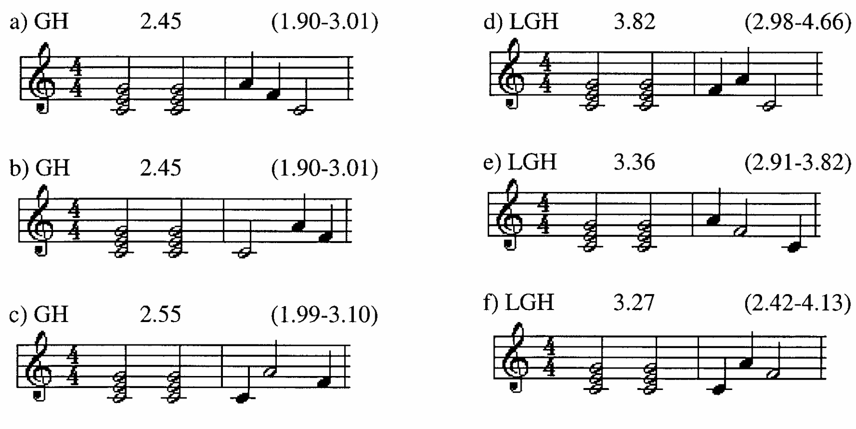


Figure 2. Large contrast effects in means for happiness. 95% CI specified in brackets. Good as happy version (GH) a-c and less good as happy version (LGH) d-f.

Similar results were found for A-flat (a note supporting the minor scale in progression 3) in judgments of sadness and for B natural (a tense note in progression 4) in judgments of anger.

Correlation analysis. Correlation analyses were used to more thoroughly examine the hypothesis that accents on target notes will increase listeners' ratings of emotional expression. A variable called *sum-on-target* was created in which different immanent accents — metric (on first and third beats), pivot, jump, duration and syncopation — were allocated scores, either 1 (one) if they appeared on target-note or 0 (zero) if not, see Table 2. The rightmost column shows the sum of these scores. It can vary between 0 (zero), no accent at all on the target-note, and 3, that is, cases with various forms of *joint accent* (e.g., metric, melodic and duration). These scores (sums) were then analyzed for their correlation with the mean ratings for each of the 18 melodic versions within each progression, one analysis for each judgment scale, happy-sad and tender-angry.

Table 2. *Allocation of scores for different accents on on-target note over contour versions added to a sum in the right most column (Examples shown only for versions with long-short-short rhythm, n=6)*

Contour version	Metric first beat	Metric third beat	Pivot	Jump	Dura-tion	Syn-cope	Sum on target
123	0	0	0	0	0	0	0
321	1	0	0	0	1	0	2
132	0	1	1	1	0	0	3
231	0	1	1	0	0	0	2
213	0	0	0	1	0	0	1
312	1	0	0	0	1	0	2

For comparison, analogous correlations were also calculated regarding off-target-notes. A new variable called sum-off-target was created, in which the different immanent accents were allocated scores according to the principles described above. High correlations between sum-on-target scores and mean ratings would provide support for the hypothesis, whereas the opposite would be the case if there were high correlation with sum-off-target scores.

The correlations were relatively low across all conditions and scales. However, the highest correlations, positive or negative, appeared in correlations with on-target notes (r = -.25 to +.29) compared to correlations with off-target notes (r = -.16 to +.14). This gives a certain support to the hypothesis that the target notes were superior to induce perceived expression. The variable ”sum on-target” had the strongest correlation for B-natural, relating to ratings of anger (r =.29) as predicted in the earlier discussion of largest contrast effects. However, it can be seen that accent on B-natural also may be related to ratings of sadness (r = .25). Corroborating assumptions in earlier discussions, the target note A in progression 1 showed a correlation with ratings of happiness (r = -.25), but not with tenderness (r = -.08). The results concerning the target-note A flat did not conform to the predictions for any of the judgment scales (r = -.07 in the happy - sad and r =.09 in the tender-angry scale). Since all correlations above are relatively low and not statistically significant, the results should be regarded as suggestions for further study.

Experiment 2. The design was a 3 x 2 x 5 x 2 x 2 mixed split-plot factorial design. Within-subject factors were three progressions in blocks (happy, sad, angry), two levels of ”target note position”, five levels of ”type of accent” (loud, loud+legato, loud+staccato, loud+legato+delay, loud+staccato+delay), and two levels of ”performed accent ” (on target note, on other note). The remaining factor, pitch (high, low), was a between-subjects factor meaning that the 18 participants were divided into six groups across three orders of blocks and two levels of pitch. The dependent variable was a (unipolar) scale comprising seven steps (from 0 to 6), one in each block for happy, sad, and angry, respectively.

The listeners' mean ratings were subjected to multiple regression analyses, one for each block. The predictor variables were dummy coded (one-zero). Target note position on the fourth beat was coded as 1, on the third beat as 0. The type-of-accent factor was divided into two variables: (a) articulation, staccato = 1, legato = 0, and (b) timing, with delay =1, without delay = 0. Performed accent on the target note was coded = 1, on another note = 0; pitch level was coded as high pitch = 1, low pitch = 0.

Results and discussion. About 2/3 of the variance (R^2 = .69) in the ratings of anger could be accounted for by a linear regression model, , about 1/3 (R^2 = .36) in the ratings of sadness and about 1/2 (R^2 = .51) in the ratings of happiness. The hypothesis was that ratings of emotional expression would increase when performed accents occurred on target notes assumed to be important for the expression in question. With regard to anger, performed accent on the target note B-natural was strongly significant (*Beta* = .71, $p<.001$). There were in general no significant effects of stress on the target note A flat in ratings of sadness. However, inspection of the data in versions with the target note on the third beat (middle) position revealed that a performed accent on the target-note had effect, $F(1,12) = 4.78$, $p<.05$ (calculated in General Linear Model, repeated measures). Figure 3 shows this additive effect of stress on the target note but also shows that loudness plus legato and delay contributed most to the perception of sadness. This result corroborates the suggestion obtained from the study of large contrast effects that emphasis on A flat may not benefit from all types of accent.

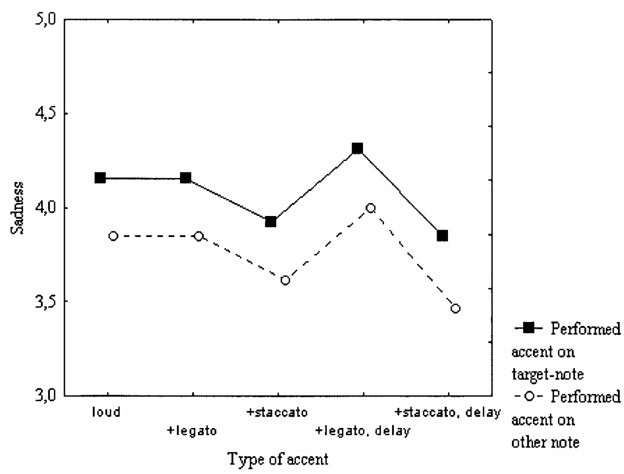


Figure 3. Mean ratings for sadness. Performed accent on target-note (solid line), on another note (dashed line), for five types of accents. (Note that these results apply only to target note on the third beat).

High pitch was favorable for expression of anger (*Beta* = .42, p<.001) a finding in accordance with earlier results (e.g., Scherer & Oshinsky, 1977), whereas low pitch was preferable for sadness (*Beta* = -.39, *p*<.01), a common finding in earlier research (Gabrielsson & Lindström, 2001). A significant effect of target note position (*Beta* = -.29, *p*<.05) that target note on the third beat was superior to target note on the fourth beat. Significant effect of articulation (*Beta* = .31, *p*<. 05) meant that legato articulation increased perception of sadness. The only significant factor regarding happiness was pitch, (*Beta* = .66, *p*<.001), indicating that high pitch increased perception of happiness. These findings are in line with earlier empirical findings (Gabrielsson & Lindström, 2001).

3.3.4 Conclusions

In accordance with the general hypothesis (see hypothesis (a) in Introduction) the results indicated that all notes in a melodic structure are not of equal importance for emotional expression. Assumed target notes affected perceived emotional expression when they were stressed by immanent and performed accents, thus supporting hypothesis (g). Hypothesis (b) concerning the importance of implicit harmonic function of certain notes (target-notes) also received support, more precisely in that happiness was emphasized by stressing an important note in major mode, sadness by stressing an important note in minor mode, and anger by stressing an unexpected and tense note, thus supporting most of the remaining hypotheses (hypotheses (c), (e), and (f), however not hypothesis (d) concerning perceived tenderness.

In general, the effects found in the two experiments were usually small, which may be expected since three-note patterns represent very impoverished musical stimuli. On the other hand, even this very simple musical material provided evidence for the interplay between musical features.

There are numerous different ways to provide emphasis on focal notes in music, both inherent in the composed score and controlled by the performer. Some data in this study suggest that emphasis on assumed important notes may not be achieved by all types of accent. Further experiments are needed to investigate possible optimum arrangements of accents on specific notes to bring about intended emotional expressions.

4 General discussion

4.1 Conclusions

The aim of this thesis was to elaborate some aspects of a dynamic view of melodic organization and performance, which may have implications for perception of musical structure and, especially, for emotional expression in music. This idea arose from discussions in literature of music analysis and music cognition concerning positions and functions of single notes within the tonal hierarchy and related concepts as chords and harmonic progressions. To appreciate the relative importance or functions of single notes is thus not a novel view, but this has been limited to their role in the musical structure. I have argued here that certain notes in a given tonal context may be called ”expressive notes” in the sense that they carry information about the emotional expression in music. In fact common usage of terms as tension and resolution/relaxation, or tense and relaxed notes, in discussions on musical structure also hints at emotional aspects, but such aspects have usually not been considered or elaborated.

The models of tonal hierarchy have furthermore been depicted as static. However, some empirical studies have questioned this approach by demonstrating examples of dynamic activation of certain notes due to various factors in the melodic and rhythmic structure. Activation of such notes—expressive notes—by immanent and performed accents may have consequences for the perceived emotional expression.

To explore these questions three studies were conducted. Study I showed evidence that complex interactions involving harmonic progression, rhythm, melodic contour and melodic direction could be linked to how certain notes received emphasis through accent structures. For instance, when stable notes (tonic and major/ minor thirds) were emphasized by metric, melodic and rhythmic accents this affected listeners' perception of stability and tenderness, whereas stress on unstable tones, passing tones, and tones unrelated to implied harmony led to judgments of instability and anger. Emphasis on notes important for identification of the major mode (e.g., major third) increased perceived happiness, and emphasis on notes defining the minor mode (e.g., minor third) increased perceived sadness. These interactions effects were however at times very small and the interpretations were only tentative. Therefore the question was whether these effects could contribute to explain variance in listeners' ratings of both structure (instability, complexity and tension) and emotion (happiness, sadness, tenderness and anger) in comparison with well-established musical factors such as mode, harmonic progression, rhythm and melodic contour and musical training. To investigate this, predictor variables representing joint accents (metric, melodic, duration) on stable (triadic) notes or on unstable (non-triadic) notes were included, together with the well-established factors just mentioned, in stepwise multiple regression analyses and contributed significantly to all structural and emotional scales. The analyses further indicated that the level of joint metric, melodic and rhythmic accents had an effect.

A further question was to what extent performers may modify their performance dependent on the inherent emotional expression in melodies, for instance if a "less happy" melody could be made to sound more happy using means available in the live performance. A related question was if potential expressive notes would be emphasized by accents achieved in performance. Study II showed that for a large majority of performances musicians adjusted the average loudness to compensate for versions that were less adequate for the intended expression. They also used articulation for the same purpose, although not to the same extent, but not tempo. It also seemed that the musicians paid special attention to notes considered to be important for the emotional expression. Note-by-note inspection of performances of contrasting expressive intentions (happy vs. sad and tender vs. angry) revealed that certain notes were stressed by increase in dynamics, articulation or duration, but not in performances intended to express the opposite emotion.

Study I and II have hence showed that not only immanent accents but also performed accents were crucial in emphasizing certain essential notes. To what extent are different immanent accents in melodic structure (duration, melodic pivot metric etc) and different performed accents (loudness, staccato, legato, and delay) efficient in emphasizing expressive notes? These questions were further investigated in Study III using systematic manipulations of short tone sequences as well as quasi-performances of the sequences. The results showed that certain selected notes (so-called target-notes) seemed to be more important for perceived emotional expression than other notes, which could be explained by reference to the target notes' relation to implied harmonic functions. The more types of immanent accents that appeared on a major third (note A), the major mode in an implied subdominant chord, the higher was the correlation with perceived happiness. A flat, an essential note in C minor mode, seemed important for perceived sadness when stressed by performed accents in certain melodic versions. The unstable and tense B natural was associated with negative emotion (anger), especially when it was emphasized by performed accents. However, a further conclusion from Study III was that important notes may not benefit from all types of accent.

Altogether it can be concluded that certain notes seem more important than others for perceived emotional expression, which is in accordance with the most general hypothesis in this thesis. This may to some extent explain, for instance, why a piece of music in minor mode which usually would be perceived as sad, sometimes may be perceived as happy, as pointed out by Hevner (1935). That is, the notes central in minor mode (hence also for sadness) may be suppressed in certain rhythmic and melodic conditions. Earlier findings concerning melodic contour and direction (e.g., Hevner, 1936; Rigg 1939, 1964; Scherer & Oshinsky, 1977; Cooke, 1959; Gabriel, 1978) and rhythm (e.g. Gundlach, 1935; Hevner, 1936; Watson, 1942; Thompson & Robitaille, 1992) may also be better understood by further analysis of immanent and performed accents in music and their relationship to potentially important notes.

4.2 Methodological issues

Earlier empirical studies have either used existing pieces of music or applied systematic manipulation of selected musical factors (Gabrielsson & Lindström, 2001). Using existing pieces of music in real performance provides ecological validity, but the conclusions concerning relationships between musical factors and emotional expression remain less definite than when experimental manipulation is applied. Studies I and II used systematic manipulations of a simple tune and Study III of short three-note patterns. The effects of the manipulations were at times small, which may be expected since all melodic patterns represented impoverished musical stimuli especially in Study III. On the other hand, even these very simple musical stimuli provided evidence for interactions among musical features.

The factorial designs used in these studies allowed study of interrelationships among different variables in the musical structure. Most important was to achieve a rich tonal material, accomplished by different tonal progressions, and variation in the accent structure accomplished by rhythm and melodic contour, thereby yielding dynamic influences on melody notes with assumed expressive characteristics.

However, even the most serious attempts at systematization in musical contexts are vulnerable to confounding factors. First, it is impossible to change the latent harmony in a monophonic tune without changing pitch levels. Thus, the different prototypical versions of Frère Jacques (Study I) and three-note patterns (Study III) were by necessity (slightly) varied in average pitch height and range, which may have expressive effects even though changes were set to a minimum. Furthermore, the prototypical melodic versions have their implicit harmonic functions, but these may vary as function of serial orderings of notes, and of rhythmic and melodic stress as used in the manipulations of each progression. Thus, it is often not possible to keep the harmonic functions *per se* constant in manipulations of monophonic pieces.

Both non-musicians and musicians were used as listeners. Researchers in music psychology often employ musicians as subjects, especially to test hypotheses from music theory concerning perceptual and cognitive abilities and processes. Smith (1997) has expressed some concern that musical novices receive far less empirical and theoretical attention, and Bigand (1997) noted differences between musicians and non-musicians. Anyhow, a study of musical expression should preferably also include musically naive listeners to increase ecological validity. The results in Studies I and II in Study I and II showed only minor differences between the listener groups, musicians versus non-musicians. In Study III only non-musicians were used as listeners.

4.3 Implications for future research

There are many factors affecting perceived emotion in music (Gabrielsson & Lindström, 2001; Gabrielsson & Juslin, 2003) which have got considerable attention in earlier research. This thesis adds still another factor, that is, the effects that may be achieved by immanent and performed accents on notes assumed important for perceived emotional expression.

The present thesis also provides interesting challenges for future research. One topic would be to figure out possible optimum arrangement of accents for different emotional expressions. There are numerous different ways to provide emphasis on focal notes in music, both inherent in the composed score and controlled by the performer. Maybe musicians and composers learn how to achieve certain sophisticated types of accents to signal notes of importance for perceived emotion. With regard to composed music this may be investigated by statistical analysis of such music stored in databases, for instance, to check whether the type of certain notes that seemed important for emotional expression in the present thesis appear in musical contexts intended or supposed to express a certain emotion.

4.4 Final remarks

In this thesis it was claimed that there is interdependency between musical features and that focus on interaction and interplay between factors is fruitful for the understanding of emotional expression in music. Three studies shed some light on dynamic aspects of melodic organization and performance which not only may have implications for music and emotion but also for novel discussions of melodic organization and tonal perception. These insights may attract the attention of music theorists, musicologists, music psychologists as well as of composers and musicians.

5 References

Bengtsson, I. (1974). Empirische Rhythmusforschung in Uppsala. *Hamburger Jahrbuch für Musikwissenschaft, 1,* 195-219.

Bengtsson, I., & Gabrielsson, A. (1980). Methods for analyzing performance of musical rhythm. *Scandinavian Journal of Psychology, 21,* 257-268.

Bengtsson, I., & Gabrielsson, A. (1983). Analysis and synthesis of musical rhythm. In J. Sundberg (Ed.), *Studies of music performance* (pp. 27-59). Stockholm: Publications issued by the Royal Swedish Academy of Music, No. 39.

Bengtsson, I., Gabrielsson, A., & Thorsén, S. M. (1969). Empirisk rytmforskning (Empirical rhythm research). *Swedish Journal of Musicology, 51*, 49-118.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16*, 485-518.

Bigand, E. (1997). Perceiving musical stability: The effect of tonal structure, rhythm and musical expertise. *Journal of Experimental Psychology: Human Perception & Performance, 23*, 808-822.

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

Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics, 53,* 585-600.

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

Butler, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception, 5,* 219-242.

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

Clarke, E. F. (1988). Generative principles in music performance. In J. A Sloboda (Ed.), Generative processes in music. *The psychology of performance, improvisation, and composition* (pp. 1-26). Oxford: Clarendon Press.

Clynes, M. (1977). *Sentics: The touch of emotions*. New York: Doubleday.

Cook, N. (1994). *A guide to musical analysis*. New York: Oxford University Press.

Cook, N., & Dibben, N. (2001). Musicological approaches to emotion. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research.* New York: Oxford University Press.

Cooke, D. (1959). *The language of music*. London: Oxford University Press.

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

Cuddy, L., Cohen, A., & Mewhort, D. (1981). Structure in melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance, 7,* 869-883.

Davies, S. (1994). Musical meaning and expression. New York: Cornell University Press.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music-Perception, 10*(3): 343-378.

DePoli, G., Roda, A., & Vidolin, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performances, *Journal of New Music Research, 27,* 293-321.

Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharucha, and Krumhansl. *Journal of Experimental Psychology: General, 113*, 413-416.

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

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. Academic press, Orlanda, Florida.

Ekman, P. (Ed.) (1973). *Darwin and facial expression.* New York: Academic Press.

Friberg, A. & Battel, G. U. (2002). Structural communication. In R. Parncutt, & G. E. McPherson (Eds.), *The science and Psychology of Music Performance.* New York: Oxford University Press.

Gabriel, C. (1978). An experimental study of Deryck Cooke's theory of music and meaning. *Psychology of Music, 6,* 13-20.

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

Gabrielsson, A. (1988). Timing in music performance and its relations to music experience. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition.* (pp. 27-51). Oxford: Clarendon Press.

Gabrielsson, A. (1994). Intention and emotional expression in music performance. In A. Friberg, J. Iwarsson, E. Jansson, and J. Sundberg (Eds.), *Proceedings of the Stockholm Music Acoustics Conference, July 1993* (pp. 108-111). Publications issued by the Royal Swedish Academy of Music, No. 79.

Gabrielsson, A. (1995). Expressive intention and performance. In R. Steinberg (Ed.), *Music and the mind machine. Psychophysiology and psychopathology of the sense of music* (pp. 35-47). Berlin: Springer-Verlag.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music.* (2nd ed.). (pp. 501-602). San Diego: Academic Press.

Gabrielsson, A. (2002). Emotion perceived and emotion felt: same or different? *Music Scientiae, special issue 2001-2002.,* 123-147.

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music, 31*, 221-272.

Gabrielsson, A., & Lindström, E. (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology, 14*, 94-116.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotion. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research.* New York: Oxford University Press.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music, 24*, 68-91.

Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In H.H. Goldsmith, R. J. Davidsson, & K.R. Scherer (Eds.), *Handbook of affective sciences.* New York: Oxford University Press.

Graybill, R. (1989). Phenomenal accent and meter in the species exercise. *In Theory Only, 11,* 11-43.
Gundlach, R.H. (1932). A quantitative analysis of Indian music. *The American Journal of Psychology, 44,* 133 - 145.
Gundlach, R.H. (1935). Factors determining the characterization of musical phrases. *American Journal of Psychology, 47,* 624-644.
Hanslick, E. (1891). *The beautiful in music*. London: Novello.
Hevner, K. (1935). Expression in music. A discussion of experimental studies and theories. *Psychological Review, 42*, 186-204.
Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48*, 246-268.
Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology, 49,* 621-630.
Huron, D., & Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. *Music Perception, 13*, 489-516.
Izard, C. E. (1972). *The face of emotion.* New York: Appleton-Century-Crofts.
Jansen, E. L. & Povel, D. J. (1999). Mechanisms in the perception of accented melodic sequences. *Proceedings of the 1999 Conference of the Society for Music Perception and Cognition.* Evanston, III., p.29.
Juslin, P. N. (1997a). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception, 14,* 383-418.
Juslin, P. N. (1997b). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae, 1,* 225-256.
Juslin, P. N. (1998). A functionalist perspective on emotional communication in music performance. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 78.* Uppsala, Sweden: Uppsala University Library.
Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research*. New York: Oxford University Press.
Juslin, P. N., & Laukka, P. (2003). Communication of emotion in vocal expression and music performance: Different channels, same code? *Psychological bulletin. 129,* 770-814.
Juslin, P. N., & Sloboda, J. A. (2001). Music and emotion: introduction. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research.* New York: Oxford University Press.
Jones, M. R. (1987). Dynamic patterns structure in music: Recent theory and research. *Perception and Psychophysics, 41,* 621-634.
Jones, M. R., & Boltz, M. (1989). Dynamic attending and responding to time. *Psychological Review, 96*, 459-491.
Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Coginitive Psychology, 11*, 346-374.
Krumhansl, C. L. (1990). *The cognitive foundations of musical pitch.* New York: Academic.
Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334-368.
Kivy, P. (1980). *The corded shell.* Princeton, N. J.: Princeton University Press.
Laden, B. (1994). Melodic anchoring and tone duration. *Music Perception, 12*, 199-212.

Langer, S. K. (1957). Philosophy in a new key. (3rd ed.). Cambridge, Massachusetts: Harvard University Press.
Lerdahl, F. (1988). Tonal pitch space. *Music Perception, 5*, 315-350.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA:MIT Press.
Meyer, L.B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago press.
Narmour, E. (1990). *The analysis and cognition of basic melodic structures*. Chicago: University of Chicago press.
Nielsen, F.V. (1983). *Oplevelse af musikalsk spandning. (The experience of musical tension)*. Copenhagen: Akademisk forlag.
Nielzén, S., & Cesarec, Z. (1982). Emotional experience of music as a function of musical structure. *Psychology of Music, 10*, 7-17.
Parncutt, R. (in press). Accents and expression in piano performance. *Systemische Musikwissenschaft. Festschrift Jobst P. Fricke zum 65. Geburtsdag*, Frankfurt: Peter Lang.
Patel, A. D., & Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. In I. Deliège & J. A. Sloboda (Eds.), *Perception and cognition of music.* Hove: Psychology Press.
Peretz, I. (2001). Listen to the brain: A biological perspective on musical emotions. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research.* New York: Oxford University Press.
Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68*, 111-141.
Povel, D. J., & Egmond, R. (1993). The function of accompanying chords in recognition of melodic fragments. *Music Perception, 2,* 101-115.
Rigg, M. G. (1939). What features of a musical phrase have emotional suggestiveness? *Publications of the Social Science Research Council of the Oklahoma Agricultural and Mechanical College, No. 1*.
Rigg, M. G. (1964). The mood effects of music: a comparison of data from former investigations. *Journal of Psychology, 58,* 427-438.
Schellenberg, E. G., Krysciak, A. M.., & Campell, R. J. (2000). Perceiving emotion in melody: Interactive effects of pitch and rhythm. *Music Perception, 18*, 155-171.
Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion, 1 (4),* 331-346.
Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*, 109-150.
Schmuckler, M. A. (1990). The performance of global expectations. *Psychomusicology, 9*, 122-147.
Seashore, C. E. (1938). *Psychology of music*. New York: McGraw-Hill. (Reprinted 1967 by Dover Publications, New York.)
Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music.* Oxford: Oxford University Press.
Smith, J. D. (1997). The place of musical novices in music science. *Music Perception, 14,* 227-262.
Sundberg, J., Friberg, A., & Frydén, L. (1991). Common secrets of musicians and listeners: An analysis-by-synthesis study of musical performance. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical structure* (pp. 161-197). London: Academic Press.

Thomassen, J. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America, 71*, 1596-1605.

Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Art, 10,* 79-89.

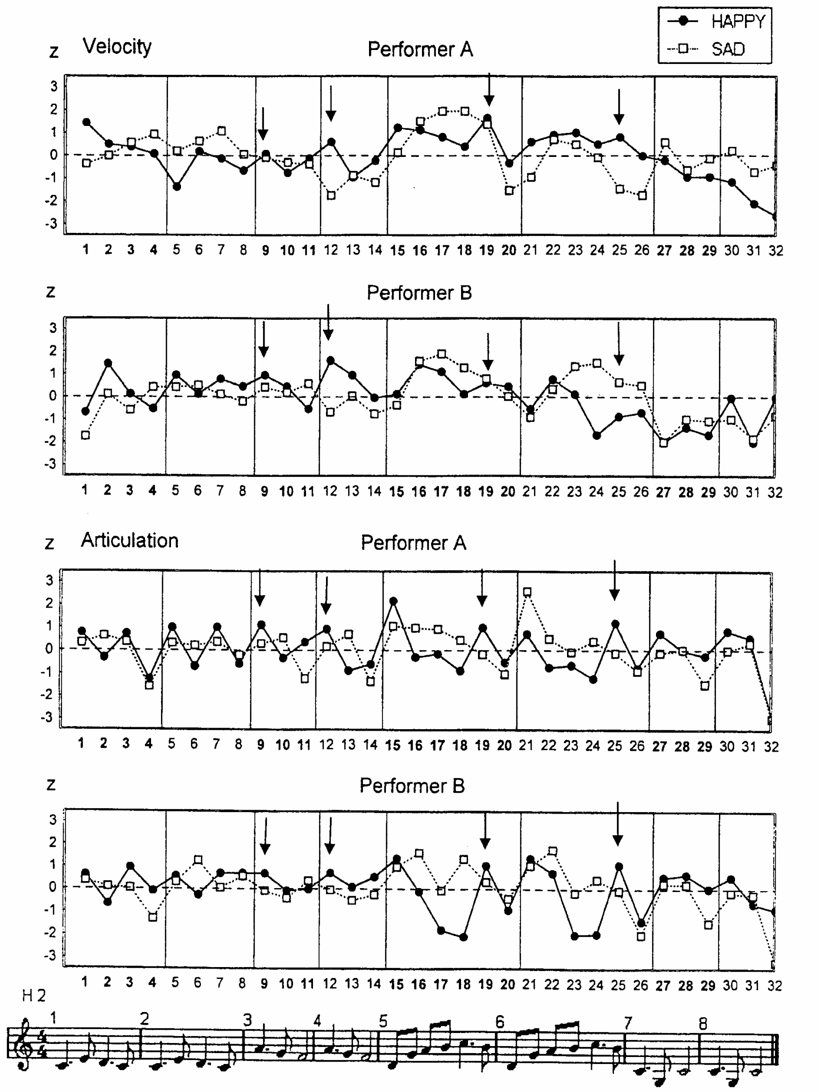
Tomkins, S. (1962). *Affect, imagery, and consciousness. Vol. 1: The positive affects.* New York: Springer.

Watson, K. B. (1942). The nature and measurement of musical meaning. *Psychological Monographs, General and Applied, 54*, 1- 43.

Wedin, L. (1969). Dimension analysis of emotional expression in music. *Swedish Journal of Musicology, 16,* 118-140.

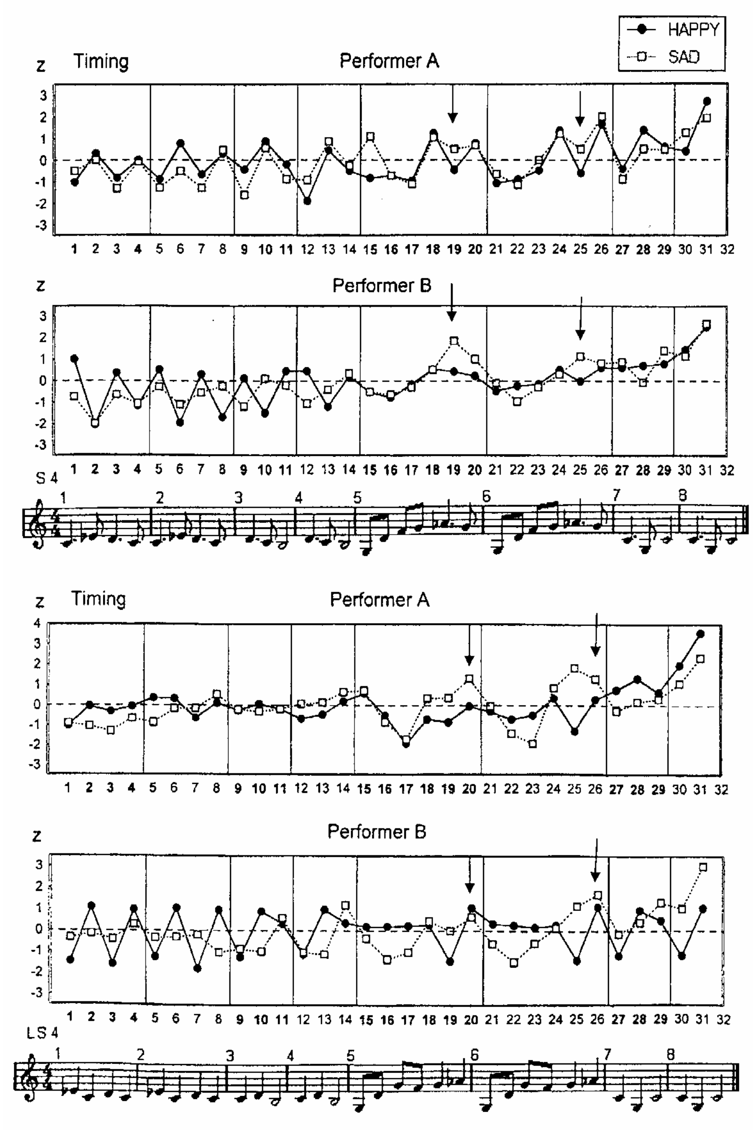
Wedin, L. (1972). A multi-dimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology, 13,* 241-257.

Appendix A



Appendix A. Note by note variation, expressed in z-values, in happy and sad performances, of the H2 version for both performers in velocity (upper panels) and articulation (lower panels)

Appendix B



Appendix B. Note by note variation, expressed in z-values, in happy and sad performances, for both performers in timing of S4 version (upper panels) and LS4 version (lower panels)

Acknowledgements

There have been many lucky circumstances that have given me the opportunity to develop this thesis. I have been surrounded by nice people and have been employed in three research projects during my years in the department. I would like to express my gratitude and special thanks to my patient and always supportive supervisor Alf Gabrielsson. His ability to achieve a mix and balance between tutorial and careful listening has been successful for my work. Patrik Juslin has made my work and time at the department enjoyable. He is a good friend and a good project leader and has given me decisive support at several moments. I would also like to thank all other colleagues I have had opportunity to cooperate with: The expressive performance group (i.e. Gertrud, Guy, Ingrid, Maria, Marie), the FeelMe group (Anders, Erwin, Jessika, Petri, Roberto) and MEGA group (Antonio, Gualtiero).

The research of this thesis has been supported by the Swedish Tercentenary Foundation through a grant to Alf Gabrielsson, and through a grant to Patrik Juslin, as well as by the EU-IST-funded MEGA project through a grant to Antonio Camurri.

I would like to thank Erik Börjesson for careful reading of the preliminary manuscript, Siv Wedung for support in the library, Peter Hammarlund for giving me all I needed and Elsa Sjöberg and Rosemari Finn for economical administration, Peter Thunberg for computer assistance and finally Mildred Larsson for reminding me of everything. Thanks also to my parents for encouraging me once for higher education and for always being at hand. Last and with affection I would like to thank my lovely family: Anette, Linus and Elin, for being the essentials in my life.

Erik Lindström
March 2004

Acta Universitatis Upsaliensis

Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences

Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is either a monograph or, as in this case, a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences*. (Prior to July 1985, the series was published under the title "Abstracts of Uppsala Dissertations from the Faculty of Social Sciences".)

Distributor:
Uppsala University Library,
Box 510, SE-751 20 Uppsala, Sweden
www.uu.se, acta@ub.uu.se

ISSN 0282-7492
ISBN 91-554-5953-6