Structural Communication

Anders Friberg and Giovanni Umberto Battel

Abstract

The communication of structure in musical expression has been studied scientifically by analyzing variations in timing and dynamics in expert performances, and by analysis by synthesis. The underlying principles have been extracted, and models of the relationship between expression and musical structure formulated. For example, a musical phrase tends to speed up and get louder at the start, and to slow down and get quieter at the end; mathematical models of these variations can enhance the quality of synthesized performances. We overview the dependence of timing and dynamics on tempo, phrasing, harmonic and melodic tension, repetitive patterns and grooves, articulation, accents, and ensemble timing. Principles of structural communication (expression) can be taught analytically, by explaining the underlying principles and techniques with computer-generated demonstrations; or in traditional classroom or lesson settings, by live demonstration.

Introduction

Variations in timing and dynamics play an essential role in music performance. This is easily shown by having a computer perform a classical piece exactly as written in the score. The result is dull and will probably not affect us in any positive manner, although there may be plenty of potentially beautiful passages in the score. A musician can, by changing the performance of a piece, totally change its emotional character, e.g. from "sad" to "happy" (Juslin and Persson, this volume). How is this possible, and what are the basic techniques used to accomplish such a change? The key is how the musical structure is communicated. Therefore, a good understanding of structure–whether theoretic or intuitive–is a prerequisite for a convincing musical performance (Clarke, 1999).

This chapter surveys the basic principles and techniques that musicians use to convey and project music structure (see also overviews in Gabrielsson, 1999; Palmer, 1997). We will only consider auditory communication, and leave out visual cues in concert performances (see Davidson & Correia, this volume). Another issue only briefly covered is perception–the extent to which subtle variations in performance are perceived by the listener. Our focus will be on topics that have been the subject of systematic research and that can be useful for music students and teachers. Largely for practical reasons, this research has tended to focus more on traditional classical music than other styles, timing more than other parameters such as dynamics and articulation, and piano performance more than other instruments. Nonetheless, most of the results obtained, and hence most of the principles presented here, seem to have a universal character and may be applied to a wide variety of genres and instruments.

Concepts and terms

How can music performance be studied scientifically? We base our analysis primarily on information available in the sound alone. All cues for musical communication are contained in sound, which in turn can be described and quantified in terms of physical variables such as the duration and sound level (physical intensity) of each tone.

Tone duration is the time interval between the physical start of the tone (onset) and the end of the same tone (offset); that is, the sounding duration of a tone. More important for timing in music is the interonset interval (IOI), defined as the time interval between the onset of the tone and the onset of the immediately following tone. In other words, IOI is the sum of a tone's physical duration and the pause duration between the offset of the tone and the onset of the next, see Figure 1. IOI is easy to measure in MIDI recordings, and can also be estimated from commercial recordings using computer software.
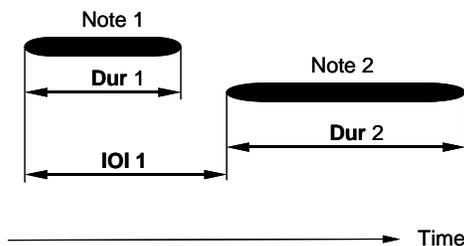

Figure 1. Definition of interonset interval (IOI) and duration (Dur) for two successive tones.

To find out if a tone is lengthened or shortened in performance, measured IOI values are compared with the note values given in the score. A nominal performance (or deadpan performance) is defined as a direct translation of the score into physical variables, where all notes of the same note value have the same nominal IOI, derived from the global tempo (tempo marking or mean tempo). In a nominal performance, for example, an eighth-note is always exactly half as long as a quarter-note. A nominal performance often serves as a reference point for research on musical timing. Timing variations relative to a nominal performance can be analyzed either tone by tone or as changes in local tempo, conceived as a continuously varying function of time (tempo curve).

Conventional western notation developed historically within the physical and cognitive constraints of performance and sight-reading. It deliberately fails to describe music in too much detail, since that would make it too difficult to read. This means that it is not possible to define nominal values for articulation and dynamics in conventional scores. Another consequence is the development of performance conventions–standard (but generally style-specific) interpretations of notational symbols that are not evident in the score itself. For example, strings of eighth-notes in a jazz score are typically performed unevenly in long-short patterns (swing).

What are the basic building blocks of musical structure? Most tonal music has a hierarchical phrase structure, sometimes simply called grouping. The slowest level is the entire piece, which is then divided and subdivided into sections, phrases, subphrases and melodic groups. Superimposed upon this is usually a metrical hierarchy: the beat or tactus (corresponding to when you tap your feet) is grouped, usually in groups of two or three, into measures and groups of measures. The beat can also be divided into sub-beats. Phrasing and meter are theoretically independent, although phrase and metric boundaries often coincide, reinforcing the overall perceived grouping (Lerdahl & Jackendoff, 1983), see Figure 2.

2

Figure 2. Phrase structure and metrical structure in an excerpt from Haydn's Symphony no. 104. The dots represent the hierarchical metrical structure. The top level in the figure is the beat level, the second the measure level, and the third and fourth are hypermetric levels. The hierarchical phrase structure is shown with brackets. The top level in the figure is the fastest, with only a few tones in each phrase or group. The slowest level corresponds to the whole phrase. (From Lerdahl and Jackendoff, 1983. Copyright 1983 by MIT Press. Used by permission.)

Commonalities among performers and between repetitions

What are the similarities and differences in separate performances of the same piece? When discussing interpretation, the emphasis is often placed on performance differences. However, even musicians who are considered to interpret music quite differently may produce remarkably similar patterns of timing and dynamics (Repp, 1992).

Musical structure is reflected in physical variables in a number of ways. Figure 3 illustrates the variation of IOI and dynamics in two performances of the same piece. We see that both pianists express the phrase structure by lengthening and softening the tones (ritardando and diminuendo) at the end of each melodic gesture, in measures 4 and 8. The slowing and softening are more pronounced at the end of the whole phrase, and is quite substantial, and thus clearly perceptible. The differences in interpretation between the pianists are largely seen in variations within phrases and on a note-to-note level.

It is often argued that a repeated passage should be performed differently in both cases. This is, however, not generally confirmed in measurements. On the contrary, there are often striking similarities between the first and second presentation of a thematic group. This is also true for the repetition of a whole piece on different occasions. In Figure 3, the differences between the two performances are very small and, at least for timing, are below the perceptual limits in most cases.

Basic Principles and Techniques

Tempo

Global tempo can vary substantially in different renditions of the same piece by different performers. In commercial recordings of both Schumann's Träumerei and Chopin's Etude in E major Op. 10 No. 3, the fastest tempo was about twice as fast as the slowest (Repp, 1992, 1998a). Clearly, tempo is influenced not only by tempo indications but also by the performer's interpretation, in particular by the intended motional and emotional character (Juslin & Persson, this volume).

Collier and Collier (1994) investigated tempo in jazz by analyzing a large number of commercial jazz recordings. They found that double time (doubling of tempo, sometimes introduced into an improvisation to produce a contrasting passage or to increase the intensity) corresponded to a tempo ratio of 2.68:1–well above a mathematical doubling of tempo.

When the global tempo of a performance is changed, patterns of local timing variations may also change. For example, there may be a tendency toward more expressive timing variation (relative to tempo) at slower tempi (Repp 1995). Also, the perceptual or motor limits of tone duration may alter the expressive pattern.
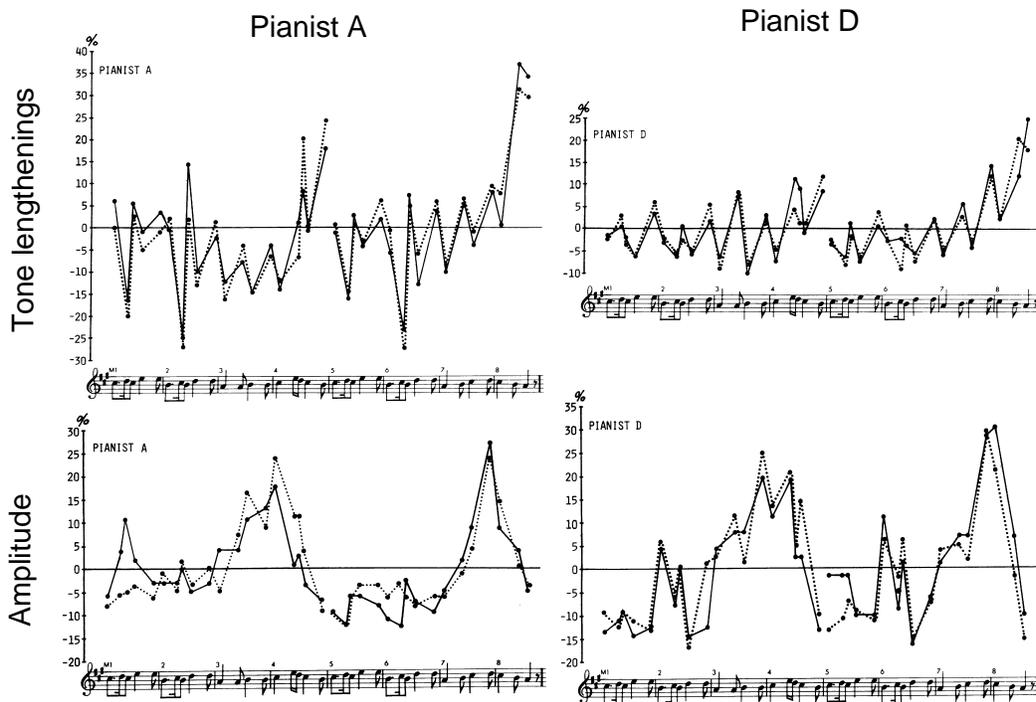

Figure 3. The first eight measures of Mozart's Piano Sonata K331, as performed by two pianists. The upper graphs illustrate the timing, with the vertical axis showing the deviation in IOI of each tone relative to nominal duration. The lower graphs illustrate the dynamic variation, with the vertical axis showing the peak amplitude of each tone, relative to the mean of all tones in each performance. The first performance is represented by full lines, and the repetition by dashed lines (Adapted with permission from Gabrielsson, 1987. Copyright 1987 Royal Swedish Academy of Music.)

Phrasing

In music from the Romantic period, large variations in local tempo are an essential part of the performance tradition. Phrases often start slow, speed up in the middle, and slow down again toward the last tone (e.g., Henderson, 1936; Repp, 1992). Dynamic variations tend to follow a similar pattern: soft in the beginning, loud in the middle, and softer towards the end of the phrase, see Figure 3 (Gabrielsson, 1987).

These typical shapes of timing and dynamics are observed in a majority of performances of Romantic music and are important for conveying the basic phrase structure to the listener. The ritardando at the end can communicate the phrase level, with typically a more pronounced ritardando at the end of a musical unit of longer duration, or at a slower hierarchical level. This is clearly the case in Figure 3, where pianists A and D both lengthen the tones more at the end of the example than in

4

the middle. The dynamic variation follows a similar pattern. In this way, not only the phrase boundaries, but also their hierarchical level–and hence the hierarchical phrase structure of the whole piece–can be communicated, just by changing tempo and dynamics. Similar principles are found in speech, where lengthening is used to communicate phrase and sentence boundaries.

The exact amount and shape of the variation over the phrases is an important issue for the performer to decide. In Figure 3, both pianists slow down at the phrase boundaries, but follow different tempo curves. For example, pianist A lengthens the final tones in each sub-phrase more than pianist D. Different shapings of local tempo or dynamics can entirely change the character of a performance, signaling different expressive intentions (Battel & Fimbianti, 1998).

Several models of these typical tempo variations have been developed. The first computational model was presented by Todd (1985, 1989; see also Windsor & Clarke, 1997; Penel & Drake, 1998) and accounted for the variation of measure duration over phrases. Todd later developed a revised model based on a different mathematical function, which he argued was more closely related to a metaphor of physical motion (Todd, 1992, 1995). Friberg (1995) modified Todd's first model so that it may be applied at the note level, introducing several extra parameters to account for individual variations. In Figure 4, this phrase arch model is fitted to three different piano performances of Schumann's Träumerei measured by Repp (1992). As can be seen in the figure, the model catches most of the individual variation regarding phrasing, but misses local variations on the note level.

Phrasing in Baroque music typically involves smaller variations in local tempo than in Romantic music. Baroque music tends to have a more motoric, metrical character (as do most contemporary jazz and pop), suggesting the metaphor of a mass moving at a constant speed, creating a kind of musical momentum. This natural coupling of motion and music has been investigated in an intuitive way in a number of publications (overview in Shove & Repp, 1995). Looking for direct couplings between the physical world and music, Friberg and Sundberg (1999) discovered a close connection between how a runner stops and how Baroque music stops at a final ritardando. They found that the average velocity of the runner and the average local tempo closely followed the same curve: their model could account for both the deceleration of individual runners and of individual music performances. Moreover, participants in listening experiments preferred final ritardandi that corresponded to the stopping runners, suggesting, in this case, a close coupling to physical motion.

The fastest level in the phrase hierarchy consists of small melodic units of a few notes each. Grouping (i.e., segmentation) at this level tends to be quite ambiguous, often with several possible interpretations. So communication of this structure can be subject to more individual interpretation than, say, communication of longer phrases. One example is found in Mozart's A-major sonata (Figure 3) where many performers chose the first five notes as a group while others chose the first four notes, giving the second group an upbeat. This ambiguity may be due to contradictory perceptual cues from different aspects of the musical structure, such as the melodic contour or the meter, and can be resolved in performance by inserting a micropause between the last tone of one phrase and the first of the next, which both interrupts the sound and delays the onset of the following tone (Friberg, Sundberg & Frydén, 1987; Friberg, Bresin, Frydén & Sundberg, 1998; Clarke, 1988).

A deceptively simple performance principle is the higher, the louder (Sundberg et al., 1991). The origin of this principle appears to be physical: wind instruments (including the voice) tend to produce louder tones at higher pitches, even though effort or input pressure is held constant. Often, the most important tone in a phrase is also the highest in pitch. In this case, the high-loud principle produces natural-sounding phrasing (cf. Windsor & Clarke, 1997; Palmer, 1996a, Krumhansl, 1996).
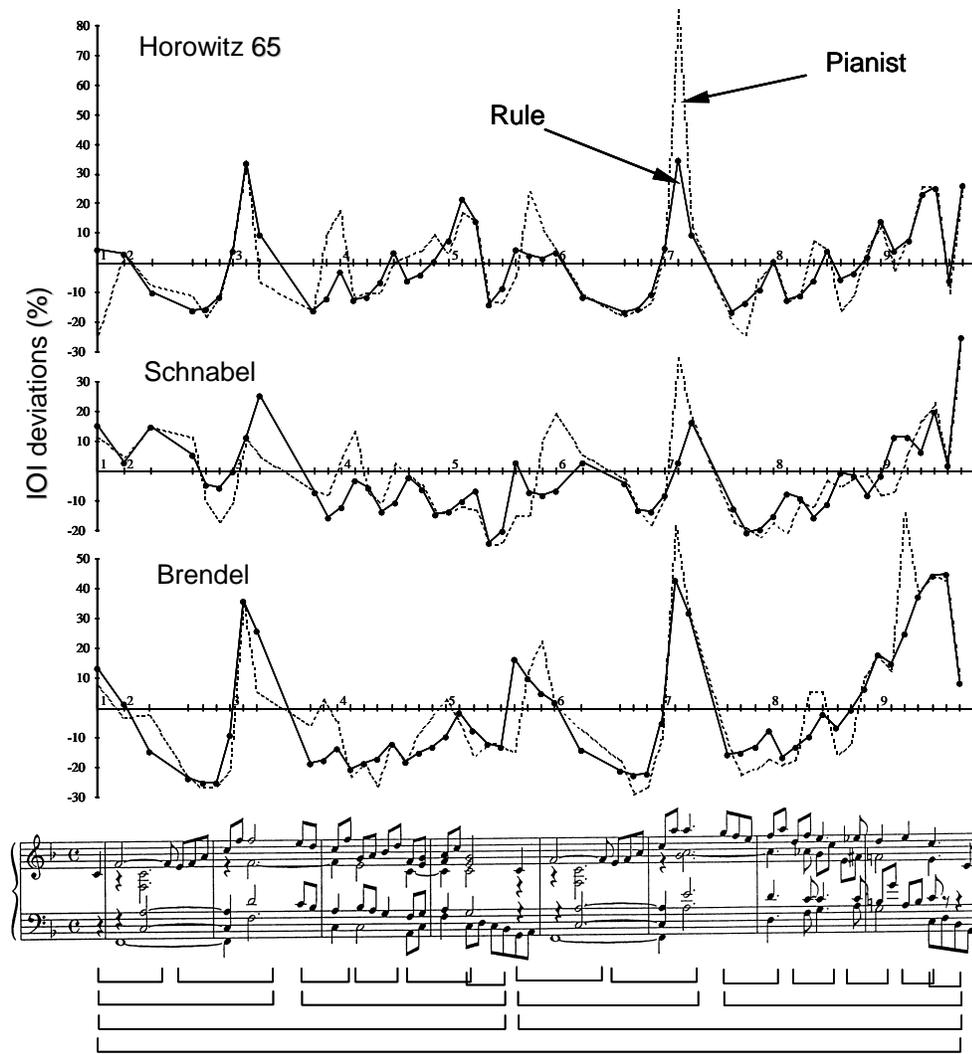
5

Figure 4. The dotted lines show the deviations in IOI from nominal values for performances of Schumann's Träumerei by three different pianists. The solid lines show predicted IOI deviations according to the Phrase arch model. The predictions have been fitted to the three performances by adjusting the parameters of the model. The brackets below indicate the grouping analysis used in the model.

Harmonic and melodic tension

The notion of musical tension is common to both music theory and music psychology. Many authors claim that the contrast between tension and release is a major source of musical interest. Tension is coupled to expectancy; an unexpected tone or chord creates tension (Krumhansl, 1990). A number of sources contributing to harmonic and melodic tension have been identified. A partial list is given in Table 1 (cf. Bigand, Parncutt & Lerdahl, 1996).

6

Table 1. A list of tonal relationships contributing to melodic and harmonic tension.

Case	Relationship	Increased tension for:
1	between keys	(i) a modulation to a key that is distant on the circle of fifths, or (ii) to a scale with few tones in common with the original scale
2	between chords; chord relative to key	(i) chords more distant to the key on the circle of fifths (comparing roots) (ii) chromatic chords, or chords that include tones foreign to the prevailing scale (iii) successive chords having few tones in common
3	tone relative to chord	(i) tones that are more distant on the circle of fifths from the root of the chord (ii) tones foreign to the diatonic scale associated with the chord
4	simultaneous tones in a chord	chords containing more dissonant intervals
5	melodic contour	unexpected melodic turns

Several models have been developed to address tension. For case 2(i), Sundberg et al. (1991) defined the <u>harmonic charge </u>of a chord as a weighted sum (root most important) of the distance between the tones of the chord and the prevailing tonic on the circle of fifths. They then allowed local tempo to slow down and sound level to increase in areas of high harmonic charge (Friberg, 1991), creating a kind of harmonic phrasing, see Figure 5.

For case 3(i), Sundberg et al. (1991) defined <u>melodic charge</u> as an increasing function of the distance on the circle of fifths. The performance model increases the sound level, IOI, and–if applicable–the extent of vibrato in proportion to the calculated melodic charge (Friberg, 1991).

A complex model taking into account most of the aspects in Table 1 was formulated by Lerdahl (1996). This model explained the majority of variation in subjects' ratings of perceived tension in a Mozart sonata (Krumhansl, 1996).

The most common way to communicate tension seems to be to emphasize notes or areas of relatively high tension, as in the models of harmonic and melodic charge described above. However, it is difficult to trace the origins of variations of timing and dynamics measured in real performances, since the various tension concepts often are coupled to each other and to the phrasing structure. For example, chords more distant to the key are more often found in the middle of phrases, while chords close to the key are more often found in the beginning or in the end of the phrase. Also, phrasing tends to dominate performance expression, which makes it hard to isolate the more subtle details such as the expression of melodic or harmonic tension.

Analyzing a performance of a Mozart sonata, Palmer (1996a) found a weak correspondence between predictions of Lerdahl's model and observed timing (IOI): positions of high tension were emphasized by lengthening. An additional finding was that relatively dissonant chords were often performed by delaying the melody by as much as 100 ms. The purpose of this performance strategy could either be to reduce the perceptual dissonance of the chord (as Palmer suggested), or to emphasize the melody tone by delaying it (see Accents below).
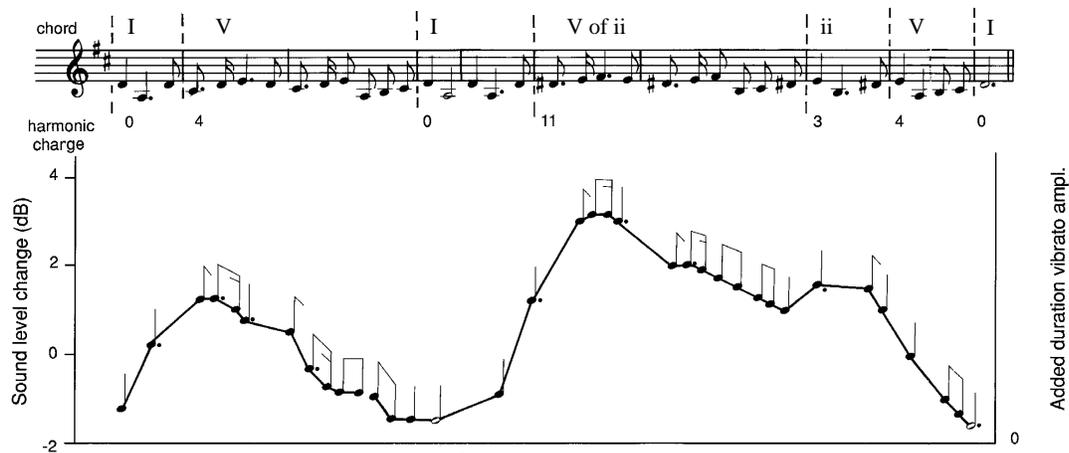
Figure 5. The harmonic charge model applied to a theme from Schubert's "unfinished" symphony. Sound level and vibrato rate increases, and tempo decreases, with increasing harmonic distance from the tonic. Note the peak in harmonic charge at the most distant chord (V of II) on the circle of fifths.

Repetitive metrical patterns and grooves

Music that is rhythmically regular often exhibits consistent patterns of timing and dynamic patterns within metrical units such as the measure. For example, if the first beat in each measure is accentuated, a dynamic pattern is formed that is repeated in each measure. This kind of patterning is often associated with dance, suggesting that these patterns serve to characterize the motional character of the piece.

Patterns in triple meter. In performances of both a Beethoven minuet (Repp, 1990) and a Chopin Nocturne (Henderson, 1936), the pianists played the second beat late and the third beat early, forming a long-short-long pattern.

In patterns of a half-note followed by a quarter-note in 3/4 time, or a quarter-note followed by an eighth-note in 6/8, the ratio of the IOIs of the long and short tones is usually in the range 1.7:1 to 1.9:1–consistently smaller than the nominal 2:1. This has been found in performances of Swedish folk music as well as Mozart and Chopin (Gabrielsson, 1987; Gabrielsson, Bengtsson & Gabrielsson, 1983; Henderson, 1936), see Figure 3 (note the characteristic zigzag patterns in the timing graphs). Sundberg et al. (1991) called this effect double duration and incorporated it into a performance model.

A different pattern is found in Viennese waltzes: they are usually performed with an early second beat, thus forming a short-long-intermediate timing pattern for the three beats in the measure. This pattern was first systematically investigated by Bengtsson and Gabrielsson (1983), who also observed a hypermetric timing pattern spanning two measures that reflects how these waltzes are danced, as well as the underlying harmonic structure.

Two-note patterns. In jazz, Baroque, and folk music, it is common to apply a long-short pattern to consecutive eighth-notes (or the shortest prevalent note value). In Baroque music, this performance convention is referred to as notes inégales (Hefling, 1993). Drummers playing typical comping patterns in jazz perform different ratios between consecutive eighth-notes depending on global tempo. At slow to medium tempi, the ratio tends to be about 3:1 (dotted eighth + sixteenth), dropping to 1:1 (even eighth-notes) at fast tempi (Friberg and Sundström, 1997). At the same time the soloist uses smaller ratios than 2:1. Surprisingly, the ratio of 2:1, implied by the often-mentioned triplet feel of swing, was not observed in these experiments.

Duration contrast. Contrary to the double duration principle described above, the contrast between long and short tones is often increased by a lengthening of comparatively long notes and a

8

corresponding shortening of comparatively short notes (Taguti, Mori & Suga, 1994*).* However, duration contrast applies to <u>all</u> notes, while double duration applies only to the 2:1 pattern. In the case of a repeated rhythmical figure, duration contrast will appear as a repetitive timing pattern. Duration contrast can also be applied backwards: depending on the performers intention, the contrast may instead be decreased (Juslin and Persson, this volume). A model was formulated by Sundberg et al. (1991).

Articulation

Articulation–at least in the sense of <u>staccato</u> versus <u>legato</u>–may be defined mathematically as the ratio of tone duration to IOI. Articulation strongly affects motional and emotional character (De Poli, Rodà & Vidolin, 1998; Battel & Fimbianti, 1998, Juslin & Persson, this volume).

The duration of <u>staccato</u> tones has been found to correspond on average to about 40% of the IOI (or note value) in typical performances of a Mozart Andante movement for piano (from sonata K545; Bresin & Battel, forthcoming). When pianists were asked to play the same piece <u>brillante</u> or <u>leggero,</u> the duration decreased to 25% (<u>staccatissimo</u> range), while in <u>pesante</u> performances the duration approached 75% of IOI (<u>mezzostaccato</u> range). A typical value of 40% is in nice agreement with C. P. E. Bach's (1753) observation that <u>staccato</u> notes should be played with less than 50% of their nominal duration.

In <u>legato</u> playing on the piano, successive tones often overlap–both keys are down for a short period of time. The amount of overlap (in milliseconds) increases with IOI and pitch interval size (Repp, 1997; Bresin and Battel, forthcoming). Bresin and Battel also found that the overlap time was used for expressive purposes: the pianists used more overlap when instructed to play <u>appassionato</u> than <u>piatto</u> (flat). Overlap is also dependent on direction of melodic motion; descending melodic patterns are usually played with more overlap than ascending (Repp, 1997; Bresin, 2000).

Bresin (2000) also found an interesting link between articulation and physical locomotion. The overlap time of the feet in walking was found to vary qualitatively in the same way as in <u>legato</u> articulation. Flying time in running (the time during which neither foot is in contact with the ground) was similarly found to be related to detached time in <u>staccato</u> articulation. Based on these measurements, Bresin (2000) was able to formulate models for <u>staccato</u>, <u>legato,</u> and tone repetition.

Accents

The meaning of the term accent varies considerably in music-theoretic literature. Two main categories can be identified: immanent accents and performed accents (cf. Parncutt, in press; Lerdahl & Jackendoff, 1983).

An <u>immanent accent</u> is evident from the structure of the score itself, meaning that even if the score is performed nominally, these positions will be perceived as accented. For example, immanent accents may occur on notes in metrically strong positions, comparatively long note values, on the second tone of an upward leap, on the top tone of a melodic turn, or at increased harmonic tension (melodic immanent accents: Thomassen, 1982; Huron & Royal, 1996).

A <u>performed accent</u> is added by the performer relative to the nominal performance. This definition is closer to the common use of the term. Performed accents seem primarily to be used to reinforce immanent accents. This was confirmed in simple melodies (Drake & Palmer, 1993). However, in more complex (real) music, only weak couplings between performed and immanent accents have been found: only when positions interacting with the grouping structure were disregarded, was it possible to get statistical significance in piano performances of Schumann's <u>Träumerei</u> (Penel & Drake, 1998). This finding may indicate that the concept of melodic accent is only a cue leading to the forming of melodic groups. In fact, cues for immanent accents and cues for making automatic melodic grouping are similar (e.g. Friberg, Bresin, Frydén & Sundberg,1998; Cambouropoulos, 1998).

Perhaps the most obvious way to perform an accent is to increase loudness. In the case of instruments where the loudness envelope can be varied (such as winds, strings and voice), there are

9

several possibilities, including a loudness increase at the beginning of the tone, or a shorter attack time. Timing can also be used in several ways to emphasize a tone: (1) by lengthening the tone; (2) by delayed the onset, i.e. lengthening the preceding tone (IOI) and possibly inserting a micropause before the accented note; and (3) by playing the accented note more legato (Henderson, 1936; Clarke, 1988; Drake & Palmer, 1993). In addition, a change in articulation such as one legato note surrounded by staccato notes may signal an accent (cf. Lerdahl & Jackendoff, 1983).

Ensemble timing

In polyphonic classical music, it is often important to highlight the melody. An obvious way is to play it louder. However, timing is also an effective method. If two tones are presented at almost the same time, the first will be perceptually emphasized (Rasch, 1978). This technical device, commonly referred to as melody lead, has been observed in string and wind trios, as well as in piano performances. The melody is typically played about 20 ms ahead of the other voices, within the range of about 7 to 50 ms (Rasch, 1979; Palmer, 1996b; Vernon, 1936).
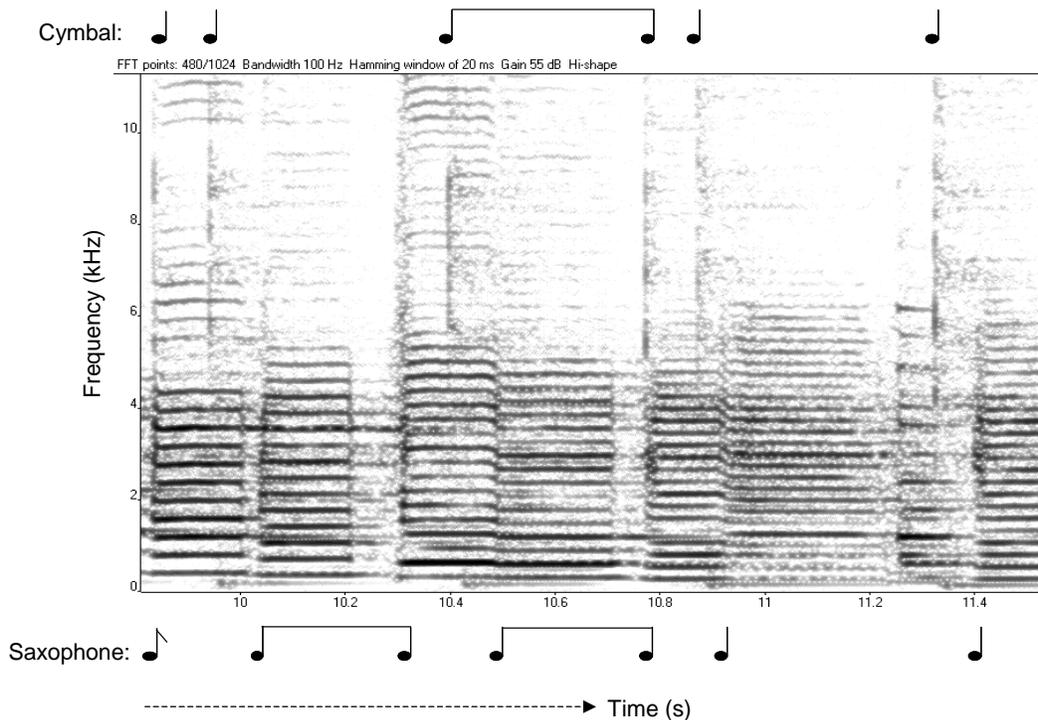

Figure 6. A spectrogram of a short passage of "My Funny Valentine" performed by Miles Davis Quintet, 1964, illustrating the timing relation between the ride cymbal and the soloist in jazz. The cymbal onsets appear as vertical lines in the high frequency part of the graph. The saxophone onsets appear as breaks, or vertical shifts, in the horizontal lines in the lower part of the graph. In this example, cymbal is being played with a swing ratio of about 4:1 and the saxophone with a swing ratio of approximately 3:2. The downbeat saxophone tones are delayed relative to the cymbal by about 100 ms, but on the upbeats, the cymbal and the saxophone are synchronized.

The opposite often happens in jazz: soloists deliberately play behind the beat. At each quarter-note beat, the soloist was delayed relative to the ride cymbal by up to about 100 ms at slow

tempi (Ellis, 1991; Friberg, in preparation). At the same time the soloist and drummer were synchronized at the upbeats, i.e., on the eighth-notes between the beats. The purpose here is probably not to highlight the soloist, which in this style is clearly audible, due to either large spectral differences or the use of microphones. Rather, this timing combination creates both the impression of the laid-back soloist often strived for in jazz, and at the same time an impression of good synchronization, see Figure 6.

Classification and Purpose of Performance Variations

Non-notated variations in timing and dynamics (deviations from the nominal performance) can be divided into two main types (cf. Juslin, Friberg & Bresin, forthcoming). Expressive variations are deliberately meaningful or communicative (but not necessarily conscious). Non-expressive variations are of two kinds: variations due to technical limitations of the instrument and the performer, and random variations (including imperfections in the perceptual timing and motor system). Neither of these last two have a deliberate communicative function; they may tell us something about the performer's abilities, but this is of course unintentional. Nevertheless, they can be important for the naturalness of a performance.

Expressive variations can be classified according to their apparent communicative purpose. They may either communicate the structure of the music, or express its character. We use the term character to refer either to the emotional (happy, sad, etc.) and motional (urgent, calm, swingy, etc.) implications of the music. Emotional character is dealt with in more detail by Juslin and Persson (this volume).

Sundberg (2000) identified two main underlying principles for communicating musical structure. The first involves the differentiation of pitch and duration. Categorical perception is improved by increasing the difference between categories, such as stretching the frequencies of scale tones, or playing short notes even shorter (e.g duration contrast, high loud). The second principle involves the grouping of notes in phrases, metrical units, or harmonic areas. For example, phrases are often performed with a diminuendo at the end. This increases efficiency of the musical communication by introducing redundancy; the phrase boundaries are often recognized even without this cue.

Structure and character are not necessarily independent. Character can be characterized in terms of how the structure is communicated. In fact, most of the cues described by Juslin and Persson (this volume) for communicating different emotions can be realized by the techniques for structural communication described in the present chapter (Bresin & Friberg, 2000; Battel & Fimbianti, 1998).

Applications in Music Pedagogy

As outlined in this chapter, music psychology research has over the past years accumulated a substantial body of knowledge about music performance that could be incorporated into music teaching. Knowledge of the basic principles of structural communication is possibly more important in earlier stages of instrumental learning; once these principles are mastered, both in theory and in performance, it may be easier to develop an individual voice and allow the more interesting development of an artist's individual personality to come to the fore, see Figure 7 (Andreas C. Lehmann, personal communication).
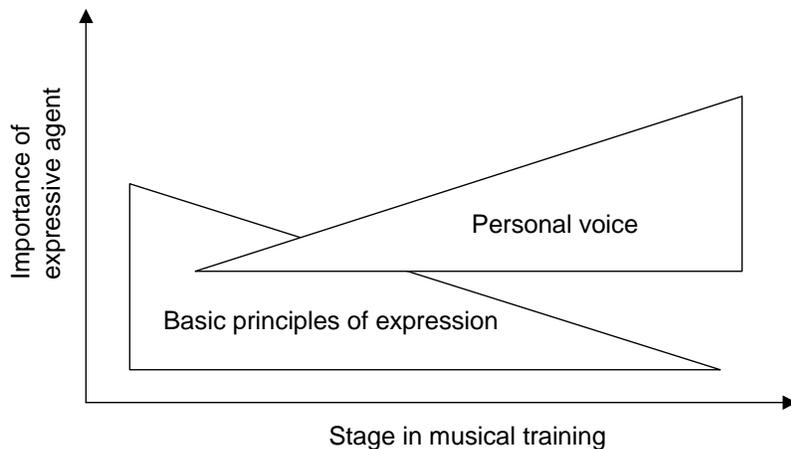
11

Figure 7. Relative importance, at different stages during performance studies, of the basic principles of communicating musical structure and character, and the musician's personal voice, or individual approach to musical interpretation (adapted from Andreas C. Lehmann, personal communication).

Teaching theory of structural communication

Each of the basic principles and techniques described above can be explained and discussed in terms of generality and individual differences. Consider phrasing as an example. A teacher might explain how performers introduce small ritardandi or diminuendi at the end of phrases, and increase their extent at the end of sections. This archetypal phrasing is not so easy to hear from just listening to recordings, and can come as a surprise, even to musicians. It is also useful to show measurements of performances by well-known musicians such as in Figure 3 and 4 so that students understand that the same principles are also used in top-level performances. Other aspects of structural communication, such as articulation or timing patterns, can be taught in a similar way.

The use of sound examples generated by computer-based performance models (e.g. www.speech.kth.se/music/performance/) allows each aspect (e.g. phrasing) to be studied and listened to in isolation. Graphs showing the variation of dynamics and timing also help the student to hear and understand what is happening. A computer program such as Director Musices (Friberg et al., 2000) allows a teacher to directly apply phrasing rules to any music example, with the flexibility of changing the extent and parameters–all in real time and all in the classroom. Students can then judge for themselves how much variation of timing and dynamics is appropriate for the interpretation of a given phrase. Regardless of their musical training or ability to read notation, students can listen to differences between different renditions of a given passage and learn to focus their auditory attention on one aspect. Thus, the suggested pedagogical approach may be useful not only for teaching performance but also for teaching relevant auditory skills.

Performance studies

Once the theoretical concepts have been explained, the various performance principles could be practiced separately. If musical tension is selected, a student or teacher might first make an analysis of tension-relaxation patterns in a piece, and think about different ways in which tension could be appropriately communicated or expressed in a particular performance on a given instrument. Then the student could practice emphasizing the tension-relaxation patterns in different ways (e.g. timing, dynamics). For example, one could, as an exercise, emphasize harmonic tension but otherwise play without expression. This can extend the student's expressive vocabulary and facilitate the creation of different emotions or patterns of implied movement (cf. Juslin & Persson, this volume).

The importance of feedback for efficient learning is well known. Using computer analysis tools, the variations of timing and dynamics in a student's performance can be measured and shown graphically. This could help the student to evaluate, for example, how well her or his variations of dynamics and timing communicate intended patterns of tension and relaxation. Commercial recordings can also be analyzed in this way, but the methods of analysis and the required software tools depend on which instrument and structural aspect is being studied.

A majority of the principles of structural communication can be studied with just a microphone connected to a computer using commercial or free software. Ensemble timing, such as observing the amount of lead/lag for each performer, can be illustrated in a spectrogram program (e.g. Soundswell, www.hitech.se; Wavesurfer, www.speech.kth.se/wavesurfer). An example was given in Figure 6. Illustrations of dynamics and articulation can be obtained by computing a smoothed RMS value of the audio signal, a feature found in several audio wave editors.

Studying interonset timing and local tempo variations is currently a little bit more complicated. Here it is necessary to compute for each note the deviations of IOI relative the nominal value. The IOIs and durations can be obtained from a MIDI-equipped instrument (e.g. Disklavier piano from Yamaha, Zeta string instruments, electric guitar MIDI controller, or an acoustic instrument with an audio-to-midi converter) connected to a computer. A program such as POCO (Desain et al., 1997; Honing, 1990; stephanus2.socsci.kun.nl/mmm/) can match the values to the score so that deviations are obtained.

Pianists may find an instrument such as the Yamaha Disklavier useful for the study of structural communication. This is essentially an acoustic piano in which the timing and dynamics of each key pressure can be registered on a computer. The computer can also control the instrument so that a recorded performance can be listened to directly on the instrument. The Disklavier has been used regularly by the second author in teaching advanced courses in piano performance (Battel et. al., 1998). One such course consisted of two parts: (1) music performance analysis, and (2) analytical methods for performance. The participants played the first 10 bars of Mozart's Sonata K333 in Bb major on the Disklavier. Variations in dynamics and timing produced by the players were displayed on a computer screen. The data were compared to those of the second author's performance recorded earlier on a Disklavier, and to those produced by a performance rule system implemented in MELODIA (Battel & Bresin 94). Finally, the musical meaning of performance variables in the historical period, style, and performance tradition of the composer were analyzed and discussed.

In summary, recent research on music performance has opened up a range of new possibilities for teaching expression in the sense of structural communication. The development of new software tools for this purpose would further facilitate the use of the computer as a complement in the classroom.

References

Bach, C. P. E. (1957). Facsimile of first edition, 1753 (Part One) and 1762 (Part Two). Ed. L. Hoffman-Erbrecht. Leipzig: Breitkopf & Härtel.

Battel G. U., Canazza S., Fimbianti R., & Rodà, A. (1998). III. Modalités de l'expérience. In Gruppo Analisi e Teoria Musicale: Quand une analyse schenkerienne est-elle utile aux interprètes? Revue Belge de Musicologie, Vol. LII, 331–334.

Battel G. U., & Fimbianti R. (1998). How to communicate expressive intentions in piano performance. In A. Argentini and C. Mirolo (Eds.), Proceedings of the XII Colloquium on Musical Informatics (pp. 67–70).

Battel G. U., Bresin R. (1994). Analysis by synthesis in piano performance: A study on the theme of the Brahms' Variations on a Theme of Paganini op. 35. In A. Friberg et al. (Eds.), Proceedings of the Stockholm Music Acoustics Conference 1993 (pp. 69-73). Stockholm: Royal Swedish Academy of Music.

Bengtsson, I., & Gabrielsson, A. (1983). Analysis and synthesis of musical rhythm. In J. Sundberg, (Ed.), Studies of Music Performance (pp. 27-60). Stockholm: Royal Swedish Academy of Music.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. Perception & Psychophysics, 58(1), 125-141.

Bresin, R. (2000). Virtual virtuosity: Studies in automatic music performance. Unpublished doctoral dissertation, Royal Institute of Technology, Stockholm.

Bresin, R, & Battel, G. U. (forthcoming). Articulation strategies in expressive piano performance. Journal of New Music Research.

Bresin, R., & Friberg, A. (2000). Emotional coloring of computer controlled music performance. Computer Music Journal, 24(4), 44-62

Cambouropoulos, E. (1998). Towards a general computational theory of musical structure. PhD Thesis, Faculty of Music and Department of Artificial Intelligence, University of Edinburgh.

Clarke, E., F. (1988). Generative principles in music performance. In J. A. Sloboda, (Ed.), Generative processes in music (pp. 1-26). Oxford: Clarendon press.

Clarke, E., F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), The psychology of music (2nd ed.) (pp. 473-500). New York: Academic.

Collier G. L., and Collier, J. L. (1994). An exploration of the use of tempo in jazz. Music Perception, 11(3), 219-242.

De Poli G., Rodà, A., & Vidolin, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. Journal of New Music Research, 27(3), 293-321.

Desain, P., Honing, H., & Heijink, H. (1997). Robust score-performance matching: Taking advantage of structural information. In Proceedings of the 1997 International Computer Music Conference (pp. 337-340). San Francisco: CMA.

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

Ellis, M. C. (1991). An analysis of "swing" subdivision and asynchronization in three jazz saxophonists. Perceptual and Motor Skills, 75, 707-713.

Friberg, A., Colombo, V., Frydén, L., & Sundberg, J. (2000). Generating musical performances with Director Musices. Computer Music Journal, 24(3), 23-29.

Friberg, A. (1991). Generative rules for music performance: A formal description of a rule system. Computer Music Journal, 15(2), 56-71.

Friberg, A. (1995, May). Matching the rule parameters of Phrase arch to performances of "Träumerei": A preliminary study. In A. Friberg & J. Sundberg (Eds.), Proceedings of the KTH Symposium on Grammars for Music Performance (pp. 37-44).

Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners. Journal of the Acoustical Society of America, 105(3), 1469-1484.

Friberg, A. and Sundström, A. (1997). Preferred swing ratio in jazz as a function of tempo. Speech Music and Hearing Quarterly Progress and Status Report, 4/1997, 19-28.

Friberg, A., Bresin, R., Frydén, L., & Sundberg, J. (1998). Musical punctuation on the microlevel: Automatic identification and performance of small melodic units. Journal of New Music Research, 27(3), 271-292

Friberg, A., Sundberg, J. & Frydén, L. (1987). How to terminate a phrase. An analysis-by-synthesis experiment on the perceptual aspect of music performance. In A. Gabrielsson (Ed.), Action

and Perception in Rhythm and Music (pp. 49-55), Stockholm: Royal Swedish Academy of Music.

Gabrielsson, A. (1973). Adjective ratings and dimension analysis of auditory rhythm patterns. Scandinavian Journal of Psychology, 14, 244-260.

Gabrielsson, A. (1987). Once again: The theme from Mozart's piano sonata in A major (K. 331). A comparison of five performances. In A. Gabrielsson (Ed.), Action and Perception in Rhythm and Music (pp. 81-103). Stockholm: Royal Swedish Academy of Music.

Gabrielsson, A. (1999). Music performance. In D. Deutsch (Ed.), The psychology of music (2nd ed.) (pp. 501-602). New York: Academic press.

Gabrielsson, A., Bengtsson, I., & Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. Scandinavian Journal of Psychology, 24, 193-213.

Hefling, S., E. (1993). Rhythmic alteration in Seventeenth- and Eighteenth-Century music: Notes inégales and overdotting. New York: Schirmer.

Henderson, A., T. (1936). Rhythmic organization in artistic piano performance. In C. E. Seashore (Ed.), Objective analysis of musical performance. University of Iowa Studies in the Psychology of Music, Vol IV (pp. 281-305). Iowa City: University of Iowa.

Honing, H. (1990). POCO: An environment for analysing, modifying, and generating expression in music. In Proceedings of the 1990 International Computer Music Conference. (pp. 364-368). San Francisco: CMA.

Huron, D., & Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. Music Perception, 13, 489-516.

Juslin, P., Friberg, A., & Bresin, R. (forthcoming). Toward a computational model of performance expression: The GERM model. Musicae Scientiae.

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

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's piano sonata K282: Segmentation, tension, and musical ideas. Music Perception, 13, 401-432.

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

Lerdahl, F. (1996). Calculating tonal tension. Music Perception, 13, 319-364.

Palmer, C. (1996a). Anatomy of a performance: Sources of musical expression. Music Perception, 13(3), 433-453.

Palmer, C. (1996b). On the assignment of structure in music performance. Music Perception, 14(1), 23-56.

Palmer, C. (1997). Music performance. Annual Review of Psychology, 48, 115-138.

Parncutt, R. (in press). Accents and expression in piano performance. Systemische Musikwissenschaft. Festschrift Jobst P. Fricke zum 65. Geburtsdag, Frankfurt: Peter Lang.

Penel, A., & Drake, C. (1998). Sources of timing variations in music performance: A psychological segmentation model. Psychological Research, 61, 12-32.

Rasch, R. A. (1978). The perception of simultaneous notes such as in polyphonic music. Acustica, 40, 21-33.

Rasch, R. A. (1979). Synchronization in performed ensemble music. Acustica, 43, 121-131.

Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. Journal of the Acoustical Society of America, 88(2), 622-641.

Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Träumerei". <u>Journal of the Acoustical Society of America, 92</u>(5), 2546-2568.

Repp, B. H. (1995). Quantitative effects of global tempo on expressive timing in music performance: Some perceptual evidence. <u>Music Perception, 13(</u>1), 39-57.

Repp, B. H. (1997). Acoustics, perception, and production of legato articulation on a computer-controlled grand piano. <u>Journal of Acoustical Society of America</u>, <u>102</u> (3), 1878-1890.

Repp, B. H. (1998a). A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. <u>Journal of the Acoustical Society of America, 104</u>, 1085-1100.

Repp, B. H. (1998b). The detectability of local deviations from a typical expressive timing pattern. <u>Music Perception, 15</u>(3), 265-289.

Shove, P., & Repp, B., H. (1995). Musical motion and performance: Theoretical and empirical perspectives. In J. Rink (Ed.), <u>The practice of performance: Studies in musical interpretation</u> (pp. 55-83). Cambridge: Cambridge University Press.

Sundberg, J. (2000). Grouping and differentiation. Two main principles in the performance of music. In T. Nakada (Ed.), <u>Integrated human brain science: Theory , method application (music).</u> (p. 299-314). Amsterdam: Elsevier.

Sundberg, J., Friberg, A., & Frydén, L. (1989). Rules for automated performance of ensemble music. <u>Contemporary Music Review</u>, 3, 89-109.

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.), <u>Representing musical structure</u> (pp. 161-197). London: Academic.

Taguti, T., Mori, S., & Suga, S. (1994). Stepwise change in the physical speed of music rendered in tempo. In I. Deliège (Ed.), <u>Proceedings of the 3rd International Conference on Music Perception and Cognition, Liège</u> (pp. 341-342). Liege, Belgium: ESCOM.

Thomassen, M. T. (1982). Melodic accent: experiments and a tentative model. <u>Journal of the Acoustical Society of America, 71</u>, 1596-1605.

Todd, N. P. McA. (1985). A model of expressive timing in tonal music. <u>Music Perception,</u> 3, 33-58.

Todd, N. P. McA. (1989). A computational model of rubato. <u>Contemporary Music Review, 3</u>, 69-88.

Todd, N. P. McA. (1992). The dynamics of dynamics: A model of musical expression. <u>Journal of the Acoustical Society of America, 91</u>(6), 3540-3550.

Todd, N. P. McA. (1995). The kinematics of musical expression. <u>Journal of the Acoustical Society of America, 97</u>(3), 1940-1949.

Vernon, L. N. (1936). Synchronization of chords in artistic piano music. In C.E. Seashore (Ed.), <u>Objective analysis of musical performance. University of Iowa studies in the psychology of music, Vol. 5</u> (pp 306-345). Iowa City: University of Iowa Press.

Windsor, W. L., and Clarke, E. F. (1997). Expressive timing and dynamics in real and artificial musical performances: Using an algorithm as an analytical tool. <u>Music Perception</u>, <u>15</u>(2),127-152.