*Cognitive Science Quarterly*

# On music performance, theories, measurement and diversity[1]

**Renee Timmers**
*University of Nijmegen, The Netherlands*[2]

**Henkjan Honing**
*University of Amsterdam, The Netherlands*
*University of Nijmegen, The Netherlands*

[Published as:
Timmers, R. & Honing, H. (2002) On music performance, theories, measurement and diversity. In M.A. Belardinelli (ed.). *Cognitive Processing* (International Quarterly of Cognitive Sciences), **1-2**, 1-19.]

[1] This research was funded by the Netherlands Organization of Scientific Research (NWO) in the framework of the *Music, Mind, Machine* project. Special thanks to Charles de Weert, Dirkjan Povel, Peter Desain and Ric Ashley for their helpful comments on an earlier version of the article and to Jane Gower for her corrections of the language. Thanks to Rob Broek for his performances of the Brahms' piece.

[2] Address for correspondence: Renee Timmers or Henkjan Honing, NICI, University of Nijmegen, P.O. Box 9104, NL-6500 HE Nijmegen, E-mail: renee74@xs4all.nl or honing@hum.uva.nl, Tel: +31(0)24-3612650, Fax: +31(0)24-3616066.

*Received 20.03.03* 

Measurement of musical performances is of interest to studies in musicology, music psychology and music performance practice, but in general it has not been considered the main issue: when analysing Western classical music, these disciplines usually focus on the score rather than the performance. This status seems to be at odds with the central position of music performance in musical behaviour. In this paper, we therefore argue for an increased focus on performance data in music research disciplines. What kind of science do we get, and what methods and techniques do we need, when we make it the central object of music research? What are the issues and what are the problems and solutions? The first issue that will be addressed is the definition and measurement of expressive timing. Defining expression in different ways highlights certain aspects of a performance and obscures others. The second issue is the interpretation of expressive patterns: what knowledge and decisions play a role in constructing a performance? This interpretation is complicated by the many perspectives that performers have towards music. Both lead to the issue of a multitude of equally acceptable performances of a single piece and the differences between them. From a comparison of the methods used to approach these issues, the contours of an empirical musicology of performance may arise.

Keywords: music performance, empirical musicology, expressive timing, methodology

## Introduction

Since the 1980s, the research questions and methodologies of musicology have changed and developed considerably. Most notable are trends towards more complete and formalized theories of music that are perception oriented, such as Lerdahl & Jackendoff's (1983) "A Generative Theory of Tonal Music" and Narmour's (1990) "Implication-Realization Model". Another new approach is performance oriented, like Clarke's (1988) "Generative principles in music performance", Epstein's (1995) "Shaping Time" and Rink's (1995) "The Practice of Performance". What we see in this literature is a rise of interest in music as a result of human perception and production, and a rise in experimental and empirical methodology; the importance of music being the way it is produced and perceived by experienced or less-experienced listeners or musicians. This viewpoint owes much to psychology, since perception and production depend on human cognitive processes.

The novel perspective emphasizes the complexity and flexibility of music. Music continually transforms itself, because musicians and listeners interpret music. Listeners attribute functions to music, such as the expression of emotions, the setting of a mood or the urging to movement, and musicians perform it in differing situations and different ways. This variety of interpretations complicates the identity of a musical piece and asks for a reconsideration of norms. Can music research be determinative, or should it deal with diversity in an adaptive way?

For the investigation of the performance data, new methodologies are needed; unfortunately, these are not yet fully developed. The concern in this paper is therefore to give an overview and evaluation of the new methods as far as they have been developed. The focus will be on *expressive timing* in piano performances of Western classical music. Pianists' interpretive choices and conceptions shape the expression of music, even when performances from musical notation are concerned. This interpretive input results in deviations from the score and in differences between performances of the same piece. In sum, the question treated in this paper is how to incorporate this creative aspect of performance into the study of classical music.

We will first compare different definitions of expression and then discuss the meaning or interpretation of the expressive patterns in performance data. This will be followed by a comparison of analysis methods used to study differences between renditions. The comparison will be illustrated with examples from musicology and psychology, with a specific focus on the temporal aspects of music, such as rhythm, timing and tempo (Honing, 2002).

## Definition of Expression

Expression is an important aspect of music. It is the added value of a performance and is part of the reason why music sounds alive and is interesting to listen to. In order to measure the expressive characteristics of a musical performance, an exact definition of expression is needed. Several definitions have been used in the music performance literature. Below, the viewpoints on the definition of expression are listed and their use for the analysis of a performance example is explored.

### Expression as microstructure

A very broad and in some sense neutral definition of expression is that expression completes what the score leaves unspecified. This is the notion of expression as microstructure (Repp, 1990; 1992a; Palmer, 1997). The microstructure consists of the large and small variations in timing, intensity, timbre and pitch (Palmer, 1997). These variations exist besides the score without a necessary dependence.

### Expression as deviation from a musical score

The most common definition of expression defines expression as deviations in the performance data from a mechanical rendition of a score. As early as the 1930s, Seashore (1938) observed that artistic performances tend to deviate from the “fixed and regular”. In the 1970s, Gabrielsson (1974; 1987) more explicitly stated that performances of rhythm are characterized by deviations from the norm as stated by the musical notation. He analysed the performance of rhythms in terms of their deviation from a hypothetical mechanical performance. Within the study of generative performance models, the definition has practical use. For example, Sundberg, Friberg & Frydén (1991a) presented a score note to a performance note conversion with the aid of a set of performance rules that act on various levels of the musical structure. The score has a constant tempo, a constant intensity and a constant intonation. The performance rules introduce deviations from this regularity, such as local increments of duration and loudness, and insertion of pauses.

**Expression as deviation within a performance**

Desain & Honing (1991) elaborated on the definition of expression as deviation from a norm by defining the norm within the performance. According to their definition (which only applies when a hierarchical structural description of the music is available), “expression is the deviation of a lower order unit from the norm as set by a higher order unit”. For example, the expressive variations of the durations of beats is expressed as ratios of the bar duration. Clarke (1995) explored this definition further and suggested that a norm can be defined by common music practice, such as the long/short interpretation of equal quarter notes, from which performances can be said to deviate (in the sense of exaggeration or diminution of the characteristic pattern).

**Comparison of definitions**

To compare these viewpoints on expression, we will take an example performance and analyse its expressive timing behaviour using these definitions. The example is a performance by a professional pianist of the theme from Brahms' *Variations on an Original Theme* (D major, Op. 21, No. 1, 1861) for piano solo (see Figure 1). The piece is in 3/8 meter, eighteen measures long, and consists of two halves that are both repeated in the performance. The halves consist of two sub-phrases of which the first spans four measures, and the second five measures. The harmonic structure consists of, for example, a harmonic suspension and harmonic pedal in the first measures, followed by faster harmonic progressions and dissonances in measures 7-10 (see for a detailed description Timmers, Ashley, Desain & Heijink, 2000).

The timing data to be analysed consists of time intervals between succeeding onsets of melody notes. The calculation of inter-onset intervals (IOI's) of succeeding notes has been common practice in expressive timing research and has usually been done in the way indicated in Figure 2.

The selection of melody notes only is a quite arbitrary choice and could have been made differently (e.g. each first note at each eighth note beat). Nevertheless, it is a sensible choice, since the notes belong to the same structural unit (melodic line) and performers are found to make very few mistakes in the melody (see e.g. Palmer & van de Sande, 1993). The calculation results in a list of note IOI's that indicates the duration between successive note onsets. This means that longer score intervals (e.g. quarter notes) generally have longer note IOI's (ca. 800 ms) and shorter score intervals (e.g. the ornamental notes in measure 6) have smaller note IOI's (ca. 80 ms).


*Figure 1*. Theme of Variations on an original theme for pianoforte, J. Brahms, Op. 21, No. 1 (1861).

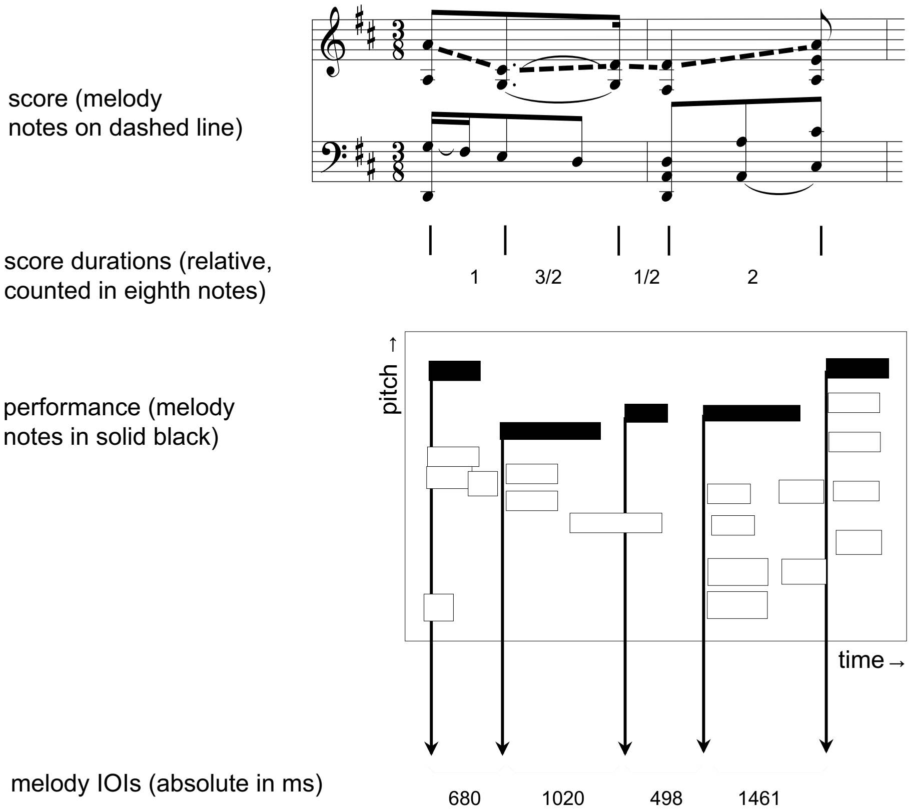


*Figure 2*. Extraction of a melody timing pattern from a piano performance. The score is given at the top, followed by the relative (symbolic) note durations of the melody. Subsequently, the performance data is presented as a piano roll notation, illustrating how the inter-onset intervals (IOs) are derived from this data. Finally, the absolute (measured) performance IOIs are given.

*Microstructure*

From the viewpoint of expression as microstructure of a performance, variations in note IOI are examined on a normalized scale (see e.g. Repp, 1992a), which means that the IOI's are corrected for their score duration. This normalization is done by dividing each note IOI by its corresponding score interval. In Figure 3, the score interval is given in multiples of eighth note beats. It shows the normalized note IOI (y-axes) for each note in the score. For readability, measures are indicated on the x-axes and phrase boundaries are indicated as vertical dashed lines.

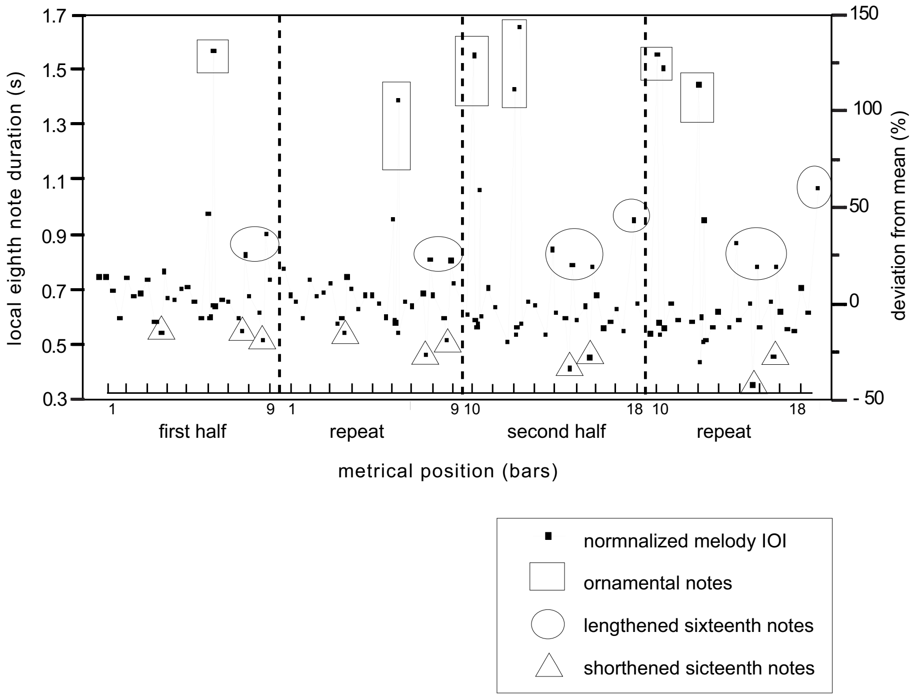


*Figure 3*. Timing pattern of the melody. It shows the relatively large deviations (lengthenings) of sixteenth notes and ornaments.

Surprisingly, the largest relative lengthenings within the timing pattern concern the smallest notes: the ornamental notes (squares in Figure 3). The clearest relative shortenings also concern the smaller notes: the sixteenth notes in the sixteenth note leap and in the sixteenth changing note (triangles in Figure 3). Intermediate lengthenings of notes occur at sixteenth notes just before phrase boundaries and at other sixteenth note upbeats (see circles in Figure 3). The smallest variations appear at the quarter notes and eighth notes, even when the quarter note is the last note of the phrase (measures 9 and 18). This is surprising, because expressive variations are assumed to increase with the structural importance of the time unit it closes or spans and not to decrease (see e.g. Todd, 1989). In sum, as Palmer (1997) noted, there are small and large variations in the onset timing of which the relatively small variations concern the larger note values, and the relatively large variations concern the smaller note values. If this is a general trend, it is misleading to examine variations in duration on a normalized scale, since the variations in duration of eighth and sixteenth notes may be proportionally large with respect to their average duration, though absolutely they are small.

An alternative way of representing these expressive variations is to calculate IOI's at a certain metrical level, most often the beat level (see e.g. Shaffer, Clarke & Todd, 1985) or the bar level (see e.g. Todd, 1985; Repp, 1992a). In this way, the timing pattern becomes a local tempo indicator, be it one over tempo (large IOI's correspond with low tempi). If there is no melody note onset at an eighth note beat, the IOI between the last onset on an eighth note beat and the next onset of an eighth note beat has to be interpolated (see Figure 4 and Repp, 1992a). This results in the same IOI measurement for several eighth note beats (this spans three eighth notes maximum in the example, see Figure 5, e.g. last three measures of the first half).

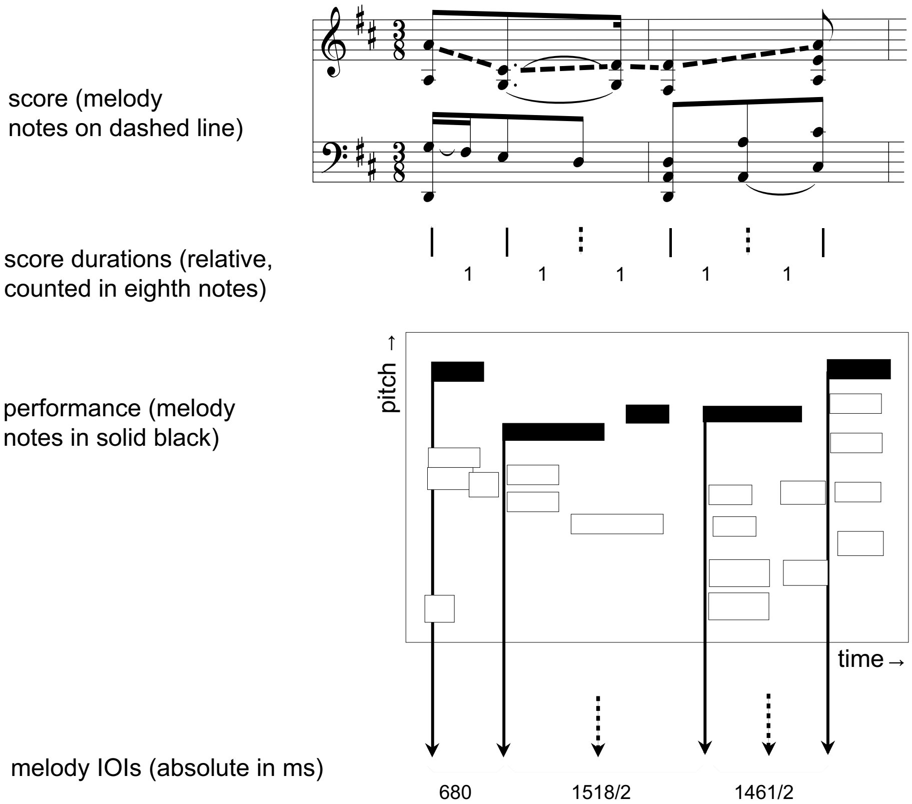


*Figure 4*. Extraction of a regular beat from a piano performance. The score is given at the top, followed by the relative (symbolic) note durations of the melody. Subsequently, the performance data is given as a piano roll notation, indicating how the IOI pattern is derived from this data. The absolute (measured) performance IOIs are given at the bottom. The beat pattern consists of time intervals between note onsets at succeeding beats. If no note onset is present at a certain beat, the interval to the onset at the following beat is interpolated (dotted arrows).

The normalized eighth note IOI pattern of the example (Figure 5) reveals a clear lengthening of the ornamented eighth note in measures 6, 10 and 12, which are preceded and followed by relative short IOI's. Lengthenings also occur in the proximity of phrase boundaries at measures 9 and 17. Locally shortened beats are the third beat of measure 1 (passing note), the second beat of measure 3 (start of sixteenth leap), the second beat of measure 16 (changing chord) and only in the repeat the third beat of measure 14 (sixteenth changing note and upbeat). There are two global variations in duration. A gradual shortening of beats followed by a gradual lengthening of beats that accompanies phrase structure. And a global steady mean duration of beats in the first half and a global decrease of the length of beats in the second half. From this perspective, Palmer's (1997) remark on small and large timing variations might be interpreted differently, in the sense that small variations refer to local shortening and lengthening of notes, while large variations refer to global trends.

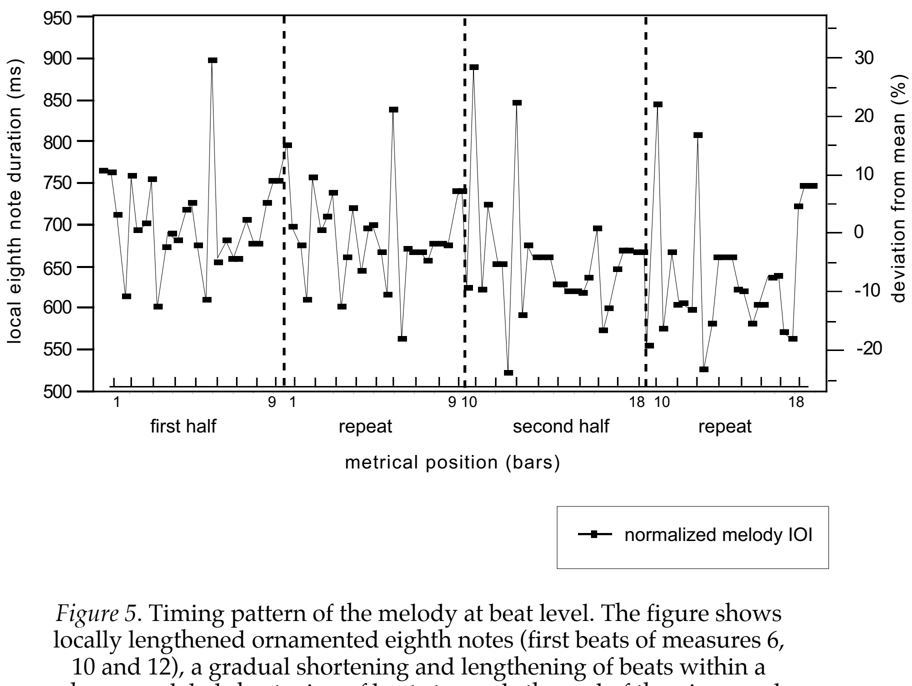


*Figure 5*. Timing pattern of the melody at beat level. The figure shows locally lengthened ornamented eighth notes (first beats of measures 6, 10 and 12), a gradual shortening and lengthening of beats within a phrase, a global shortening of beats towards the end of the piece, and additional local lengthenings and shortenings (e.g. lengthened first beats and shortened second beats of measures 2 and 3).

*Deviation from the norm given by the score*

In practice, the definition of expression as deviations from the score implies that variations in normalized note or beat IOI can be represented as percentages (or fractions) below and above the mean (see e.g. Clarke, 1985; Gabrielsson, 1974), which asks for a re-scaling of the normalized IOI's. The resulting timing patterns are identical to the pattern according to the microstructure definition, besides a scaling of the y-axes (% instead of ms; see Figures 3 and 5).

For the performance example used here, it may be clear that using the mean note duration as a reference is not very sensible. This reference functions over a too wide and global time span and the deviations are too large for the reference to be meaningful. It is unlikely that the duration of beats at e.g. the start and the end of the piece are interpreted in relation to the same average beat duration. Instead, it is desirable to reduce the norm to a more perceptually valid construct that can be derived directly from the performance, which is what Desain & Honing (1991) advocate when they talk about deviations from the norm as given by a higher order unit in a performance, instead of the score.

*Deviation from a norm within the performance*

As an example of the definition of expression as the deviation from the norm as given by a higher order unit, we express the timing of the beat in reference to the timing of the bar. Again, normalized IOI's are used in this measurement. In this representation, the timing pattern shows for each score eighth note the measured beat IOI as a fraction of the measured bar IOI.

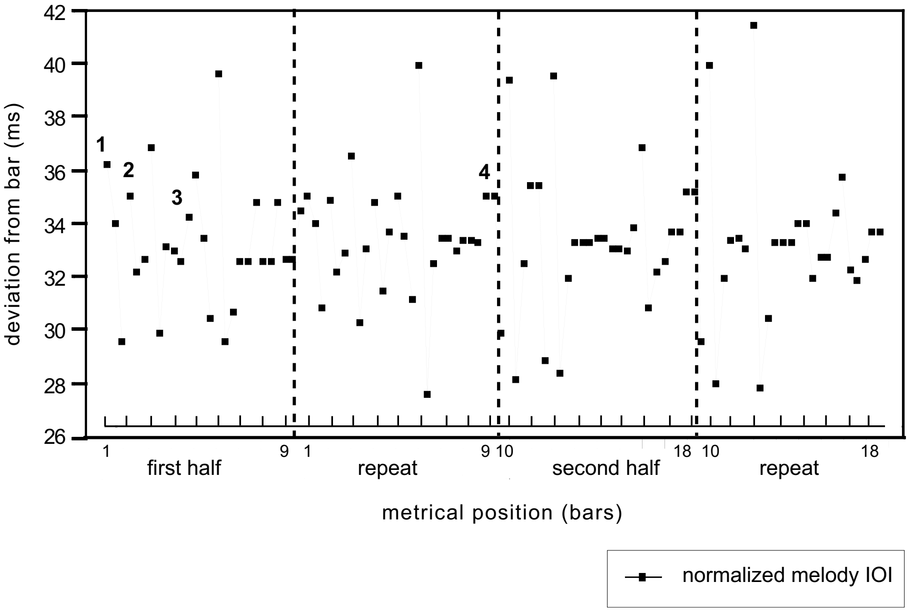


*Figure 6*. Timing pattern of the melody: deviation of beats within bars. The figure shows a systematic alternation of four timing patterns: (1) a gradual shortening of the beats over the measure at the start of phrases (first half); (2) a long first beat and a short second beat, which is characteristic for ternary measures; (3) a long third beat at the end of phrases; and (4) a short third beat at the end of phrases as compensation for a following lengthened first beat.

Figure 6 shows a rather systematic relation between beats within a bar. Four characteristic timing patterns of the beat within a measure alternate in a systematic way, especially in the first half of the piece: (1) a gradual shortening of the beats over the measure, (2) a long first beat and a short second beat, (3) a long third beat and (4) a short third beat.

In the first half, pattern 1 occurs at the beginning of phrases and sub-phrases. Pattern 2 seems to be a typical pattern for a ternary measure (see also Clynes, 1983). This pattern is especially strong in measures with the sixteenth leap on the second and third beats (measure 3) or with a lengthened ornamented eighth note on the first beat (measures 6, 10, and 12). Pattern 3 occurs at the end of a phrase, except in cases in which a highly lengthened eighth note follows the third beat. These last cases show pattern 4. In the second half, the patterning is much less clear because of the many interpolated beat onsets.

To summarize, the different representations of the expressive timing variations showed that it matters at which structural level expressive timing is examined. The normalization of variations in note IOI to a single scale emphasized the rubato on small sub-beat levels and (in this performance) overruled the more global trends. The beat IOI showed both global trends and local variations. It revealed patterns of gradual shortening and lengthening that accompany phrases, and a gradual shortening of beats in the second half of the piece. The local variations in beat IOI were better interpreted in relation to the bar IOI. Within this representation of beat-bar fraction, the local variations turned out to be rather systematic.

In the examples, a change in the scale of representation influenced only the interpretation of the variations. The representation of the variations in note and beat IOI as deviations from the score (that is as deviation from a mean IOI) proved, however, not to be very meaningful, since the mean IOI was not a suitable reference candidate.

In general, the perceptual validity of certain representations was a point in question and an aspect to be improved upon. Although all representations seem to have some perceptual validity, they bias the interpretation in specific ways (cf. Honing, 2001). The different representations emphasized certain aspects of the timing profile, while diminishing or even ignoring others. In this respect, it should be mentioned that most models of expressive timing make use of one type of representation. For example, Todd's (1985) model of expressive timing uses the bar duration representation. In contrast, Clynes' (1983) composer's pulse models the variations of beat IOIs within a higher conceptual unit (bar or hyperbar). And Sundberg et al. (1991a) focus on duration and onset timing variation from note to note.

## Explaining expression

The explanation of an expressive gesture is complicated, because a variety of factors play a role in the accomplishment of a performance. A performer brings to music a variety of sources and perspectives. For example, she/he makes a conceptual interpretation of the music, positions it within a certain stylistic period and practices the movement actions. The question is whether all aspects of the act of performance are of equal importance for the resulting expressive interpretation and whether one may investigate the influence of one aspect without reference to the others. What is the role of each aspect in a performance? To what extent do they explain expressive variations?

The theory that has had most success in explaining the regularities underlying expressive variations is the generative theory of expression (Clarke, 1988). According to this theory, it is the performer's interpretation of the musical structure that generates expressive variations; expression serves to highlight musical structure (see Clarke, 1988; Palmer, 1989, 1996b; Sloboda, 1983). Evidence for the theory is the observation that musicians are able to perform a piece of music in a highly similar way without rehearsing the piece even after a number of years have passed. This suggests that expression is closely related to the performer's mental representation of the musical piece, because it is unlikely that a performer has memorized the music with all details of the performance included. Further evidence comes from experiments that show a direct relation between the interpretation of musical and the expressive variations, such as relative lengthening of notes at the end of a phrase (Palmer, 1992; Todd, 1985) or articulation differences between strong and weak beats of a measure (Sloboda, 1983).

A complication in the analysis of expressive variations with respect to the musical structure is the composition of expressive patterns and the variety of music structural descriptions that could be the source of it. To analyse the structure of an expressive timing pattern, Desain & Honing (1997) propose what is called the structural expression component theory (SECT). This theory is based on the observation that generative models of expressive timing formalize a relation between expression and one specific kind of structure, while a human performance may express an interpretation of several structural aspects at the same time. The theory suggests to integrate existing generative models (Clynes, 1983; Todd, 1989; Sundberg, Friberg & Frydén, 1989) and optimise the parameters of the generalized model to fit the timing pattern of human performances. The resulting fit will reveal the extent to which the components are present in human performances, as such dissecting the expressive signal into its most strongly communicative structural information components.

To decompose an expressive timing profile into different sources of variation was also the aim of Penel & Drake's (1999) psychological segmentation model. Penel & Drake (1999) distinguished between variability due to perceptual bias, variability due to motor noise and variability due to expressive intention. They separated the sources of expression by way of an experimental paradigm. In this paradigm, the musician is instructed to first adjust the durations of a musical performance to become strict in tempo. The next step is to perform the music in a mechanical way, without expression, and finally to perform it with expression. Each deviation of the note durations from score durations that resulted from the adjustment task was interpreted to be a compensation for perceptual biases caused by the musical material. For example, the last note of a group of short notes is perceived as relatively short and therefore lengthened to sound equally long. The variations added in the second condition (i.e. mechanical performance) with respect to the first condition (i.e. perceptual adjustment) were interpreted to be due to motor constraints or motor noise. The variations added in the third condition (i.e. expressive performance) were interpreted as intended expression.

The distinction between variations due to expressive intention on the one hand and those due to motor noise and perceptual bias on the other hand may seem evident at first sight, but is not made explicitly by all expressive performance researchers. For example, Sundberg *et al.* (1991) model the deviations of expressive timing from mechanical timing without making a distinction between variations that compensate for perceptual biases of musical structure and variations that are intended. Likewise, the variation due to motor constraints is not always interpreted as unintentional. In fact, an embodied quality may be part of the performer's aesthetic. As an example, Windsor, Aarts, Desain & Timmers (2001) reported the intention of a pianist to perform a descending leap with time delay to imitate the constraints that singers encounter in performing the leap. In addition, there are several models that actually see the imitation of physical movements as the main source of expression (Kronman & Sundberg, 1987; Todd, 1992, 1995). So, in conclusion, it is very likely that the encoding of movements (of the performer or otherwise) also attribute to the expressiveness of a performance.

Even if the distinctions between intentional, motoric and perceptual cannot be drawn so sharply, they are nevertheless useful concepts. The perception task discussed in Penel & Drake's (1999) may distinguish perceptible from imperceptible variations, with the perceptible variations being the important variations for the performer's expression. The variations due to motoric constraints may be intentional if they are reiterated in repeated performances, or may be unintentional if they occur relatively at random (motor noise). In both cases, the motor condition should reveal motorically rather than conceptually governed variations. Palmer & Meyer (2000) cleverly differentiated between conceptual and motor constraints by asking pianists to perform different versions of a simple and short tonal melody as fast as possible with the same hand or with different hands and by examining the effect of change of hand on the speed of the production. If there was a large effect of change of hand, motor constraints played a relatively large role. Less experienced pianists were greatly influenced by motor constraints, while experienced pianists only showed effects of conceptual variations (see Palmer & Meyer, 2000).

In addition, two other explanations of expressive variations have been proposed. The first is expression as communication of emotional content, which is advocated by for example Gabrielsson & Juslin (1996), who found that the intention of musicians to perform a western classical piece “sadly” or “happily” influences the timing, intensity and timbre variations of the performance (independent of structure). They also found that listeners were generally able to perceive the expressed emotion, at least if relatively basic emotions were concerned.

Another possibility is that expressive variations comment on the musical structure and provide it with additional characterization. This possibility is advocated by Shaffer (1992) and Clarke (1995), who introduced this possibility as an additional communicative intention, besides the communication of musical structure. According to Shaffer (1992), only the patterning of expressive variations is determined by the structural characteristics of the music, while the shape of expressive gestures and the choice of expressive features are a function of musical character.

How exactly the variety of expressive intentions is communicated to the listener is still quite unclear. A first proposal was made by Juslin, Friberg & Bresin (2002) who developed a generative model consisting of an addition of distinct variations due to separate sources. In a first version of this model, emotional expression affects the average tempo, while variations due to the generative expression of musical structure affect local tempo. Variations that relate to motor movements affect the shape of the variations related to the musical structure. Variations due to motor or clock noise are random and have a very small extent.

Timmers et al. (2000) also showed that although pianists may agree on a certain interpretation of the musical structure, they show clear differences in their use of tempo rubato. For example, the performer of the example detailed in the previous section performed with “give and take” rubato and made considerable use of asynchrony, while a second pianist performed with gradual changes in tempo and showed very little asynchrony. The extent of the rubato was further shown to be a variable that changes with musical texture. A perceptual study (Timmers, 2002) showed the importance of global features of a performance as rubato extent, average articulation, use of dynamic shaping and use of asynchrony in characterizing a pianist's interpretation.

To summarize, the multi-tasking of a performer consists of (at least) (1) a conceptual interpretation of the structure of the musical piece that affects the way the music is performed, (2) a planning of movements that may cause additional variations or may shape the variations, and (3) an interpretation of the musical emotion and/or the musical character. In addition, the performer may compensate for duration and loudness accents that are caused by the musical structure and are inappropriate. The performer may also comment on or add to the musical structure, instead of only expressing it. These comments could take the form of shaping the expressive pattern differently, accenting ambiguity or characterizing the musical material.

## Diversity/commonality

One of the characteristic features of Western performance practice is that a single piece of music (e.g. Brahms' Variation Op. 12 No. 1) is performed over and over, by different performers and/or by the same performer at different times. These different renditions of a piece relate to each other in the sense that they render the same piece of music, and in the sense that several performance features are roughly the same. For example, slowing down at the end of phrases and melody lead are general phenomena (Palmer 1996b). However, as a result of the performer's expression and interpretation of a musical piece, and of differences in the acoustical characteristics of the instrument and the room etc., the diversity between several renditions of a piece is considerable as well. This diversity concerns for example the choice of global tempo, the specific intensity levels of notes, choices of articulation, etc. The question is how these different renditions relate to the musical piece and to each other. Are they merely attempts to realize a prototypical “ideal” performance? Or do they have independent validity? If so, how should we differentiate between representative performances and noise? If we construct theories, should they be based on dozens of performances, or can they be more specific? Can we explain the differences, or is there more freedom and variety to it?

In performance literature several approaches towards this diversity of performances are in use:

1) Only a small number of performances of a piece of music are analysed, and the performance features are taken as examples of expert behaviour (see Clarke, 1995; Desain & Honing, 1994; Palmer, 1996a). When there exist repeated performances of a single musical fragment, the consistency of characteristics can be analysed between repetitions. The idea behind this approach is that expert behaviour is interesting in itself, at least when consistency within this behaviour is shown. These studies suggest that relationships within a single performance are important, meaningful and specialized.

2) Performances are classified into groups. Within a single group, performances are expected to have similar characteristics and measurements are averaged. For example, motor and memory constraints are expected to influence the performance of beginners more than the performance of experienced players, while the latter are

expected to behave in a more conceptually driven manner (Palmer & Meyer, 2000). The idea is that people who fall into a single group behave in approximately the same way, while the behaviour of people from different groups can easily be differentiated. Traditional group divisions are made along the lines of gender, age and experience. Especially the difference between the musically experienced and inexperienced is often made.

3) The analysis concerns several performances of a single piece. A grand average is made over a considerably large group of performances, assuming that by averaging over a sufficiently large set common characteristics are amplified while disagreements are weakened or averaged out. Repp (1992a) uses a grand average timing profile (i.e. measured note IOI patterns of hundreds of performances) that contains common timing characteristics, to which individual performances can be compared.

4) Common and distinct features of different performances are detected and their relatedness is formalized. For longer stretches of music this is done by a principal components analysis, while curve fitting is used for brief musical fragments. In this last case, expressive shapes are described by mathematical functions with adjustable variables (for both methods, see Repp, 1992a).

Now we return to the question how different renditions relate to the musical piece and to each other. In our view, the challenge is to balance the perspective of a representative performance that exemplifies a prototypical performance, such as used in the third approach, and the view of all expert performances having their own identity and contributing a specific expressiveness, such as taken in the first approach. What we think is needed is a clear insight into the relation between different renditions of a piece and into the way differences between performances are categorized. This aim relates to the aims of the second and the last approach. For a successful second approach, different lines to group performances are however needed than the conventional groupings such as age, gender and experience, because too many differences exist within the group of experienced performers of a single culture, with (approximately) the same age and similar music education. The last approach could help defining these lines. This approach can be characterized as open and explorative with respect to the kinds of differences between performances. Especially the principal component – analysis – is data-driven and explores the number of components needed to explain the variability within the data. This number of components reflects the diversity within, for example, timing patterns. The curve fitting assumes (and often confirms) that there are patterns that underlie all performances. In Repp (1992a), the only differences allowed consist of the degree of emphasis of the patterns, which causes a limitation of the diversity to one degree of freedom.

The use of a Minskyan representation framework may be helpful in illustrating the relation between different renditions of a piece. This knowledge representation framework suggests that long-term memory knowledge can be captured in units (called frames) that contain nodes and relations between nodes (Minsky, 1975). Incoming events activate a certain frame, which in turn activates other frames.

The terminals of frames point towards objects, persons or other frames. (However, note that there are restrictions on the modularization of musical knowledge; see Honing, 1993). Where musical performances are concerned, we may assume a general scheme for Romantic piano music, potentially with a specialization for Brahms' music and even for Brahms' Variation Op. 21. Different performances of this piece may now be said to highlight different aspects of the knowledge representation frame, with the addition that they share most terminals (same notes, same object categories). In other words, the relationships may change but the notes remain the same. More specifically, performers may agree on the structural interpretation of a piece but have different strategies to express the interpretation (such as "use tempo variation" or "use dynamic accents"). On the other hand, performers may have the same strategies but differ in their opinion about the important structural aspects of the music (such as "use melody lead", but differ with respect to the choice of melodic lines). In the first case, it is as though a room were seen from a certain perspective in the dark and then in light: it would have the same walls and objects, but different colour shading. In the second case, it is as though a room were seen from different perspectives: the same walls and objects within the room would get different measures on given dimensions. The impact of the first might be as large as that of the second, at least if art is concerned (as exemplified by different versions of the same still life).

If we want to group related performances and define group boundaries between qualitatively different performance aspects, we could take advantage of some guidelines indicated by previous research on the categorization of objects and object attributes (e.g. Rosch & Mervis, 1975). Objects are categorized according to the commonality of attributes within the category and the diversity of attributes between categories. Some objects are prototypical of the category, while the inclusion of others is ambiguous. There is a hierarchy of categories, but it is not strict: an object may belong to several higher order categories (multiple inheritance). The attributes of objects (e.g. colour or shape) are also subject to categorization. In a similar way, the example performance falls within the category "music" (a relative basic category) and within the sub-category "romantic music by Brahms". Meanwhile, the example is an instance of a variation, and, more specifically, an instance of Variation Op. 21 by Brahms. It is also an instance of a performance by pianist X in the Y style. This performance may be more/less representative of this pianist and of this style. The performance aspects themselves may, for example, be categorized as "slow" or "fast" or be found to speed up or to drag. In addition, specific musical fragments may be performed in several characteristic ways. Listeners may learn to recognize and categorize such specific treatment. For example, the performer of the example may be recognized by his performance of arpeggio chords in a rather brisk and fast way, in contrast to the other characteristic treatment of arpeggio chords, which is round and fluent.

To conclude this section, the variety in performances of musical pieces raises the question of meaningful differences and similarities between performances and the relevant relationship between performance characteristics and musical structure. The approaches to rendition differences that are used in performance studies range from the study of single performances to the averaging over dozens of tempo patterns belonging to equally many performances. Our preference is first to explore and categorize the relationship between the performances. We propose to do this along the lines of object categorization and with the aid of a knowledge representation framework. We reject the idea of a single ideal performance of a musical piece in favour of the concept of a network of performances that are equal at one level and different at another level of categorization.

## Discussion

In the introduction, we stated that the emphasis on music as a product of human perception and production raises the issue of the complexity and transformability of music. The identity of music is not a stable concept, but changes with use and practice. In the previous discussion about the expressive component of musical performance, we indeed saw that performers have considerable creative input in the interpretation and shaping of music. This leads to a variety of renditions of a musical piece that may have some consistency within a specific period of time, but which may change considerably over time (see e.g. the changing choice of tempi in performances of Beethoven's symphonies). If we include expression as a constituent part of a musical piece, the identity of that piece is obscured. What is *the* piece if several categorically different versions exist? It may be the most prototypical performance as identified by the majority of people. The concept may also be more flexible than that. While some stylistic features and characteristics of the piece remain, others change or are specified by the performance. For example, performance may not affect the style forms in relation to which the music is interpreted (e.g. the presence of harmony, melody and rhythm). Also the style structure of a piece and its period may remain unaffected (e.g. the kind of chord progressions, melodic motives, rhythms). Performance will, however, affect the idiostructure of the piece, which is the typical structure for that piece. For example, performance may change the inclusion of notes within groups, or it may change the function of notes by stressing the melodic function above the harmonic function of, for example, tenor notes. Also, the rhythmic function of notes is easily varied by a change in global tempo or *tactus* level (e.g. from downbeat to upbeat). If we interpret the identity of a piece within a knowledge representation framework, it should be clear that a representation of a musical piece is flexible to the extent that it incorporates variations in the relation between notes on different musical dimensions. A representation framework of a musical piece may have some default settings as far as harmonic, melodic and rhythmic relations are concerned. It also contains some assumptions about the relative speed and dynamic of notes. In Minsky's model, however, default settings are easily displaced without losing the link to higher order representations or lower level terminals.

The role of notation and its relation to performance is an issue closely related to the issues raised in this paper. The discussion on the definition of expression made clear that musical notation does not represent all characteristics of a musical performance, and – more importantly – that musical notation does not predict or imply all aspects of a performance. Instead, musicians add certain aspects of a performance that are undefined in the score, and deviate from other aspects such as the indicated duration ratios in the score. It also became clear that the score should not be overestimated as a norm based on which the performance aspects can be interpreted.

The notion of musical structure was indeed important as a reference for performance. This notion is not equal to music as represented in a score, but it is closely related. An annotated musical score that explicitly indicates phrase structure, metrical structure, or harmonic structure represents the musical structure of a classical piece fairly well. The difference between common music notation and musical structure is that the latter contains more information and is more flexible. A representation of musical structure varies among interpreters and performers. For example, the pitches and rhythm ratios may be fixed, while the grouping and metrical interpretation vary.

The function of an annotated score is one of reference or explanation. It functions as a reference for the location and interpretation of performance aspects. In addition, it sometimes explains an expressive gesture, in the sense that it provides the *raison d'être* (as in the case of the communication of phrase structure). It should be noticed, however, that musical structure (or an annotated score) does not explain all aspects of a performance. This is because, first, the performer also shapes the character and mood of the performance, and second, this shaping is further influenced by context-dependent factors. There is too much freedom within the realization of a performance to relate *all* performance aspects to structural interpretation of music.

## Conclusion

An empirical musicology of performance deals with performers' expression and multisidedness in such a way that expressive features are examined from an informed standpoint and on different structural levels. Experimental manipulation and computational modelling are necessary to separate the sources of expression. A performance-based musicology analyses and categorizes performances, and defines a system to classify rendition differences. It uses annotated scores to represent musical structure and show the communicative function of the expressive behaviour. The score does not function as a norm; instead, if annotated, it is used as reference.

## References

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

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. Oxford: Science Publications, pp. 1-26.

Clarke, E. F. (1995). Expression in performance: generativity, perception and semiosis. In J. Rink (ed.), The Practice of Performance: Studies in Musical Interpretation. Cambridge: CUP, pp. 21-54.

Clynes, M. (1983). Expressive microstructure in music, linked to living qualities. In J. Sundberg (ed.), Studies of music performance. Stockholm: Royal Swedish Academy of Music, pp. 76-181.

Desain, P., & Honing. H. (1991). Towards a calculus for expressive timing in music. Computers in Music Research, 3, 43-120.

Desain, P., & Honing. H. (1994). Does expressive timing in music performance scale proportionally with tempo? Psychological Research, 56, 285-292.

Desain, P., & Honing. H. (1997). How to evaluate generative models for expression in music performance. In Issues in AI and Music Evaluation and Assessment. International Joint Conference on Artificial Intelligence, 5-7. Nagoya, Japan.

Epstein, D. (1995). Shaping Time; Music, the Brain and Performance. New York: Schirmer Books.

Gabrielsson, A. (1974). Performance of rhythm patterns. Scandinavian Journal of Psychology 15, 63-72.

Gabrielsson, A. (1987). Once again: the theme from Mozart's Piano Sonata in A major: A comparison of five performances. In A. Gabrielsson (ed.), Action and Perception in Rhythm and Music. Stockholm: Royal Swedish Academy of Music, pp. 81-103.

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 (1), 68-91.

Honing, H. (1993). A microworld approach to the formalization of musical knowledge. Computers and the Humanities, 27, 41-47.

Honing, H. (2002) Structure and interpretation of rhythm and timing. Tijdschrift voor Muziektheorie [Journal of Music Theory], 7 (3), 227-232.

Honing, H. (2001) From time to time: The representation of timing and tempo. Computer Music Journal, 35 (3), 50-61.

Juslin, P. N., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM Model. Musicae Scientiae special issue 2001-2002, 63-122.

Kronman, U., & Sundberg, J. (1987). Is the musical ritard an allusion to physical motion? In A. Gabrielsson (ed.) Action and Perception in Rhythm and Music. Stockholm: The Royal Swedish Academy of Music.

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

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (ed.) The Psychology of Computer Vision. New York: McGraw Hill Book Company, pp. 211-277.

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

Palmer, C. (1989). Mapping Musical thought to musical performance. Journal of Experimental Psychology: Human Perception and Performance, 15 (12), 331-346.

Palmer, C. (1992). The role of interpretive preferences in music performance. In M. R. Jones & S. Holleran (ed.) Cognitive bases of musical communication. Washington: American Psychological Association, pp. 249-262.

Palmer, C., & van de Sande, C. (1993). Units of Knowledge in Music Performance. Journal of Experimental Psychology: Learning, Memory and Cognition, 19, 457-470.

Palmer, C., & van de Sande, C. (1995). Range of planning in skilled music performance. Journal of Experimental Psychology: Human Perception and Performance, 21, 947-962.

Palmer, C. (1996a). Anatomy of a Performance: Sources of Musical Expression. Music Perception, 13, 433-454.

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.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. Psychological Science, 11, 63-68.

Penel, A. & Drake, C. (1999). Seeking “one” explanation for expressive timing. In S. W. Yi (Ed.), Music, Mind & Science. Seoul: Seoul University Press, pp. 271-297.

Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by 19 famous pianists. Journal of the Acoustical Society of America, 88, 622-641.

Repp, B. H. (1992). Diversity and commonality in music performance - an analysis of timing microstructure in Schumann's Traumerei. Journal of the Acoustical Society of America, 92 (5), 2546-2568.

Repp, B. H. (1998). Obligatory 'expectations' of expressive timing induced by perception of musical structure. Psychological Research, 61 (1), 33-43.

Rink, J. (1995). *The Practice of Performance: Studies in Musical Interpretation*. Cambridge: CUP.

Rosch, E., & Mervis, C. (1975). Family Resemblances: Studies in the Internal Structure of Categories. Cognitive Psychology, 7 (4), 573-605.

Shaffer, L. H., Clarke, E., & Todd, N. P. (1985). Meter and rhythm in piano playing. Cognition 20, 61-77.

Shaffer, L. H. (1992). How to Interpret Music. In M. R. Jones & S. Holleran (ed.) Cognitive bases of musical communication. Washington: American Psychological Association, pp. 263-278.

Seashore, C. E. (1938). Psychology of Music. New York: Dover.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. Quarterly Journal of Experimental Psychology, 35 (A), 377-396.

Sundberg, J., Friberg, A., & Frydén, L. (1989). Rules for Automated Performances of Ensemble Music. Contemporary Music Review, 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, and I. Cross (eds.), Representing Musical Structure. London: Academic Press, pp. 161-197.

Timmers, R., Ashley, R., Desain, P. & Heijink. H. (2000). The influence of musical context on tempo rubato. Journal of New Music Research, 29 (2), 131-158.

Timmers, R. (2002). On the contextual appropriateness of performance rules. In R. Timmers, Freedom and constraints in timing and ornamentation; investigations of music performance. Maastricht: Shaker Publishing, pp. 85-109.

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

Todd, N. P. (1989). A Computational Model of Rubato. In "Music, Mind and Structure", edited by E. Clarke and S. Emmerson. Contemporary Music Review, 3 (1), 69-88.

Todd, N. P. (1992). The dynamics of dynamics: a model of musical expression. Journal of the Acoustical Society of America, 91 (6), 3540-3550.

Todd, N. P. (1995). The kinematics of musical expression. Journal of the Acoustical Society of America, 91, 1940-1949.

Windsor, W. L., Desain, P., Aarts, R., Heijink, H., & Timmers, R. (2001). The timing of grace notes in skilled musical performance at different tempi: a case study. Psychology of Music, 29, 149-169.