Psychol Res (1998) 61: 12–32

ORIGINAL ARTICLE

Amandine Penel · Carolyn Drake

Sources of timing variations in music performance: A psychological segmentation model

Received: 22 November 1996 / Accepted: 9 June 1997

Abstract Systematic timing variations observed during music performance have usually been attributed to a *musical expression hypothesis*, related to relatively high-level processes, by which musicians emphasize certain events in order to transmit a particular musical interpretation to the listener. We propose, in addition, a *perceptual hypothesis*, related to lower-level processes, in which some observed variations would be related to functional constraints of the auditory system. (Some intervals would be heard shorter and thus played longer as a phenomenon of perceptual compensation.) We present a psychological model of temporal organization proposing two types of process (regularity extraction and segmentation into groups) operating parallel that allow listeners to parse complex auditory sequences such as music. Each type of process operates at both a low processing level (beat extraction and segmentation into basic groups) and a higher processing level (hierarchical metric organization and hierarchical segmentation organization).

The analysis of musical and mechanical performances of Schumann's *Träumerei* demonstrated performance variations in relation to both hierarchical segmentation and hierarchical metric organizations, and to rhythmic groups. Variations were not systematically observed in relation to melodic groups. Regression analyses quantified these effects and demonstrated that hierarchical segmentation and rhythmic groups accounted for approximately 60% of the variance for musical and mechanical performances, leaving room for the description of other, as yet unidentified, processes. The percentage of variance explained by high-level processes (hierarchical segmentation) decreased from musical to mechanical performances, whereas the percentage of variance explained by lower-level processes (rhythmic groups) increased.

We conclude that it is important to go beyond the traditional approach of describing performance variations in relation to musical structure and to adopt the approach of studying performance variations in relation to the psychological processes that allow the musician to perceive the musical structure. Finally, we adapt the psychological model of temporal organization to expressive timing: similar psychological processes operate at multiple hierarchical levels – namely, those of segmentation and grouping –, and these similar processes result in the same pattern of performance variations (an accelerando/ritardando profile).

A. Penel (✉) · C. Drake
Laboratoire de Psychologie Expérimentale,
Université René Descartes, 28, rue Serpente,
75006 Paris, France
Tel: 01 40 51 98 65; e-mail: penel@idf.ext.jussieu.fr

Introduction

Sources of performance variations

The analysis of any musical performance reveals that a piece is never played exactly as it is written in the score: performers vary several physical parameters. Studies have concentrated on variations in timing (Clarke, 1985, 1988; Desain & Honing, 1994; Gabrielsson, 1987; Palmer, 1989; Repp, 1990; Shaffer, 1981), intensity (Nakamura, 1987; Repp, 1995c; Sloboda, 1985a; Todd, 1992), and articulation (corresponding to legato or staccato, Palmer, 1989; Repp, 1995b; Sloboda, 1985a), with some studies devoted to pitch (in the case of a non-fixed pitch instrument like voice or violin), timbre, and the use of the pedal (Palmer, 1996a; Repp, 1996b). In this paper, we study only timing variations – in particular, the duration between the onsets of successive events (inter-onset intervals: IOI).

Frequently, these systematic performance variations have been attributed to the performer's wish to transmit a particular musical message to the listener (Clarke, 1985, 1988; Palmer, 1989, 1996b; Repp, 1992a, 1992b, 1995c; Sloboda, 1985b). This *musical expression hypothesis* refers to relatively high-level processes involving a global analysis of information which requires musical knowledge: performers' intended interpretations (e.g., their interpretations of the musical structure) modify the way in which they perform the music, thus influencing the listeners' mental representations (i.e., their representations of the musical structure). In this way, listeners have access to a particular interpretation of the piece and to a better understanding of it, due, for example, to the resolution of structural ambiguities. One example of temporal variations used to highlight structural features of a piece involves the tendency for musicians to slow down at the end of phrase boundaries (known as phrase-final lengthening), a phenomenon reflecting high-level temporal segmentation processes which allow performers to transmit a particular phrasing to the listener (Palmer, 1989; Todd, 1985). Similarly, notes in the melody usually precede other non-melody notes in a chord (known as melody lead), a phenomenon reflecting a horizontal organization into melodic and non-melodic voices which enables pianists to transmit their perceived melodic interpretation to listeners (Palmer, 1989, 1996a, 1996b; Palmer & van de Sande, 1993; Repp, 1996a).

However, the musical expression hypothesis cannot, by itself, explain all observed variations. In addition, the *perceptual hypothesis* was proposed, which would work parallel to the musical expression hypothesis (Drake, 1993b). This is similar to the "bottom-up" hypothesis proposed by Repp (1995c). The perceptual hypothesis suggests that some performance variations are related to lower-level processes involved in the local analysis of surface acoustic features of the music – in particular, to the grouping of several notes into perceptual units: some events would be played longer because they are perceived shorter (a phenomenon of perceptual compensation). We suggest that these performance variations would result from unavoidable and universal mechanisms related to constraints of the auditory system (ten Hoopen, Hartsuiker, & Sasaki, 1995). In this paper we examine two candidates for such lower-level processes – segmentation into rhythmic groups and melodic groups.

A first source of evidence in favor of some performance variations being related to lower-level processes emphasizes their unavoidable nature. The musical expression hypothesis predicts that variations should be under the musicians' voluntary control and thus should disappear if they so desire. However, when pianists were asked to perform simple musical sequences in either a musical or a mechanical fashion, many high-level timing variations described above were reduced in the mechanical performances, but the timing variations corresponding to rhythmic grouping hardly decreased (Drake & Palmer, 1993).

Another source of evidence in favor of some performance variations being related to lower-level processes emphasizes their universal nature. The musical expression hypothesis predicts that variations should become increasingly systematic with increased age and musical training. However, the comparison of the reproduction of musical rhythms by groups of participants varying considerably in levels of passive musical acculturation (age) and musical expertise (training) indicated that the same systematic temporal variations related to rhythmic grouping were observed for 5-year-old children, adult non-musicians, and musicians. More precisely, in a series of tones separated by short intervals (rhythmic group), the last interval was systematically lengthened by 10%–20% in relation to the preceding interval (Drake, 1990, 1993c).

Evidence that these performance variations have a perceptual origin has been provided by a series of experiments (Drake, Botte, & Gérard, 1989; Drake, 1993b) using various perceptual tasks (temporal adjustment, temporal irregularity detection, and identification) which reduced motor factors to a minimum. These studies demonstrated that the detection of a temporal decrement was easier (and detection of an increment harder) in the last short interval in a rhythmic group, compared with preceding intervals. This temporal distortion was in the range of about 10%. These results lead us to reject the possibility that these variations could be explained by motor factors: performers accented the last note by playing it louder, a process that took longer than playing unaccented notes because the finger had to be lifted higher (see Semjen, Garcia-Colera, & Requin, 1984). We therefore retained the perceptual hypothesis for temporal variations related to rhythmic groups: the last interval was perceived shorter, so it was played longer in order to make it sound regular compared to other intervals. For melodic groups, results appeared less clear (Drake, 1993b). We consider that variations associated with melodic groups may be related to an intermediate processing level, higher than that associated with rhythmic groups and lower than that associated with hierarchical segmentation into phrases.

Converging support for the perceptual hypothesis can be found in a series of studies by Repp (Repp, 1992a, 1992c, 1995c) which examined listeners' abilities to detect temporal and intensity changes in performances of simplified and real music (although Repp interpreted his data in a different way[1]). For temporal changes, the

[1] In this series of studies, Repp contrasted two hypotheses to explain variations in the detectability profiles, a top-down hypothesis which suggests that they may reflect listeners' expectations of expressive performance microstructure, and a bottom-up hypothesis which suggests that they may be due to psychoacoustic stimulus factors. Thus, he considers that correlations between the detectability and the performance profiles reflect top-down (expectations) processes involved in the perception of temporal and intensity changes. We think that these correlations may also reflect bottom-up processes which would also be involved in expressive performance. For example, the last interval of a rhythmic group is played 10% longer than the preceding one, because when played mechanically, it sounds too short (as a result of the low-level process of rhythmic grouping). Perceptually, based around a mechanical performance, a lengthening is difficult to detect (and a shortening, easy). Thus, this low-level process also results in a correlation between the detectability and the expressive performance profiles. Interestingly, Repp never includes the duration of notes (responsible for rhythmic grouping) in his psychoacoustic stimulus factors, but always factors related to pitch and intensity

increment detection accuracy and false alarm profiles were significantly correlated with each other and with the temporal microstructure of expert performances. Interpreted in our theoretical framework, this means that actual lengthenings are easy to detect for certain intervals because they sound a priori (without actual lengthening) longer than others, and these intervals are played shorter. Hence, when played mechanically by a computer, intervals are not all perceived as the same length: some are perceived too short, both by listeners and performers. A lengthening of these intervals, therefore, results in the perception of regularity, and so listeners have difficulties detecting these lengthenings. In the same way, performers compensate for this perceptual distortion by playing these intervals longer. More recently, Repp (in press) examined the detection of both increments and decrements in the initial phrase of Chopin's *Etude*, Opus 10, No. 3. In Exp. 1, he found that when the detectability of increments was high, the detectability of decrements was low, and vice versa. This fits quite well with the hypothesis that some intervals are perceived shorter (or longer) than they are played, which makes a lengthening difficult (or easy) to detect and a shortening easy (or difficult) to detect.

So far, we have limited the discussion to temporal variations. A different pattern of results appears for intensity variations. Intensity variations appear to be more under voluntary control, as they almost completely disappeared in mechanical playing (Drake & Palmer, 1993). They are not universal, as patterns varied considerably, depending on levels of musical acculturation and training (Drake, 1990, 1993c). Finally, the detection of intensity variations appears to be much less influenced by grouping phenomena than the detection of temporal variations, as no systematic perceptual intensity distortions were observed (Drake, 1993b). Similarly, a much weaker relation was found between the intensity detection accuracy profiles and the intensity profiles of expressive performances than for temporal profiles (Repp, 1992c, 1995c).

To summarize this first section, the traditional musical expression hypothesis emphasizing the role of high-level processes linked to musical structure (such as phrase, metric, or voice structure) can account for some of the observed variations. However, we are convinced that some variations cannot be explained without the perceptual hypothesis, which emphasizes the role of lower-level processes linked to acoustic surface features, at least for the temporal variations. (This does not seem to be the case for intensity variations.) Previous studies have mainly highlighted relatively high-level processes at the origin of the temporal variations observed in musical performance. Hence, the first main aim of this study was to show that, in the same normal musical context, temporal variations are governed by both high- and lower-level processing. We now propose a theoretical framework which describes these high- and lower-level processes and the relations between them.

A psychological model of the temporal organization of auditory sequences

Research in cognitive psychology has proposed psychological mechanisms enabling listeners to organize complex auditory sequences. Sequential processing of the temporal structure of auditory sequences must, obviously, involve the analysis of individual event durations, but, more importantly, each event must be situated in relation to surrounding events. Traditional concatenation models propose the maintenance in memory of the characteristics of each event, which would result, in the case of long sequences, in memory overload. Several processes have therefore been proposed (Drake, 1990; Botte, McAdams, & Drake, 1995) which result in the perceptual organization of information contained in a sequence, thus reducing memory constraints. We will describe two temporal organization principles – regularity extraction and segmentation into groups –, each involving a high and a lower processing level. A summary of these organizational principles is presented in Fig. 1.

Temporal regularity extraction

The lowest level in this organization involves the extraction of temporal regularities at the beat level (Parncutt, 1994), often referred to as *beat extraction* (level A2 in Fig. 1). Once a regular beat has been identified, each event can then be situated, in relative terms, to neighboring events. At a higher processing level, the *hierarchical metric organization*, perceived beats can be situated into a hierarchical structure, whereby several beats (usually two, three, or four) are incorporated into larger units (level A3), and these units can themselves be incorporated into increasingly larger units (levels A4, A5, and A6). Subdivisions (level A1) of the beat are also possible (Drake, 1997; Povel, 1985). In Western tonal music, this type of organization corresponds to the metric structure, involving beats, measures, and hyper-measures. Perceptual and tapping tasks indicate that listeners show evidence of beat extraction from an early age (Baruch & Drake, in press) and of two hierarchical levels (Drake, 1997; Palmer & Krumhansl, 1990), but the use of more hierarchical levels is less clear and seems to be restricted to musicians (Bamberger, 1980; Drake, 1993a, 1993c). Performance variations associated with metric structure seem to depend on the musical context, varying from one study to the next: notes corresponding to the first and third beats of a 4/4 measure were lengthened compared to other notes (Shaffer, 1981); notes corresponding to the last beat of either 4/4 or 6/8 measures were lengthened compared to others (Drake & Palmer, 1993); notes at strong metrical positions were played louder and more legato than notes in less prominent metrical positions (Sloboda, 1983, 1985a; Drake & Palmer, 1993).

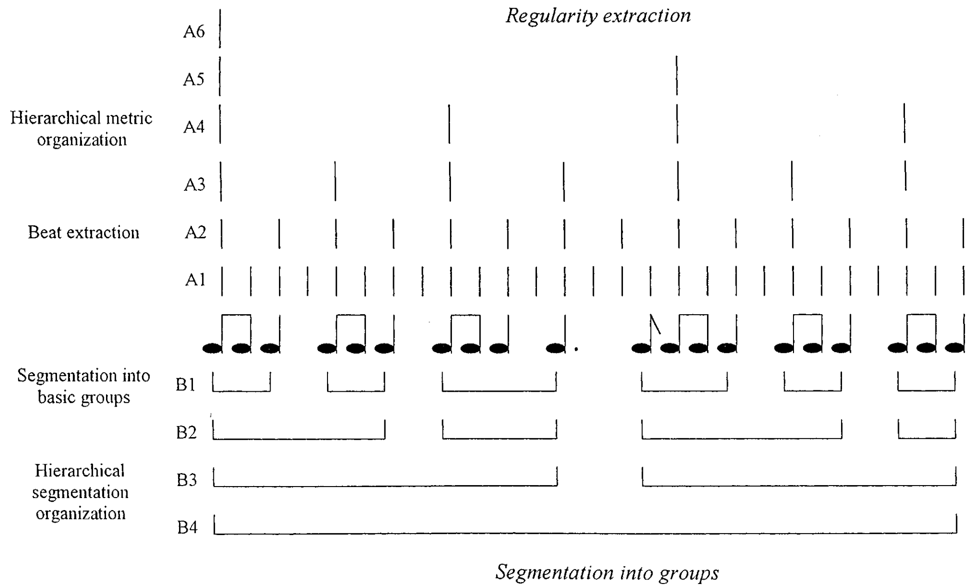


Fig. 1 Two temporal organization principles – regularity extraction and segmentation into groups –, each involving a high (hierarchical metric organization and hierarchical segmentation organization) and a lower processing level (beat extraction and segmentation into basic groups) (from Botte et al., 1995)

Segmentation into groups

The lowest level in this type of organization involves the *segmentation into basic groups*. Many studies (for example, Handel, 1989; Vos, 1977; Thomassen, 1982) have shown that listeners segment a sequence according to its surface characteristics (event duration, pitch, intensity, timbre, etc.): a change in any sound parameter leads to the perception of a break in the sequence and thus to the creation of groups separated by the changes (level B1 in Fig. 1). Perceptual evidence in support of the process of segmentation and grouping is provided by the gap detection literature: it is easier to detect a temporal lengthening within a group than between a group. This process may be universal, as similar patterns of results are observed for adult musicians and non-musicians (Fitzgibbons, Pollatsek, & Thomas, 1974) and 6- to 8-month-old infants (Thorpe & Trehub, 1989). Such segmentation processes into basic groups also appear to operate in musical sequences (Deliège, 1987, 1990). In the present study, we investigate performance variations in relation to two types of basic groups proposed previously (Drake, Dowling, & Palmer, 1991; Drake & Palmer, 1993): rhythmic and melodic groups (see Fig. 2), which appear to play an important role in the reproduction of musical sequences. A rhythmic group is a series of events separated from those surrounding them by longer durations or pauses. A melodic jump is defined as an abrupt pitch jump of five or more semitones (a stricter criterion than the three semitones used by Drake & Palmer, 1993). A melodic turn occurs at changes in pitch direction. Systematic performance variations were observed in relation to both rhythmic and melodic groups in three musical contexts. Concerning temporal variations, for the rhythmic groups the last short interval was played, on average, 10% longer than the preceding one. For both melodic jumps and turns, the interval before the jump or turn was played, on average, 3% longer than the following one. These temporal variations are summarized in Fig. 2.

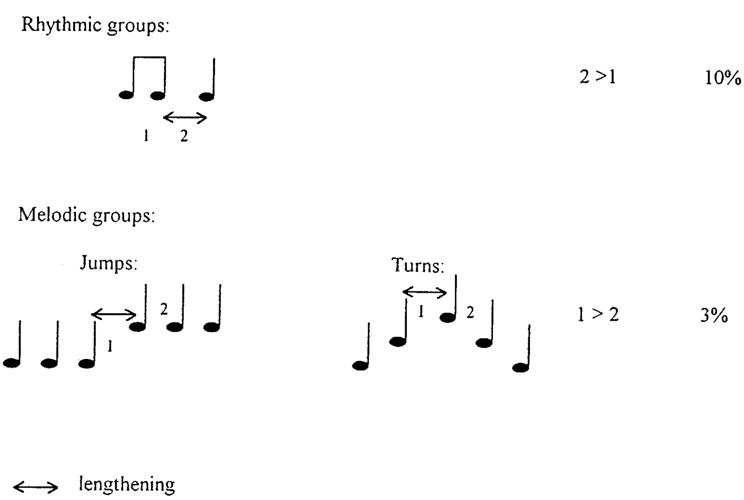


Fig. 2 Rhythmic and melodic groups and associated temporal variations produced in musical performance, according to Drake & Palmer (1993)

The corresponding higher processing level in this type of organization involves the insertion of these basic groups into a *hierarchical segmentation organization* based on the successive grouping of basic groups into increasingly larger groups until they incorporate whole musical phrases and finally the entire piece (levels B2, B3 and B4 in Fig. 1). Variations associated with these hierarchical levels probably originate from higher level processes developed by acculturation and musical training. They are probably used more by musicians than non-musicians (Clarke & Krumhansl, 1990) and they are probably absent from babies' representations. The most prominent studies have examined the phenomenon of phrase-final lengthening, whereby performers slow towards the end of phrases (Seashore, 1938; Todd, 1985; Kronman & Sundberg, 1987; Shaffer & Todd, 1987; Repp, 1990), the degree of slowing down being proportional to the depth of the hierarchical level (Gabrielsson, 1987; Shaffer & Todd, 1987). The central thesis is that the performer uses phrase-final lengthening as a device to reflect some underlying structure abstracted from the musical surface. Listeners are sensitive to this slowing (Sundberg & Verrillo, 1980). This approach has been formalized by Todd (1985), who combined Lerdahl and Jackendoff's theory of grouping (1983) with this principle of phrase-final lengthening to generate a duration structure corresponding to the rubato in a performance. We have chosen to test the predictions of this model (rather than later variants: see Todd, 1989, 1992, 1995) because it provides a simple way to simulate phrase-final lengthening using a parabolic function.

To summarize this second section, two fundamental organizational processes appear to be relatively universal and functional from an early age (beat extraction and segmentation into basic groups). Experience with a particular type of sequence such as that found in Western tonal music leads to an increasing ability to incorporate these basic principles into higher-level hierarchical organizations (hierarchical metric and hierarchical segmentation organizations). These processes result in a perceived structure similar to that proposed by Lerdahl and Jackendoff (1983). The second main aim of this paper was to investigate whether or not the results of the analysis of performance variations can be interpreted in the light of this processing model.

Principle

The first goal of this study was to show that, in the same normal musical context, both types of processes (high- and lower-level) operate in parallel – that is to say, both the musical expression hypothesis and the perceptual hypothesis apply. To demonstrate this, we analyzed both global variations (above the note level), reflecting the high-level processes of hierarchical segmentation organization and hierarchical metric organization, and local variations (at the note level), reflecting the lower-level processes of segmentation into basic groups. Global variations were analyzed in relation to Todd's 1985 model of phrase-final lengthening for variations reflecting hierarchical segmentation organization. Global variations reflecting hierarchical metric organization and local variations reflecting rhythmic and melodic grouping were analyzed in relation to the results of the study by Drake and Palmer (1993). Thus, we examined three lower-level processes producing variations (rhythmic groups, melodic jumps, and melodic turns). We recorded performances of Schumann's *Träumerei* by eight professional pianists. Each pianist played the piece three times with the same interpretation and was then asked to play it again in a mechanical fashion. Only the first eight bars of the piece, with the repeat, were analyzed (referred to as measures 1–16). In Part 1, we examined the performances and calculated intra- and inter-individual correlations. Average performances (an average "musical" performance and an average "mechanical" performance) were defined by computing arithmetic means (between the 24 musical performances and the 8 mechanical performances). In Part 2, we analyzed the timing variations of these average performances, both at global and local levels, in order to highlight the possible presence of variations caused by high- and lower-level processes. In Part 3, in order to assess the extent to which all temporal variations could be identified and to compare the relative contributions of high- and lower-level processes towards explaining the variability of musical and mechanical performances, we computed stepwise multiple regressions. Our prediction is that lower-level processes contribute towards explaining the variance more for mechanical performances than for musical performances, and that this is the opposite for high-level processes. Finally, we re-examined these findings in the light of the psychological model of temporal organization in order to assess whether a more process-oriented approach might present a harmonious model of performance timing variations.

Method

Participants. Eight professional pianists (two men and six women) took part in the experiment. They were piano teachers in music schools, accompanists, and concert pianists, with a mean age of 35 years (range: 25–46 years) and a mean of 22 years of practice (range: 16–38 years).

Materials. We analyzed the first eight measures of performances of Schumann's *Träumerei* (the 7th of the 13 short pieces that constitute *Scenes from Childhood*, Opus 15), with the repeat. The score used (from G. Henle Verlag) is reproduced in the annex. The piece was easy for the pianists concerned, thus reducing to a minimum any variations that could have been caused by technical difficulties. This piece was previously used by Repp (1992a, 1992b, 1995a), so our performances could be compared with a larger data set (either by expert pianists, see Repp, 1992b, or by student pianists, see Repp, 1995a).

Apparatus. Recordings were carried out on a Yamaha Disklavier II, monitored by a computer on which the sequencer Studiovision was the only software running at the time of the recordings.

Table 1 Intra-individual and inter-individual correlations for musical and mechanical performances

		Musical performances	Mechanical performances
Intra-individual	**mean**	**.89**	**.60**
	minimum	.67	.20
	maximum	.97	.90
Inter-individual	**mean**	**.73**	**.44**
	minimum	.46	.03
	maximum	.85	.83

Procedure. Pianists were asked to practice the piece before the recording session, hence it was a prepared performance. They performed the piece three times with the same interpretation. They noted on the score in traditional musical notation the interpretation they had wanted to convey so that it could be compared with what they actually did. At the end of the session, they were asked to play the piece again in a mechanical fashion, i.e., with the intention to play without expression, with the mechanical consistency of the musical score (Palmer, 1989).

Analyses. We analyzed temporal variations observed for the soprano voice, which corresponded in most cases to the principal voice (see Repp, 1992b). The position of a note is given by its metrical position, i.e., by the measure, the beat, and the half-beat of the note. Thus, the note in position 7-3-2 is the note of the second half-beat of the third beat of the 7th measure. Measures 9 to 16 correspond to the repeat. Using the "expressive timing" function of POCO (Honing, 1990), an environment for analyzing, modifying, and generating expression in music, text files were obtained for each performance, with the onset of each note and the ratio of the duration of the IOI in the score (fixed to an arbitrary value) in relation to the duration of the IOI as it was played. These text files were transported into spreadsheets for further analysis. Analyses of correlation, variance, and multiple regression were conducted using the Statistica software. We defined our relative temporal variation as being the normalized ratio of the duration of the IOI as it was played in relation to the duration of this IOI in the score. A value above 1 indicated that the note was played slower than the arithmetic mean of the notes in the excerpt, whereas a value below 1 indicated that the note was played faster.

Results

Part 1: The timing profiles

Intra- and inter-individual correlations

Musical performances. For each pianist, correlations were computed between the timing profiles of their six musical performances of the first eight measures (41 notes): each pianist produced three musical performances, each of which including a repeat of measures 1 to 8 (indicated in the score). Thus, for each pianist, we had six musical performances of the first eight measures. These correlations were all significant ($p < .001$), with a mean correlation of $r = 0.89$ (ranging from $r = 0.67$ to 0.97), similar in range to those found by Repp (1992b, 1995a). Thus, for each pianist, timing profiles of musical performances were consistent across the different performances and repeats. Some pianists claimed that they had chosen a slightly different interpretation for the repeat (S1: "more *piano*"; S2: "more free"; S3: "more remote"; S6: "more *piano*, flatter dynamics"). However, no evidence was found that these intentions changed anything for the temporal characteristics of the performance, as the respective correlations between the two repetitions of the same performances were in the same range as those for the pianists who claimed to have kept the same interpretation. In order to obtain an average musical profile for each pianist, i.e., a musical profile where random and non-characteristic variations were reduced, an arithmetic mean[2] was calculated over the different musical performances. In the following, the mean was computed on the three performances of measures 1 to 16 (the measures 9 to 16 being the repeat of measures 1 to 8) or on the six performances of the first eight measures (the repeats being considered as repeated performances), depending on whether or not we wanted to take into account the differences observed in the profiles of the repeats. Correlations were also computed between the average musical timing profiles of the eight pianists (41 notes). They were all significant ($p < .001$), with a mean correlation of $r = 0.73$ (ranging from $r = 0.46$ to 0.85). Here, no comparison was possible with Repp's data, but the size of the correlations remained relatively high, indicating that the different musical profiles had many characteristics in common.

Mechanical performances. Table 1 shows that similar intra-individual and inter-individual correlations calculated for the mechanical performances were lower than those observed for the musical performances. Concerning intra-individual correlations, only 6 correlations out of 8 were significant ($p < .01$). Concerning inter-individual correlations, only 20 correlations out of 28 were significant ($p < .03$). Thus, for each pianist, timing profiles of mechanical performances were less consistent across the different performances than timing profiles of musical performances, indicating less control from the pianists when trying to play mechanically. Also, timing profiles of mechanical performances by different pianists seem to have fewer characteristics in common than timing profiles of musical performances.

[2] Previous studies have used either an arithmetic or geometric mean. If values are relatively homogenous (as is the case here), the two methods provide similar results. We chose to use an arithmetic mean.

The mean timing profiles

To obtain a mean *musical* profile for our group of participants, an arithmetic mean was calculated over the eight individual mean musical profiles. The mean musical profile obtained (see Fig. 3) may not be an "ideal" performance profile; however, it incorporated what the different musical performances had in common and reduced the influence of random variations as well as the variations that could differ from one performer to another. The mean of the variations observed for measures 9 to 16 (the repeat) was almost the same as the mean of the variations observed for measures 1 to 8 ($r = 0.995$, $p < .001$). This indicated that the number of performances per participant was adequate, since the mean reduced the random variations that could have been produced. Moreover, this profile was very similar to those obtained by Repp (1992b, 1995a) when he calculated a geometric mean over 28 performances by 24 pianists or an arithmetic mean over 29 performances by 10 pianists, although he integrated in his profiles notes occurring in the other voices: positions 1-2-1, 4-2-1 to 4-4-2, and 7-2-2 to 8-4-2. This indicates that the number of participants was also adequate, as idiosyncrasies corresponding to the pianists were eliminated.

To obtain a mean *mechanical* profile for our group of participants, an arithmetic mean was calculated over the eight individual mean mechanical profiles. The mean mechanical profile obtained is also supposed to have incorporated what the different mechanical performances had in common and reduced the influence of random variations as well as the variations that could differ from one performer to another. The idea is that even when trying to play mechanically, pianists do produce systematic variations, either as residuals of variations intended when playing musically or as variations unintended even when playing musically. This mean mechanical timing profile is also presented in Fig. 3. As for the musical performances, the mean of the variations observed for measures 9 to 16 (the repeat) was almost the same as the mean of the variations observed for measures 1 to 8 ($r = 0.941$, $p < .001$). This indicated that the number of repetitions per participant was adequate (two instead of six for musical performances), since the mean reduced possible random variations.

As predicted, variations were globally reduced but did not totally disappear in mechanical performances: the standard deviation was 15.3% for the musical profile and 4.8% for the mechanical profile (averaged over the two repeats). Thus, a reduction by a factor 3 was observed. The variations observed in the mechanical profile had globally the same directionality as the variations observed in the musical profile: the correlation between both profiles was relatively high ($r = 0.890$, $p < .001$). This is in line with other findings (Palmer, 1989), where each of the patterns of chord asynchronies (particularly melody lead), rubato and overlaps was correlated with the same pattern in other types of performances (me-

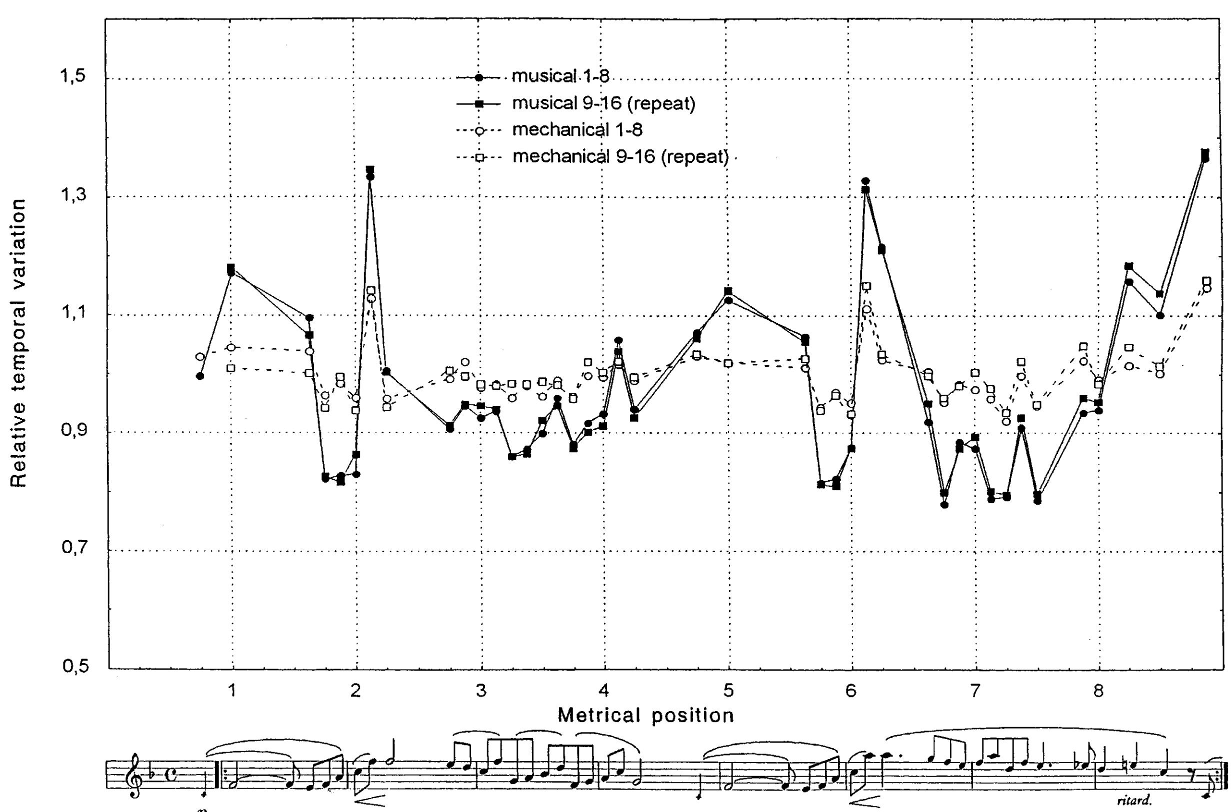


Fig. 3 Average musical and mechanical timing profiles for the two repeats

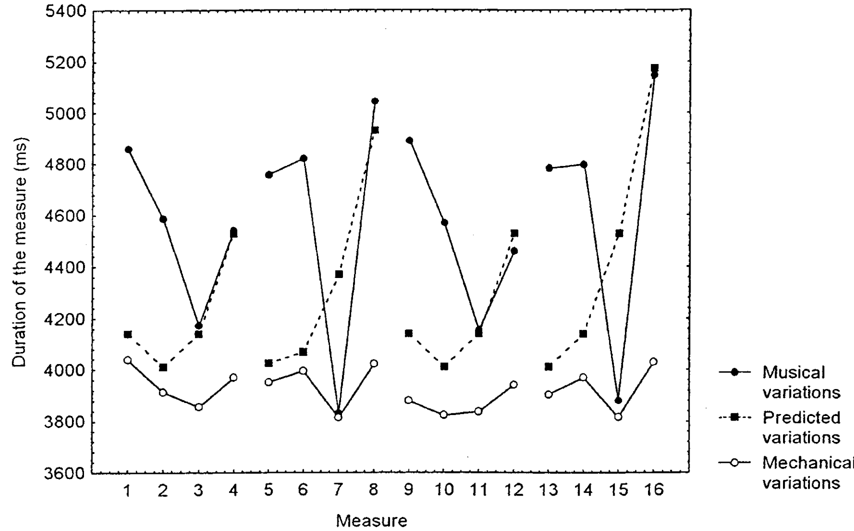


Fig. 4 Predicted and observed musical and mechanical variations in relation to Todd's 1985 model of phrase-final lengthening

chanical, musical, and exaggerated performances) and increased from unmusical to musical and to exaggerated performances. The author concluded in terms of a parameter theory of motor programs. In the following, we will examine in greater detail the differences between variations observed in musical and mechanical performances.

Part 2: Analysis of global and local performance variations

Analysis of global variations

Hierarchical segmentation organization. Global variations reflecting hierarchical segmentation organization were analyzed in relation to Todd's 1985 model predicting phrase-final lengthening proportional to the depth in the hierarchical structure.[3] Figure 4 presents the predicted and observed musical and mechanical variations at the measure level. Each point corresponds to the mean duration of the measure (that is, four beats that can contain between four and eight events). The predicted profile corresponded relatively well with the observed musical profile for the slowing towards the end of phrases. However, the model significantly under-estimated the accelerandos observed at the beginning of phrases.

To verify that phrase-final lengthening occurred in both musical and mechanical performances, we analyzed the mean durations of the last two measures in each phrase. An ANOVA on the mean duration by Instruction (musical and mechanical), Repeat (2), Phrase (2), and Position of the measure in the phrase (the second to last and the last) revealed only significant effects of phrase, $F(1, 30) = 7.6$, $p < .01$, and position, $F(1, 30) = 179.3$, $p < .0001$, significant interactions between instruction and phrase, $F(1, 30) = 4.4$, $p < .05$, instruction and position, $F(1, 30) = 78.8$, $p < .0001$, phrase and position, $F(1, 30) = 67.0$, $p < .0001$, and a significant three-way interaction between instruction, phrase, and position, $F(1, 30) = 42.6$, $p < .0001$. The significant effect of position indicates that the last measure in each phrase was on average played slower than the preceding one, showing that globally there was phrase-final lengthening. This was greater in some phrases than others and more pronounced in the musical than in the mechanical performances. We will return to these points in later analyses. To check that phrase-final lengthening also occurred in the mechanical performances, the protocol was restricted to *mechanical* performances. The ANOVA on the mean duration by Repeat (2), Phrase (2), and Position of the measure in the phrase (the second to last and the last) revealed only a significant effect of position, $F(1, 7) = 6.9$, $p < .04$, with the last measure in each phrase played significantly slower than the preceding one, showing that there was phrase-final lengthening in mechanical performances, also.

The phrase-final lengthenings can be taken as an indication of the perceived phrasings. To compare the lengthenings of each phrase, we analyzed the mean duration of the last measure in each phrase. An ANOVA on the mean measure duration by Instruction (2), Repeat (2), and Phrase (2) revealed only significant effects of instruction, $F(1, 30) = 7.3$, $p < .02$, and phrase, $F(1, 30) = 57.3$, $p < .0001$, and a significant interaction between instruction and phrase, $F(1, 30) = 35.6$, $p < .0001$. The significant effect of instruction is found because pianists adopted a faster tempo in the mechanical performances than in the musical ones. With a

[3] Several parameters had to be fixed for the application of Todd's 1985 model. The embedding depth of each structural ending E = (1, 3, 1, 4) was set by the analysis of cadential groups; the baseline tempo A = 4011 ms was set by averaging the durations of measures 3, 7, 11, and 15 in the musical profile (which gives an estimation of the minimal global tempo); and the rubato amplitude m = 1033 ms was set by averaging the values found for m to make the duration of the last measure of each phrase as predicted by the model coincide with the observed duration in the musical profile.

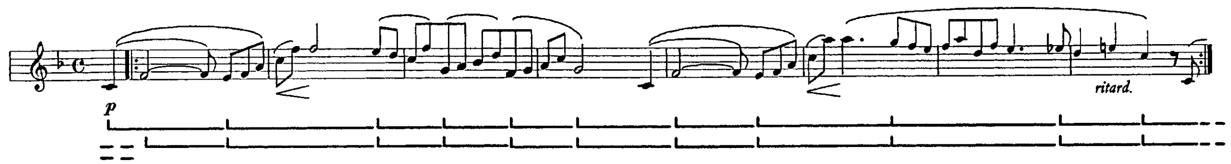


Fig. 5 Subdivisions of the musical excerpt into melodic gestures according to Repp (1992b)

protocol restricted to *musical* performances, the ANOVA on the mean measure duration by Repeat (2) and Phrase (2) revealed only a significant effect of phrase, $F(1, 23) = 228.8$, $p < .0001$, and a significant interaction between the two, $F(1, 23) = 9.66$, $p < .005$. Thus, phrase-final lengthening was greater in the second phrase of each repeat than in the first phrase, and this effect was stronger in the second repeat. Thus, the results provided evidence for three hierarchical levels: phrase-final lengthening highlighted a segmentation between phrases; a greater phrase-final lengthening in the second phrase of each repeat than in the first phrase highlighted a segmentation between repeats; and a stronger effect in the second repeat highlighted a segmentation at the end of the second repeat. With a protocol restricted to *mechanical* performances, the ANOVA on the mean measure duration by Repeat (2) and Phrase (2) revealed no significant effects or interactions. Contrary to musical performances, phrase-final lengthening was not greater in the second phrase of each repeat than in the first phrase, and this effect was not stronger in the second repeat. Thus, the results provided evidence for only one hierarchical level (a segmentation between phrases).

The measure is the basic unit in Todd's (1985) model. Another way of visualizing global variations is to take melodic gestures, as defined by Repp (1992b), as the basic unit. Figure 5 shows the subdivision into melodic gestures used.[4] Each phrase was composed of five or six melodic gestures. Figure 6 shows that, within each phrase, the melodic gestures also followed a pattern of accelerando followed by a ritardando, similar in nature to that observed at the measure level.

As previously, phrase-final lengthening was investigated by analyzing the relative temporal variation of the last two melodic gestures in each phrase. An ANOVA on the relative temporal variation by Instruction (2), Repeat (2), Phrase (2), and Position of the melodic gesture in the phrase (the second to last and the last) revealed only significant effects of phrase, $F(1, 30) = 76.0$, $p < .0001$, and position, $F(1, 30) = 34.9$, $p < .0001$, significant interactions between instruction and phrase, $F(1, 30) = 37.2$, $p < .0001$, repeat and phrase, $F(1, 30) = 5.5$, $p < .03$, instruction and position, $F(1, 30) = 16.7$, $p < .001$, and phrase and position, $F(1, 30) = 14.0$, $p < .001$, and a significant three-way interaction between instruction, phrase, and position, $F(1, 30) = 14.0$, $p < .001$. As previously, the significant effect of position indicates that the last melodic gesture in each phrase was on average played slower than the preceding one, showing that globally there was phrase-final lengthening. With a protocol restricted to *mechanical* performances, the ANOVA on the relative temporal variation by Repeat (2), Phrase (2), and Position of the melodic gesture in the phrase (2) revealed no significant effects or interactions. However, by analyzing the relative temporal variation of the third to last and last melodic gesture in each phrase, a significant effect of position of the melodic gesture was found, $F(1, 7) = 7.8$, $p < .03$, with the last melodic gesture in each phrase played significantly slower than the third to last one, showing as previously that there was phrase-final lengthening. (It was the only significant effect or interaction.)

To compare the lengthenings of each phrase, we analyzed the relative temporal variation of the last melodic gesture in each phrase. Above, we used absolute durations as implemented in Todd's (1985) model. However, as melodic gestures did not all have the same duration in the score, relative values had to be used in this new analysis. As previously, an ANOVA on the relative temporal variation by Instruction (2), Repeat (2), and Phrase (2) revealed only significant effects of instruction, $F(1, 30) = 9.3$, $p < .005$, and phrase, $F(1, 30) = 68.1$, $p < .0001$, and a significant interaction between instruction and phrase, $F(1, 30) = 33.8$, $p < .001$. Here, as relative values are used, the significant effect of instruction is not caused by a faster tempo in the mechanical performances but by differences between the profiles of musical and mechanical performances. With a protocol restricted to *musical* performances, the ANOVA on the relative temporal variation by Repeat (2) and Phrase (2) revealed only a significant effect of phrase, $F(1, 23) = 261.4$, $p < .0001$, and a significant interac-

[4] This subdivision takes into account the switch of the principal voice to the tenor voice at position 7 (15)-4-2 and to the bass voice at position 4 (12)-2-1 and 8 (16)-3-1: melodic gestures are defined from position 7 (15)-4-2 to position 8 (16)-3-1, from position 4 (12)-2-1 to position 5 (13)-1-1, and from position 8 (16)-3-1 to position 9 (17)-1-1. Indeed, we have chosen to study only timing variations observed in the soprano voice, which contains the main melodic events. However, in terms of principal voices, there is a switch to the tenor and to the bass voices at the positions mentioned above and a switch to the alto voice at position 7 (15)-2-2 (see Repp, 1992b). We have chosen to take these switches into account, as we think that principal voices carry phrasing effects (from the performer's point of view) and are responsible for phrasing perception (from the listener's point of view). Because of the overlapping of melodic gestures in the alto and in the tenor voices at position 7 (15)-4-2, and because of the paralleling of the melodic gesture in the tenor voice with the soprano and bass voices, the melodic gesture in the alto voice was not taken into account in the subdivision. This is the same for the two notes beginning each phrase in the soprano voice when overlapping with the bass voice (thus explaining that phrase 2 is segmented differently from phrase 1, and that phrase 1 is segmented differently in the repeat).

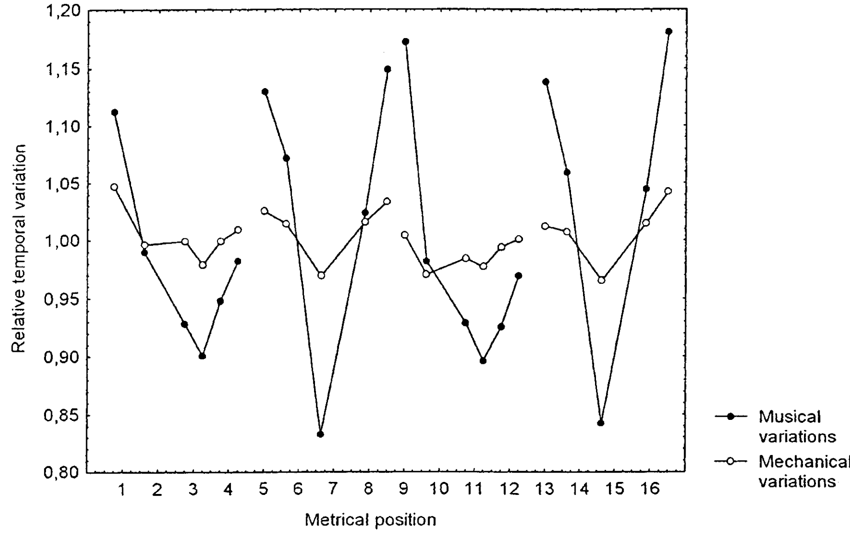


Fig. 6 Observed musical and mechanical variations in relation to melodic gestures

tion between the two, $F(1, 23) = 4.8$, $p < .04$. This analysis confirmed the preceding analysis in showing evidence for three hierarchical levels of segmentation. However, this new analysis showed more clearly than in the previous analysis the predicted accelerando/ritardando profile: the number of basic units appears optimal leading to smoother curves, in particular without the irregularities of measures 6 and 12. We suggest that this is due to the fact that melodic gestures are a more psychologically relevant basic unit than measures. With a protocol restricted to *mechanical* performances, the ANOVA on the relative temporal variation by Repeat (2) and Phrase (2) revealed no significant effects or interactions. This analysis confirmed the preceding analysis in providing evidence for only one hierarchical level (a segmentation between phrases).

Hierarchical metric organization. In order to test whether particular metrical positions undergo any systematic timing variations, each note of the excerpt situated on a beat was labelled according to its metrical position (1 to 4). From here on, data were averaged over the repeat in order to avoid biasing analyses towards rejecting the null hypothesis: the data between the two repetitions were so similar that it was as if the data were repeated twice. Also, we did not take into account the last notes from the position 7-4-2, as a large ritardando (indicated on the score) is observed because of the end of the phrase. Figure 7 shows the mean relative temporal variation for musical and mechanical performances at the beat level for each metrical position and the values obtained in relation to a piece in 4/4 observed by Drake and Palmer (1993). The shape of musical and mechanical profiles is similar to that found previously by Drake and Palmer (in musical performances), even though the style of music is completely different. Larger effects were found for the musical than for the mechanical performances. An ANOVA on the relative temporal variations by Instruction (2) and Metrical position (4) confirmed significant effects of instruction, $F(1, 30) = 30.4$,

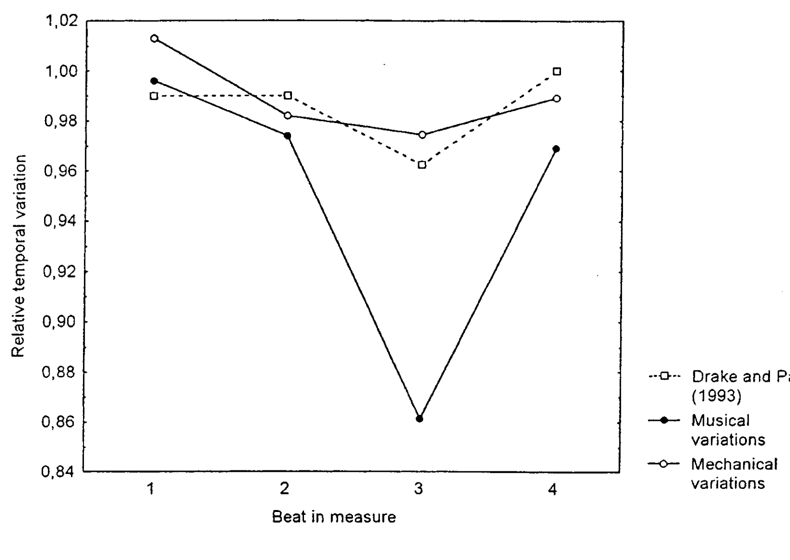


Fig. 7 Musical and mechanical variations in relation to the metric structure observed here and by Drake & Palmer (1993)

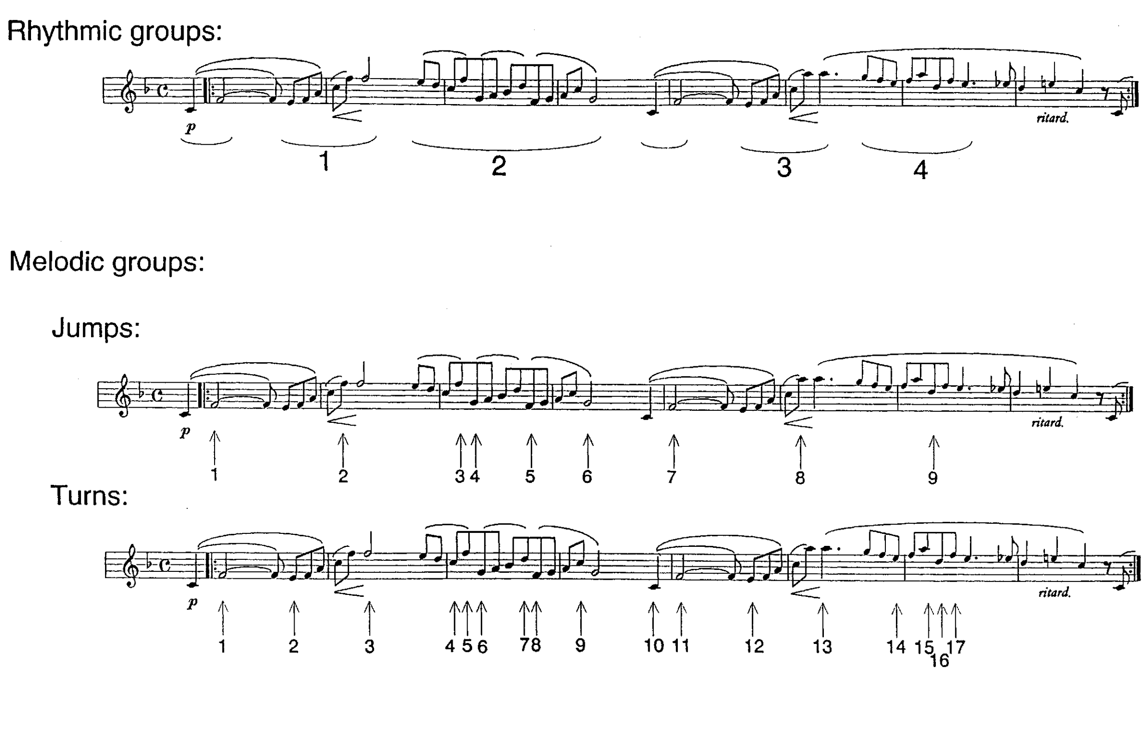
$p < .0001$, and position, $F(3, 90) = 38.4$, $p < .0001$, and a significant interaction between the two, $F(3, 90) = 17.2$, $p < .0001$. With a protocol restricted to *musical* performances, an ANOVA on the relative temporal variations by Metrical position (4) revealed a significant effect of position, $F(3, 69) = 84.3$, $p < .0001$. With a protocol restricted to *mechanical* performances, the same ANOVA also revealed a significant effect of position, $F(3, 21) = 10.2$, $p < .001$. Thus, pianists' musical and mechanical performances seemed to be influenced by the metric structure, though to a lesser extent for mechanical than for musical performances.

Analysis of local variations

We analyzed local variations observed in relation to three hypothesized organizational principles: rhythmic and melodic groups corresponding to jumps and turns.

Rhythmic groups. Four rhythmic groups containing at least three notes (two intervals) were analyzed (see Fig. 8). Figure 9 shows the relative temporal variations for musical and mechanical performances observed for the notes of each rhythmic group. In agreement with predictions, the last interval (note at position −1) was played longer than the preceding one (note at position −2). For musical performances, the deviations were: group 1 = 56.6%, group 2 = 14.0%, group 3 = 50.4%, group 4 = 15.8%. For mechanical performances, they were: group 1 = 19.6%, group 2 = 2.1%, group 3 = 20.2%, group 4 = 8.7%. An ANOVA on these relative temporal variations by Instruction (2), Group (4), and Position in the groups (the second to last and last interval) revealed no significant effect of instruction, significant effects of group, $F(3, 90) = 21.4$, $p < .0001$, and position in the group, $F(1, 30) = 121.6$, $p < .0001$, and significant interactions between instruction and group, $F(3, 90) = 6.5$, $p < .001$, instruction and position, $F(1, 30) = 21.2$, $p < .0001$, and group and position, $F(3, 90) = 36.4$, $p < .001$, and a significant three-way interaction, $F(3, 90) = 6.6$, $p < .001$. With a protocol restricted to *musical* performances, an ANOVA on the relative temporal variation by Group (4) and Position (2) revealed significant effects of group, $F(3, 69) = 41.6$, $p < .0001$, and position in the group, $F(1, 23) = 239.8$, $p < .0001$, with the last interval longer than the preceding one, and a significant interaction between the two, $F(3, 69) = 62.5$, $p < .0001$, indicating that the lengthening of the last interval in relation to the preceding one was greater for groups 1 and 3 than for groups 2 and 4. The size of the effect for groups 2 and 4 corresponded to that predicted in the literature (10%: Drake & Palmer, 1993). However, it was much higher for groups 1 and 3, which may be explained by the presence of arpeggios in the left hand on the last notes of groups 1 and 3. Groups 2 and 4, which did not contain an arpeggio, were more representative of the rhythmic grouping process. With a protocol restricted to *mechanical* performances, the same results were obtained: significant effects of group, $F(3, 21) = 6.6$, $p < .01$, and position in the group, $F(1, 7) = 14.7$, $p < .01$, with the last interval longer than the preceding one, and a significant interaction between the two, $F(3, 21) = 9.6$, $p < .001$, indicating that the lengthening of the last interval in relation to the preceding one was greater for groups 1 and 3 than for groups 2 and 4. The size of the

Fig. 8 Rhythmic groups and melodic jumps and turns of the excerpt



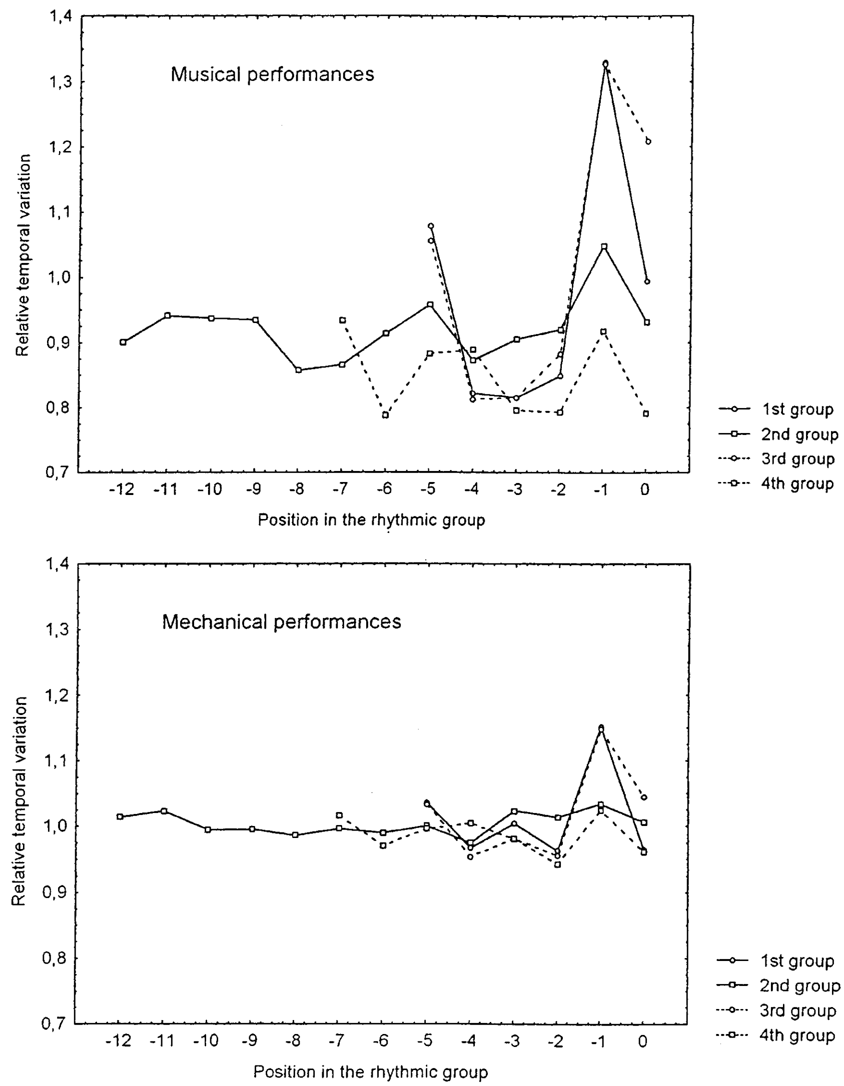


Fig. 9 Musical and mechanical variations observed in relation to rhythmic groups. On the x-axis, *0* codes the last note in each rhythmic group (that is, the long note), *−1* codes the last short note, and *−2* the second-to-last short note

effect for groups 2 and 4 was lower than that predicted in the literature (10%: Drake & Palmer, 1993), but still present. It was higher for groups 1 and 3, which can also be explained by the presence of arpeggios in the left hand on the last notes of groups 1 and 3: pianists had to play three times in a musical way (they prepared the recordings and worked the piece beforehand in this way) and then had to play once in a mechanical way, which did not enable us to fully test the degree of voluntary control. We suggest that with more practice of mechanical playing, they would be able to reduce the deviations of groups 1 and 3 to the range of deviations of groups 2 and 4. The three-way interaction observed indicates that the interaction between group and position depends on instruction. Indeed, this interaction is less pronounced for mechanical performances than for musical performance. We suggest that with more practice of mechanical playing, no interaction between group and position would be observed. The effect of position would then be a pure effect of temporal variation due to the process of rhythmic grouping.

Melodic groups. Figure 8 also presents all instances of jumps (changes in pitch of more than 5 semitones) and turns (changes in pitch direction). Here, the analyses were only conducted on *musical* performances, as even in this condition, variations found were very unsystematic.

For the *jumps*, contrary to predictions (the IOI preceding the jump or turn is, on average, played 3% longer than the following one), the interval before a jump was played on average shorter than the interval after the jump (see Fig. 10). An ANOVA on the relative temporal variations by Jump (9) and Position (2) revealed a main effect of position, $F(1, 23) = 71.8$, $p < .0001$, with the first interval shorter than the second, a main effect of jump, $F(8, 184) = 43.6$, $p < .001$, and a significant interaction between the two, $F(8, 184) = 107.6$, $p < .001$. Temporal variations related to melodic groups are weak compared with others. (In case of conflict, they are dominated by variations related to rhythmic groups and to a lesser degree by variations related to the metric structure: see Drake & Palmer, 1993). Consequently, we

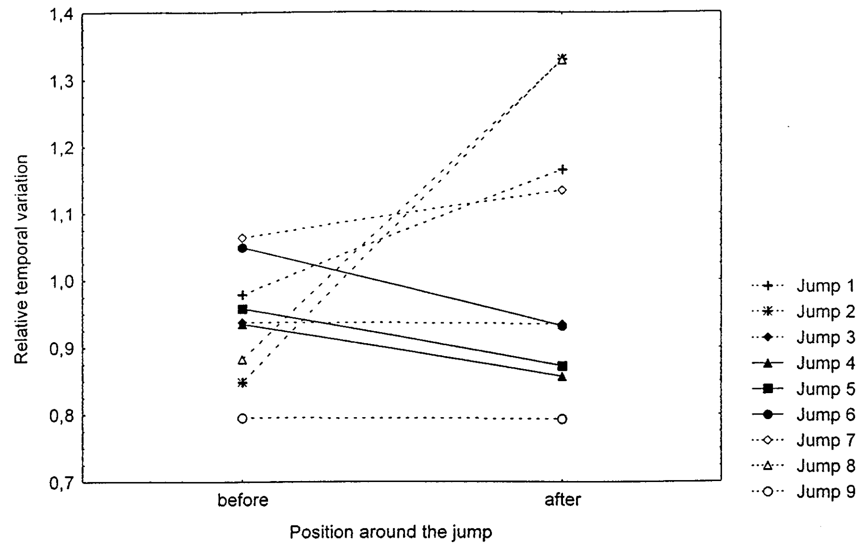


Fig. 10 Musical variations observed in relation to melodic jumps

re-ran the analysis removing the instances of jumps which were in conflict with other structural features (those related to rhythmic group: jumps 2 and 8; to the beginning and end of phrases: jumps 1 and 7; and to melodic gestures, see later: jumps 3 and 9). Such an approach using a restricted protocol limits the theoretical conclusions that can be drawn concerning jumps, but the aim was to obtain a value for the lengthening when it did occur (see Part 2). For this restricted protocol (jumps 4, 5, and 6, marked with continuous lines in Fig. 10), the predicted accelerando was observed: an ANOVA on the relative temporal variations by Jump (3) and Position around the jump (2) revealed a significant effect of position, $F(1, 23) = 144.5$, $p < .0001$, with the IOI preceding the jump being on average lengthened by 10.5% with regard to the following one. There was also a significant effect of the jump, $F(2, 46) = 26.3$, $p < .0001$, but the interaction between these two factors was not significant.

For *turns*, on average, no significant difference was observed between the duration of the interval before and after the turn (Fig. 11). In the same way as for jumps, we re-ran the analysis on a restricted protocol, removing instances of turns that were in conflict with other structural features (those related to rhythmic group: turns 3, 9, and 17; to the beginning and end of phrases: turns 1, 10, and 11; and to melodic gestures, see later: turns 4, 5, 7, 14, and 16). For this restricted protocol (turns 2, 6, 8, 12, 13, 15, marked with continuous lines in Fig. 11), the predicted accelerando was observed: an ANOVA on the relative temporal variations by Turn (6) and Position around the turn (2) revealed a significant effect of position, $F(1, 23) = 52.7$, $p < .0001$, with the IOI preceding the turn being on average lengthened by 9.2% with regard to the following one. There was also a significant effect of turn, $F(5, 115) = 98.1$, $p < .001$, but the interaction between the two factors was not significant.

Thus, evidence in support of melodic groups was weak: the associated variations were not found systematically. We suggest that some of them either tended to be masked by the conflicting presence of other types of variations or were not produced at all. The observed

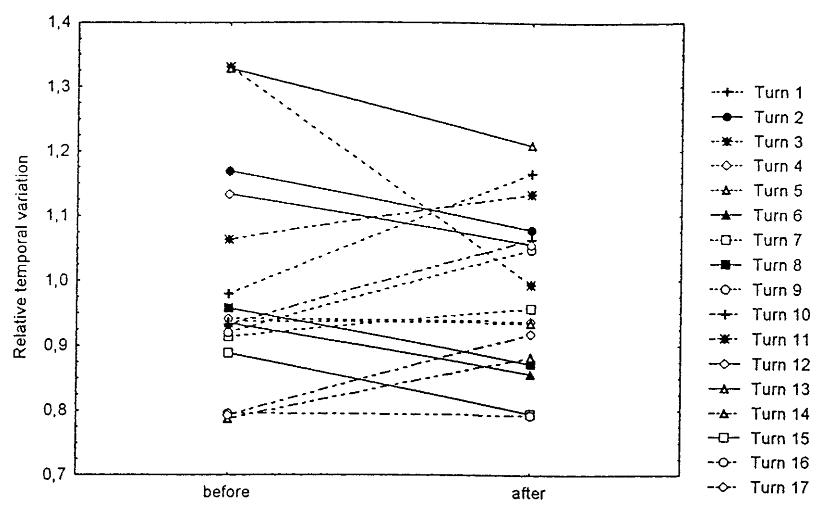


Fig. 11 Musical variations observed in relation to melodic turns

value from the restricted protocol was higher (10%) than observed in the literature (3%) where no selection was undertaken.

Discussion of Part 2: The analysis of performance variations

The analyses of the musical timing profile obtained by averaging over three performances by eight pianists have provided evidence that pianists organize the musical sequence both at a global level and at a local level. First, analyses of global variations indicated that pianists segmented the piece at the end of the second repeat, between repeats, and between phrases, providing evidence for three hierarchical levels in the hierarchical segmentation organization. For these analyses, the use of the melodic gesture as the basic unit was more appropriate than the use of the measure, as in Todd's (1985) model. Variations in relation to the metric structure were also observed. Second, analyses of local variations indicated that pianists segmented the sequence primarily in relation to rhythmic groups. Evidence for the melodic groups was weak.

The analyses of the mechanical timing profile obtained by averaging over one performance by eight pianists confirmed the high-level processing involved in hierarchical segmentation organization and hierarchical metric organization: an organization on three hierarchical segmentation levels in the musical performances became an organization on one hierarchical segmentation level in the mechanical performances, and variations reflecting hierarchical metric organization were considerably reduced. It also confirmed the low-level processing involved in rhythmic grouping: associated variations were reduced, but more for groups 1 and 3 than for groups 2 and 4. Additional lengthening is added when playing musically, but there is an initial lengthening due to perceptual compensation: the last interval of a rhythmic group is perceived too short; thus, it is played longer to sound regular.

Consequently, since hierarchical segmentation organization and hierarchical metric organization result from high-level processes and rhythmic groups from lower-level processes, this analysis showed that in the same normal musical context, both high- and lower-level processing operate in parallel. Previously, studies of very simple musical sequences have focused on low-level processing related to acoustic surface features, whereas studies of more complex musical sequences have focused on high-level processing related to the musical structure. Our analyses have demonstrated that the complexity of the sequence does not eliminate lower-level processing.

Part 3: Relative contributions of each source of variation

The relative contributions of the different types of variations described above in explaining the variance of the average musical and mechanical profiles were obtained, using forward stepwise regression analyses, by taking the average timing profile as the dependent variable and the different types of variation as the independent variables. (The cases are the 41 notes.) For musical performances, this analysis was first carried out using variation values predicted in the literature in order to establish the percentage of variance explained by these values. As these values did not entirely correspond with variations observed here, the analysis was repeated using the values obtained in Part 2 for musical performances. For mechanical performances, the use of variation values predicted in the literature made no sense, as these values corresponded to musical contexts. Thus, the regression analysis was only carried out using values obtained in Part 2 for mechanical performances.

Variation values predicted in the literature

The forward stepwise regression analysis conducted on the *musical* timing profile with variation values predicted in the literature used the following parameters: F to enter = 4; dependent variable = average musical timing profile; 4 independent variables = deviations from minimum tempo of global variations corresponding to hierarchical segmentation (in accordance with Todd's 1985 model, see Fig. 4) and hierarchical metric organization (in accordance with the values presented in Drake & Palmer, 1993; see Fig. 7), and of local variations corresponding to rhythmic and melodic groups (in accordance with the values presented in Drake & Palmer, 1993, see Fig. 2). The following regression equation was found:

$$\mathrm{Y} = 514 + 2.62 \times (\text{rhythmic groups}) + 0.99 \times (\text{hierarchical segmentation})\ \text{ms} \quad (1)$$

(variables are in order of entry in the regression), where 514 ms represents the minimum tempo. The coefficient significantly above 1 associated with rhythmic groups resulted from observed values being much higher than predicted values (mainly because of the arpeggios in the left hand), and the coefficient very close to 1 associated with hierarchical segmentation indicated that the values of A and m used in the implementation of the model were correctly estimated. This analysis indicated that the effects of variations predicted in the literature that were related to hierarchical segmentation and rhythmic groups made a significant contribution to explaining the variance of the musical timing profile (variance explained by hierarchical segmentation = 14.3%; by rhythmic groups = 17.6%), explaining a total of 31.9% of the variance.

Variation values taken from the analysis (Part 2)

The forward stepwise regression analysis conducted on the *musical* timing profile with variation values taken from the analysis used the following parameters: F to enter = 4; dependent variable = average musical timing profile; 4 independent variables = deviations from minimum tempo of global variations corresponding to hierarchical segmentation (musical observed variations, see Fig. 4) and hierarchical metric organization (musical observed variations, see Fig. 7), and of local variations corresponding to rhythmic and melodic groups (musical observed variations, see Figs. 9–11). The following regression equation was found:

$$\begin{aligned} \mathrm{Y} = 436 &+ 1.20 \times (\text{hierarchical segmentation}) \\ &+ 0.77 \times (\text{rhythmic groups})\ \text{ms} \end{aligned} \tag{2}$$

As with variations predicted in the literature, this analysis indicated that the effects of variations from Part 2 of this study that were related to hierarchical segmentation and rhythmic groups made a significant contribution to explaining the variance of the musical timing profile (variance explained: hierarchical segmentation = 33.1%; rhythmic groups = 24.9%), explaining a total of 58.0% of the variance. Here, of course, the total percentage of variance explained was greater.

The forward stepwise regression analysis conducted on the *mechanical* timing profile with variation values taken from the analysis used the following parameters: F to enter = 4; dependent variable = average mechanical timing profile; 4 independent variables = deviations from minimum tempo of global variations corresponding to hierarchical segmentation (mechanical observed variations, see Fig. 4) and hierarchical metric organization (mechanical observed variations, see Fig. 7), and of local variations corresponding to rhythmic groups (mechanical observed variations, see Fig. 9). The following regression equation was found:

$$\begin{aligned} \mathrm{Y} = 469 &+ 0.61 \times (\text{rhythmic groups}) \\ &+ 1.26 \times (\text{hierarchical segmentation})\ \text{ms} \end{aligned} \tag{3}$$

This analysis indicated that the effects of variations from Part 2 of this study that were related to hierarchical segmentation and rhythmic groups made a significant contribution to explaining the variance of the mechanical timing profile (variance explained: hierarchical segmentation = 22.5%; rhythmic groups = 38.1%), explaining a total of 60.6% of the variance.

Discussion of Part 3: Relative contributions of each source of variation

In this analysis we compared the relative contributions of different sources of variation: hierarchical segmentation organization, hierarchical metric organization, and segmentation into basic rhythmic and melodic groups. All regression analyses (according to the literature for the musical profile and following the results of the analysis for the musical and the mechanical profiles) present a clear pattern: in all cases, a significant part of variance was explained by variations related to hierarchical segmentation organization and by variations related to rhythmic groups. Hierarchical metric organization and melodic groups did not contribute significantly. The apparent contradiction between the variations observed in relation to the hierarchical metric organization (Part 2) and its lack of significant contribution in explaining the variance (Part 3) may be explained by a confounding of the metric structure with other structural features in the music. Further studies which would try to separate these effects are necessary to elucidate this point. What is of particular interest is the comparison of percentages of variance explained by hierarchical segmentation organization and by rhythmic grouping for the musical and the mechanical profiles (with variations values from Part 2). For the musical profile, hierarchical segmentation organization explained a larger part of the variance (33.1%) than rhythmic grouping (24.9%). It was the opposite for the mechanical profile: hierarchical segmentation organization explained a smaller part of the variance (22.5%) than rhythmic grouping (38.1%). This confirmed once again the high-level processing involved in hierarchical segmentation organization and the lower-level processing involved in rhythmic grouping. Variations related to hierarchical segmentation organization are under the musicians' voluntary control and thus can disappear if the musicians so desire. We suggest that with more practice of mechanical playing, pianists would be able to completely eliminate phrase-final lengthening. Variations related to rhythmic grouping are unavoidable and correspond to perceptual compensation. Thus, they can be reduced in mechanical performances if additional variations were produced in musical performances, but we suggest that they cannot disappear completely, even with more practice of mechanical playing. Indeed, the percentage of variance explained by rhythmic grouping increased from musical to mechanical performances, whereas the percentage of variance explained by hierarchical segmentation organization decreased from musical to mechanical performances. For the musical timing profile, the size of the contributions was similar for both organizational principles, either by using variations predicted in the literature or by using variations found in the analysis. This underlines the importance of including lower-level sources of variation when analyzing music performance. As predicted, the effects were larger when the values obtained in the analysis were used. In this case, the proposed organizational principles can account for 58% of the musical profile variance and 61% of the mechanical profile variance. Of the variance, 40% still remains to be identified, probably by the identification of other potential sources of variation.

General discussion

Until recently, the general approach to the study of music performance has been to emphasize how systematic variations are related to the musical structure. In

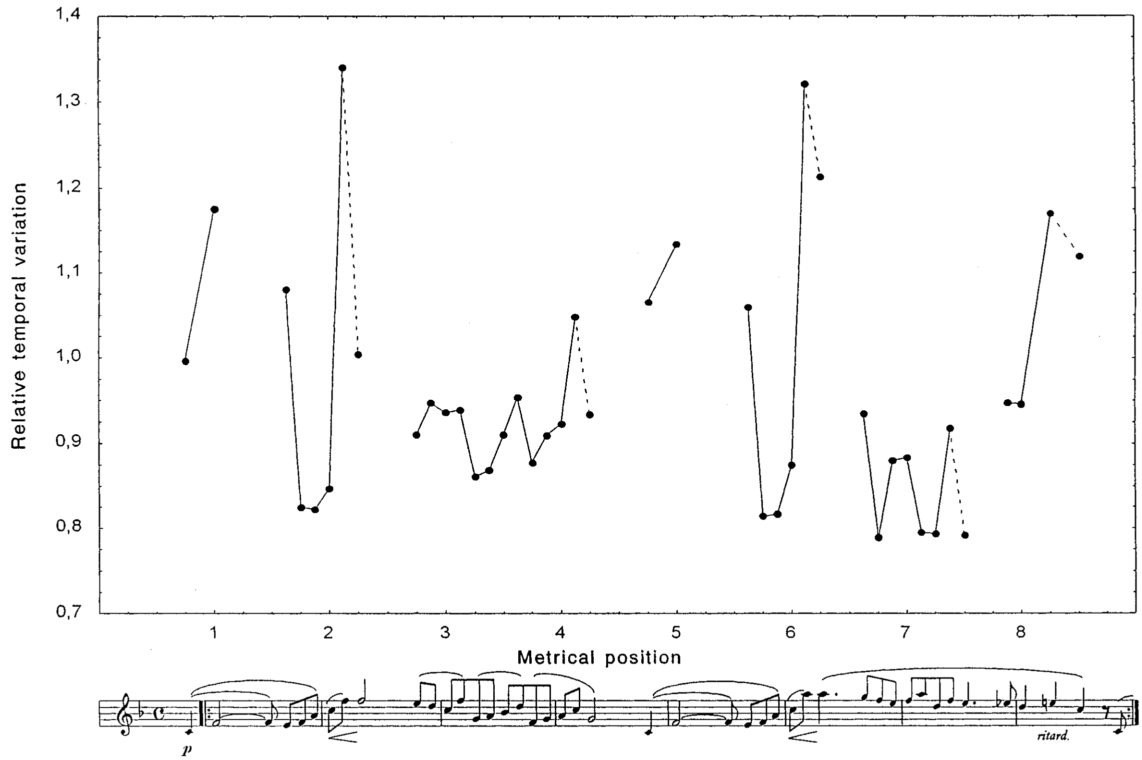


Fig. 12 Segmentation into 7 rhythmic groups

this paper we propose a change in emphasis, whereby performance variations are studied in relation to the psychological processes allowing the musician to perceive the musical structure. We have described two types of temporal organization – regularity extraction and segmentation into groups – , each of which operates at a lower processing level (beat extraction and segmentation into basic groups) and at a high processing level (hierarchical metric organization and hierarchical segmentation organization), involving the creation of increasingly larger units. Performance variations have been observed in relation to some of these organizational principles. At the lower processing level, analysis of local variations highlighted systematic variations in relation to rhythmic but not melodic groups. At the high processing level, analysis of global variations highlighted systematic variations in relation to hierarchical segmentation organization (for this analysis, the melodic gesture as the basic unit was more appropriate than the measure, as in Todd's (1985) model) and hierarchical metric organizations.[5] Moreover, the percentage of variation explained by rhythmic groups was similar in magnitude to that explained by hierarchical segmentation organization. As predicted, the comparison of the musical and mechanical performances showed that variations resulting from high-level processes decrease from musical to mechanical performances more than variations resulting from lower-level processes. These results support our position that both high- and lower-level processing underlie performance variations, and that the musical expression hypothesis and the perceptual hypothesis must be considered complementary.

We will now demonstrate how the results obtained here can be incorporated into a psychological model of music performance. This analysis is restricted to the musical performances. We will examine patterns of variations at four superimposed hierarchical levels of analysis.

Let us first examine what happens at the surface level. Figure 12 shows the average musical timing profile at the note level. How did pianists segment the sequence into basic groups? The analysis of local variations provided a first criterion: temporally long events terminated perceived groups, leading to a segmentation into seven basic units (rhythmic groups), containing 2 to 13 notes. This segmentation is demonstrated by the lengthening towards the end of each of these basic units (if one excludes the last note, which, being already long, does not need to be lengthened further). However, 13 notes is probably too long for a basic group (particularly at the slow tempo of this piece). It is likely that pianists segmented the sequence further, using additional criteria. Indeed, local lengthenings suggest that they further segmented below the rhythmic group level in the instances when the rhythmic group exceeded five notes: Fig. 12 shows several local lengthenings within the long rhythmic groups in measures 2–4 and 6–7. The local

[5] However, we do not claim that global variations always reflect high-level processes and that local variations always reflect lower-level processes. (It is obvious that some very local variations may be introduced very consciously by the performer, for instance, to emphasize a particular note.)

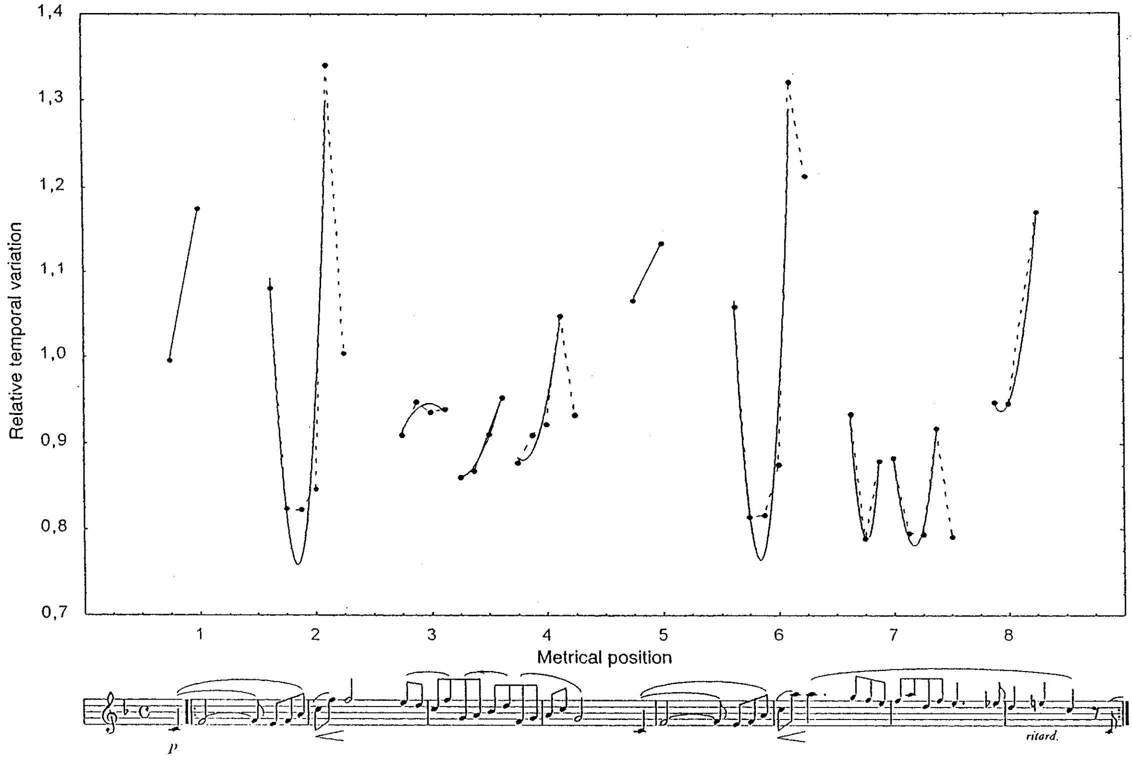
lengthenings within the rhythmic group in measures 2–4 corresponded to the segmentations the composer indicated with phrase marks. Does it mean that pianists followed his wishes? Another explication for these lengthenings, although nonexclusive, would be that pianists segmented according to jumps in positions 3-2-1 and 3-4-1. As a matter of fact, these were among the jumps for which the temporal variation predicted in the literature was observed (jumps 4 and 5). This would suggest that an additional criterion to segment the sequence would be pitch jumps or turns, despite the fact that associated performance variations are not systematic. The rhythmic group in measures 2–4 contains four instances of jumps and three instances of turns (excluding the turns which coincided with jumps). In a strict application of psychoacoustic rules, i.e., if each instance of jump and turn led to a segmentation, this would create eight melodic groups for 13 notes. Of course, this was not the case. Pitch jumps and turns did not systematically lead to a segmentation (and to the associated performance variation), but they may have constituted an opportunity for it. The second example of local lengthening within a long rhythmic group (measures 6–7) would suggest that, contrary to the very long phrase indicated by the composer, pianists segmented at the bar line, creating respectively 3- and 5-note basic groups. The criterion used for the segmentation was less clear here: it may have been the metrical position or the harmony created by the chord in the left hand. (We consider the segmentation into basic groups to reflect lower-level processes, but we do not exclude the possibility of interactions with higher levels.) The context and, more specifically, the influence of other voices have to be taken into account.

These additional segmentations of the long rhythmic groups into shorter basic units resulted in ten basic groups, containing 2–5 notes. In sum, the rules of surface segmentation seem to be much less general than in simple psychoacoustic experiments and strongly context dependent. In any case, there appears to be a very consistent pattern of accelerando/ritardando, or simply ritardando, which follow notes within the basic perceptual units, as the interpolations of 2nd-degree polynomials in Fig. 13 suggest.

These segmentations can be compared with those proposed by Repp (1992b) according to his musical intuition, forming what he calls "melodic gestures." There is a very close correspondence with the detailed segmentation described above, except that Repp did not further subdivide the long rhythmic groups in measures 6–7 (MG3b in Repp, 1992b), as suggested by the composer but not by the analysis above. Future work will need to focus on how factors (pitch jumps and turns, influence of other voices, etc.) combine to form the segmentation at this lowest level.

We can now move up the hierarchical segmentation organization to see how these basic groups create larger units. As we have already seen, pianists segmented the sequence at the end of the second repeat, between repeats, and between phrases. Thus, the basic units

Fig. 13 Segmentation into 10 basic groups, with associated polynomial fits



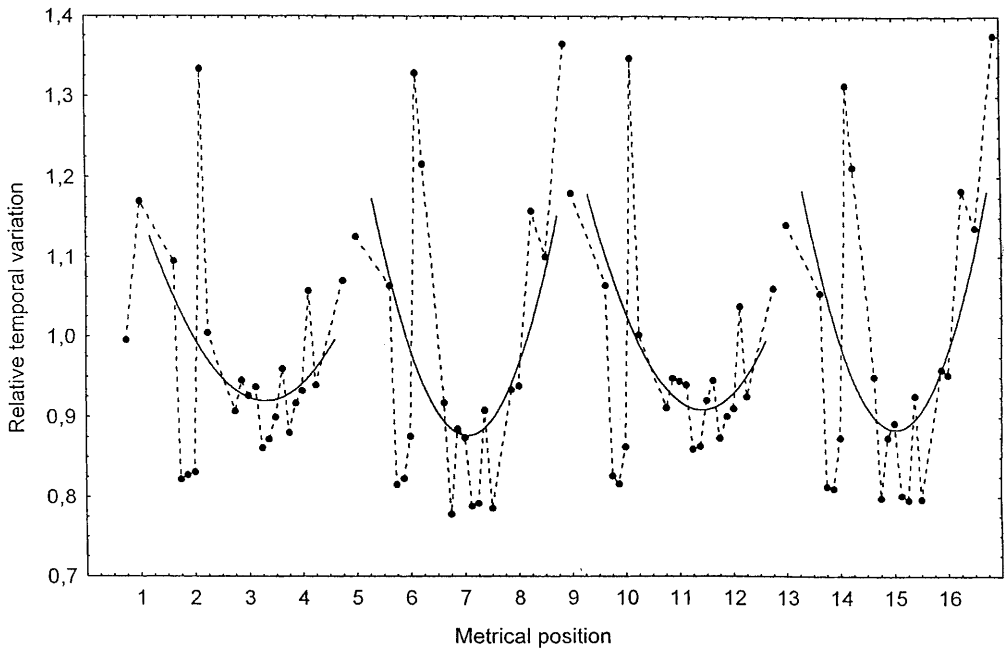


Fig. 14 Polynomials fitted to 4 phrases

grouped to form phrases, phrases grouped to form repeats, and repeats grouped to form the excerpt. These grouping processes were demonstrated by the profile of accelerando/ritardando that followed basic groups (or melodic gestures) within phrases, by the fact that the lengthenings of the second phrases were larger than for the first phrases, and by the fact that this characteristic was more pronounced for the second repeat. Of particular interest here is the similarity of the profile of notes within basic groups and of the profile of these basic groups within phrases. Figure 14 presents polynomial interpolations for the basic group profile within phrases. The best-fitting polynomials were adapted to the data, which retained an accelerando/ritardando profile: 2nd-degree polynomials for the first phrases and 3rd-degree polynomials for the second phrases. (We do not wish to enter into the debate about the best-fitting curve: parabol, linear, or other, but rather wish to emphasise the importance of slowing towards the end of groups.) The same ritardando profile (a truncated accelerando/ritardando profile) is also found at higher levels in the hierarchy: phrases within repeats (in each repeat, the second phrase was played slower than the first one) and repeats within the excerpt (the second repeat was played slower than the first one). We predict that the same accelerando/ritardando profile would be found at higher hierarchical levels of musical sections. (Unfortunately, the excerpt analyzed here was not sufficiently long to allow the demonstration.)

We therefore suggest that similar psychological processes are operating at multiple hierarchical levels – namely, those of segmentation and grouping – and that these similar processes result in the same type of performance variations (an accelerando/ritardando profile). We are currently testing this proposed model by asking musicians to segment a musical sequence from a score and from a mechanical performance into increasingly larger units and by comparing at different hierarchical levels of analysis this segmentation with the accelerandos and ritardandos they produce when performing the sequence.

What is the origin of such a universal phenomenon of slowing towards the end of perceived groups? Several authors (Kronman & Sundberg, 1987; Todd, 1995) have suggested that phrase-final lengthening is related to an even more universal principle: perceived motion inherent to all human behaviour. One extension to the present theoretical framework would suggest that lengthening at the end of basic groups is related to perceptual limitations of the auditory system (the last interval of a perceived basic group is perceived shorter and thus performed longer to sound regular, producing a performance variation related to a low-level process), and that the lengthening pattern of accelerando/ritardando would generalize to increasingly longer units as listeners become increasingly familiar with particular hierarchical organizations such as those present in music, producing variations related to higher-level processes. Further research would be necessary to address this issue in detail.

In sum, the psychological segmentation model of music performance provides a framework in which to interpret a wide range of systematic performance variations observed at multiple hierarchical levels: the same pattern of accelerando/ritardando appears to apply to segmentation at these multiple hierarchical levels.

Acknowledgements We wish to thank E. Clarke, W. Printz, B. Repp, and an anonymous reviewer for helpful comments on previous manuscripts.

Annex *Träumerei* score (G. Henle Verlag)

References

Bamberger, J. (1980). Cognitive structuring in the apprehension and description of simple rhythms. *Archives of Psychology*, *48*, 177–199.

Baruch, C., & Drake, C. (in press). Tempo discrimination in infants. *Infant Behavior and Development*.

Botte, M. C., McAdams, S., & Drake, C. (1995). Perception des sons et de la musique. In B. Lechevalier & F. Eustache (Eds.), *Perception et Agnosies: Séminaire Jean-Louis Signoret* (pp. 55–99). Bruxelles: De Boeck.

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

Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 1–25). Oxford: Science Publications.

Clarke, E. F., & Krumhansl, C. (1990). Perceiving musical time. *Music Perception*, *7*, 213–252.

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

Deliège, I. (1990). Mechanisms of cue extraction in musical grouping: A study of Sequenza VI for Viola Solo by L. Berio. *Psychology of Music*, *18*, 18–45.

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

Drake, C. (1990). *Processus cognitifs impliqués dans l'organisation du rythme musical*. Unpublished doctoral dissertation, Université René Descartes, Paris.

Drake, C. (1993a). Reproduction of musical rhythms by children, adult musicians and adult nonmusicians. *Perception and Psychophysics*, *53*, 25–33.

Drake, C. (1993b). Perceptual and performed accents in musical sequences. *Bulletin of the Psychonomic Society*, *31*, 107–110.

Drake, C. (1993c). Influence of age and musical experience on timing and intensity variations in reproductions of short musical rhythms. *Psychologica Belgica*, *33*, 217–228.

Drake, C. (1997). Motor and perceptually preferred synchronisation by children and adults: Binary and ternary ratios. *Polish Quarterly of Developmental Psychology*, *3*, 41–59.

Drake, C., & Palmer, C. (1993). Accent Structures in Music Performance. *Music Perception*, *10*, 343–378.

Drake, C., Botte, M. C., & Gérard, C. (1989). A perceptual distortion in simple musical rhythms. *Proceedings of the fifth annual meeting of the International Society for Psychophysics, Cassis, France*, 108–113.

Drake, C., Dowling, W. J., & Palmer, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists, *Music Perception*, *8*, 315–334.

Fitzgibbons, P. J., Pollatsek, A., & Thomas, I. B. (1974). Detection of temporal gaps within and between perceptual tonal groups. *Perception and Psychophysics*, *16*, 522–528.

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.

Handel, S. (1989). *Listening*. MIT Press: London.

Honing, H. (1990). POCO: An environment for analyzing, modifying, and generating expression in music. *Proceedings of the 1990 International Computer Music Conference* (pp. 364–368). San Francisco: Computer Music Association.

ten Hoopen, G., Hartsuiker, R., & Sasaki, T. (1995). Auditory isochrony: Time shrinking and temporal patterns. *Perception*, *24*, 577–593.

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* (pp. 57–68). Stockholm: Royal Swedish Academy of Music.

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

Nakamura, T. (1987). The communication of dynamics between musicians and listeners through musical performance. *Perception and Psychophysics*, *41*, 525–533.

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

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

Palmer, C. (1996b). On the assignment of structure in music performance. *Music Perception*, *14*, 23–57.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 728–741.

Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 457–470.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409–464.

Povel, D. J. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411–440.

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*, 622–641.

Repp, B. H. (1992a). Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. *Cognition*, *44*, 241–281.

Repp, B. H. (1992b). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's *Träumerei*. *Journal of the Acoustical Society of America*, *92*, 2546–2568.

Repp, B. H. (1992c). Detectability of rhythmic perturbations in musical contexts: Bottom-up versus top-down factors. *Proceedings of the Fourth Workshop on Rhythm Perception and Production, Bourges, France* 111–116.

Repp, B. H. (1995a). Expressive timing in Schumann's "Träumerei": An analysis of performances by graduate student pianists. *Journal of the Acoustical Society of America*, *98*, 2413–2427.

Repp, B. H. (1995b). Acoustics, perception, and production of legato articulation on a digital piano. *Journal of the Acoustical Society of America*, *97*, 3862–3874.

Repp, B. H. (1995c). Detectability of duration and intensity increments in melody tones: A partial connection between music perception and performance. *Perception and Psychophysics*, *57*, 1217–1232.

Repp, B. H. (1996a). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America*, *100*, 3917–3932.

Repp, B. H. (1996b). Pedal timing and tempo in expressive piano performance: A preliminary investigation. *Psychology of Music*, *24*, 199–221.

Repp, B. H. (in press). Variations on a theme by Chopin: Relations between perception and production of timing in music. *Journal of Experimental Psychology: Human Perception and Performance*.

Seashore, C. E. (1938). *Psychology of music*. New York: McGraw-Hill. (Reprinted 1967, New York: Dover Publications.)

Semjen, A., Garcia-Colera, A., & Requin, J. (1984). On controlling force and time in rhythmic movement sequences: The effect of stress location. In *Timing and time perception*, *423*, 169–182.

Shaffer, L. H. (1981). Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, *13*, 326–376.

Shaffer, L. H., & Todd, N. P. (1987). The interpretive component in musical performance. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 139–152). Stockholm: Royal Swedish Academy of Music.

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

Sloboda, J. A. (1985a). Expressive skill in two pianists: Metrical connection in real and stimulated performances. *Canadian Journal of Psychology*, *39*, 273–293.

Sloboda, J. A. (1985b). *The musical mind: The cognitive psychology of music*. Oxford: Oxford University Press.

Sundberg, J., & Verrillo, V. (1980). On the anatomy of the retard: A study of timing in music. *Journal of the Acoustical Society of America*, *68*, 772–779.

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

Thorpe, L. A., & Trehub, S.E. (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology*, *25*, 122–7.

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

Todd, N. P. McAngus (1989). A computational model of rubato. *Contempory Music Review*, *3*, 69–88.

Todd, N. P. McAngus (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America*, *91*, 3540–3550.

Todd, N. P. McAngus (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*, *97*, 1940–1949.

Vos, P. G. (1977). Temporal duration factors in the perception of auditory rhythmic patterns. *Scientific Aesthetics*, *1*, 183–199.