

The dynamics of expressive piano performance: Schumann's "Träumerei" revisited 1012

Bruno H. Repp^{a)}

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511-6695

(Received 13 October 1995; revised 25 March 1996; accepted 26 March 1996)

Ten graduate student pianists were recorded playing Robert Schumann's "Träumerei" three times on a Yamaha Disclavier. Their expressive dynamics were analyzed at the level of hammer (MIDI) velocities. Individual dynamic profiles were similar across repeated performances, more so for the right hand (soprano and alto voices) than for the left hand (tenor and bass voices). As expected, the soprano voice, which usually had the principal melody, was played with greater force than the other voices, which gained prominence only when they carried temporarily important melodic fragments. Independent of this voice differentiation, there was a tendency for velocity to increase with pitch, at least in the soprano and alto voices. While there was an overall tendency for velocities to increase with local tempo, there were salient local departures from this coupling. Individual differences in expressive dynamics were not striking and were only weakly related to individual differences in expressive timing. © 1996 Acoustical Society of America.

PACS numbers: 43.75.St, 43.75.Mn [WJS]

INTRODUCTION

Variation in timing and dynamics are the two primary means a performer has available to make music expressive and communicative. Together they account in large measure for the gestural quality or "motion" in music that engages the listener (Truslit, 1938; Todd, 1992; Repp, 1993b). This is especially true on a keyboard instrument such as the piano, which essentially fixes the pitch, timbre, and amplitude envelope of each note. The artistic control of dynamics presents a special challenge for pianists who must deal with several structural layers at once. Yet, expressive dynamics have been given far less attention than expressive timing in scientific studies of piano performance.

The relative intensities with which notes are to be played are indicated only to a very limited extent in the score. Global prescriptions of dynamics such as *forte*, *mezzoforte*, or *piano* are imprecise and relative, comparable in that respect to tempo prescriptions such as *allegro* or *andante*. However, while tempo can be indicated and measured precisely using a metronome, there is no analogous method of calibrating dynamic level in music. (Of course, such a calibration would make little sense, since dynamic level must be flexibly adjusted to the instrument, room size, room acoustics, etc.) At a more local level in a score, one finds hairpins or verbal instructions of *crescendo* and *decrescendo* (or *diminuendo*), which tell the performer to gradually increase or decrease the dynamic levels of successive tones. However, the precise manner in which this is to be done (i.e., the extent and rate of the increase or decrease) is never specified. At the level of the individual note or chord, marks such as *sforzato* or a wedge are occasionally employed to indicate a dynamic accent. Finally, there are the ubiquitous bar lines, which convey the metrical structure. Notes following a bar line usually have a strong metrical accent, and notes occurring in the

temporal center of a bar may receive a secondary accent (metrical subdivision). These are theoretical accents, however, that are not necessarily to be realized by an increase in intensity. If the music has a dance-like or motoric character, regular accents on the downbeat may be appropriate, but in more lyrical and expressive styles this usually leads to undesirable monotony (cf. Thurmond, 1983).

Clearly, there are only very rough guidelines to dynamics in the score, some of which (the bar lines) may even be misleading, and it is up to the performer to make the right choices and provide the fine detail. In particular, performers must decide on the basis of what "feels right" to them how strongly to play tones marked as accented, how to shape the dynamic progression of a *crescendo* or *decrescendo*, how to give expressive melodic gestures appropriate dynamic shapes, and how to give repetitive rhythmic figures a characteristic dynamic profile. These are aspects of "horizontal" dynamics, applied to successive notes. In addition, there are aspects of "vertical" dynamics to consider, which apply to simultaneous notes. Particularly important here are the emphasis of the principal melody over less important voices and the proper "voicing" of chords to make them sound rich and balanced. Consequently, the intricate dynamic microstructure of a performance reveals far more about the performer's skill, taste, and grasp of the musical structure than about his or her observance of prescriptions in the score.

Dynamic shaping and differentiation are difficult and often neglected aspects of the pianist's skill. They involve risks at both extremes of the range (i.e., the possible production of either an ugly or an inaudible tone) and require adaptation to the instrument and to room acoustics. Most of all, they require exquisite motor control and a sensitive ear to guide the hands and fingers. For these reasons, even highly skilled pianists' control over dynamics may be less precise than their control over timing (though level of precision is difficult to compare across different dimensions), and masters of dynamic shading are rare and highly esteemed.

^{a)}Electronic mail: repp@haskins.yale.edu

Like expressive timing, expressive dynamics can be measured at several different levels: kinematic, acoustic, and perceptual. The present investigation is restricted to the kinematic level—the varying forces of the pianist's finger movements on the keyboard, as reflected in the registered hammer velocities. Such hammer velocity registrations, using photographic means, were obtained already in Carl Seashore's laboratory (Henderson, 1936; Seashore, 1938), though they were not analyzed in great detail and derived in part from music that did not call for much dynamic differentiation. Some of Seashore's basic observations were that (1) similar musical sections tend to show similar dynamic patterns in performance, (2) metrically accented notes are not necessarily played with greater intensity than unaccented notes, and (3) melody notes are played with greater intensity than the accompaniment. (However, this last distinction was confounded with pitch register, the melody being higher than the accompaniment, and to some extent with hand as well, since the melody was usually played by the right hand.)

Similar apparatus to record the piano hammer action during performance was developed later by Henry Shaffer at the University of Exeter, but his subsequent analyses and publications dealt primarily with timing. Shaffer (1981) examined a single performance of a Chopin Etude consisting of a continuous three (right hand melody) against four (left hand accompaniment) rhythm. He noted that the right hand played louder and had a wider dynamic range than the left hand, and that accents were made independently in each hand. Particularly interesting was his finding (implicit in the fact that he fitted straight regression lines to the data) that *crescendi* and *diminuendi* exhibit a linear sequence of intensity values. He calculated these values as the inverses of the upward transit times of the hammers. Hence they were equivalent to hammer velocities in m/s, which more recently have been shown to be linearly related to the peak amplitude of recorded piano tones (Palmer and Brown, 1991).

The nature of dynamic change in piano performance was examined more closely by Todd (1992). Guided by measurements obtained in Shaffer's laboratory as well as by some acoustic data of Gabrielsson (1987), he proposed a model of expression that links dynamics closely with timing variation. Music is often described as consisting of cycles of tension and relaxation, where increasing tension is manifested overtly in both *accelerando* and *crescendo*, whereas increasing relaxation is associated with both *ritardando* and *diminuendo*. Although this coupling is often violated, Todd considers it a default mode that applies whenever there are no contrary instructions in the score. His test cases were two performances of Chopin's *Prélude* in f-sharp minor, a piece in which the dynamics indeed seemed to follow the timing variations fairly closely at the beat level: The correlation between local tempo and intensity (averaged over all tones within a beat) was about 0.7.¹ Todd also noted the consistency of two different performances by the same pianist, though it was lower for dynamics ($r=0.85$) than for timing ($r=0.97$).

On the basis of his admittedly limited data, Todd proposed a model for the covariation between timing and intensity. The model starts with an analysis of the hierarchical

grouping structure (Lerdahl and Jackendoff, 1983) of the composition, as did Todd's (1985, 1989) earlier model of expressive timing variation. Each group within this hierarchical structure is assumed to have a *crescendo-decrescendo* shape composed of two linear segments (in terms of the intensity measure used, i.e., hammer velocity), with the temporal location and magnitude of the peak intensity being free parameters. The dynamic shapes of superordinate groups are then linearly superimposed upon those of subordinate groups. Using an analysis-by-synthesis approach, Todd was able to adjust the free parameters in his model until it fit the Chopin *Prélude* intensity data about as well as the pianist's two performances resembled each other.

Todd's model is an important advance, but it is in need of testing with more extensive data. Although piano performance data are now relatively easy to obtain, thanks to MIDI technology and computer-controlled pianos, little use has been made of these facilities so far for research purposes, particularly with regard to dynamics. Some researchers, however, have obtained amplitude information from acoustic recordings by reading peak amplitudes off visual displays of the waveform envelope (Truslit, 1938; Gabrielsson, 1987), by using a level recorder (Nakamura, 1987), or by computing the root-mean-square (rms) amplitudes of digitized signals (Kendall and Carterette, 1990). These measures naturally include transformations imposed on the radiated sound by soundboard resonances and room acoustics, which introduce considerable variability specific to the instrument and the recording situation (Repp, 1993a). Also, whenever there are several simultaneous tones, their amplitudes are superimposed (but not necessarily additive).

The most interesting of these acoustic studies in the present context is that of Gabrielsson (1987), because it compared five different performances by well-known pianists with regard to both timing and intensity patterns. The music was limited to the first eight measures of Mozart's *Sonata* in A major, K.331, with a repeat available for three pianists. The similarity of the amplitude profiles for the repeated passages was striking, although no correlations were reported. There was also considerable similarity of dynamic patterns across artists. Each of the two 4-bar phrases in the music showed a pronounced amplitude peak near the end, followed by a rapid *decrescendo*. This temporal asymmetry, which inspired Todd (1992) to include a "peak shift" option in his model, reflects the motivic and harmonic structure of the composition. The time course of *crescendi* and *decrescendi* does not seem linear in Gabrielsson's data. The correlation between timing and amplitude profiles was not reported, but it is clear that the phrase-final *decrescendo* was coupled with a *ritardando*, whereas earlier in each phrase there was much less correspondence. There appeared to be a lot of fine detail in the amplitude profiles, though it is impossible to tell whether this was intended by the pianist or caused by irregularities in instrument response and sound transmission.

Another phenomenon needs to be mentioned, and that is an increase in amplitude with pitch. Already observed by Truslit (1938) in the expressive playing of scales, it is one of the intuitive rules incorporated in the performance synthesis systems of Sundberg (1988; Friberg, 1991) and Clynes

(1987). In Sundberg's system, the increase is about 3 dB per octave, regardless of the instrumental timbre it is applied to. Clynes developed his system using pure tones and only more recently applied it to realistic instrument sounds. According to him, the "pitch crescendo" should be composer-specific, with very little for Beethoven but as much as 6 dB per octave for Schubert. Different values may apply to piano sounds than to pure tones. Surprisingly, there seem to be no data in the literature on the relation between pitch and intensity (or hammer velocity) in actual piano performance, except for a recent study by Palmer (in press) who found only a negligible correlation.

From this review of studies of performance dynamics, a few consistent observations emerge:

- (1) Repeated performances of the same music generally have highly similar dynamic patterns.
- (2) Like timing, dynamic microstructure seems to reflect the hierarchic grouping structure of the music, with *crescendo-decrescendo* patterns within phrases.
- (3) There seems to be a coupling of timing and dynamics, which is most evident in phrase-final *ritardandi/decrescendi*.
- (4) The change of successive hammer velocities during a *crescendo* or *decrescendo* may be a linear function of metrical time.
- (5) Hammer velocity may increase with pitch.

All statements but the first are tentative because of the very limited amount of published data; they may be considered hypotheses in need of further test. The following analyses will address some of them in a more extensive performance data base than has been examined so far. The music chosen, Robert Schumann's "Träumerei," is particularly well suited to an investigation of expressive dynamics because of its polyphonic structure and its slow tempo, which enables pianists to exercise precise control over each individual tone. Repp (1994) analyzed two pianists' performances of "Träumerei" at three different tempos and found expressive dynamics to be quite stable within and across tempos. The focus here is on a different set of performances, played by ten graduate student pianists and recorded in MIDI format. A detailed comparison of their timing patterns with those of famous pianists' performances has demonstrated that the student pianists have excellent control over expressive timing and are distinguished from the famous artists mainly in terms of their greater conservatism and smaller individual differences (Repp, 1995). This relative homogeneity—if it applies to dynamics as well—is in fact advantageous for a study in which typical patterns of expression are of primary interest.

I. METHOD

A. The music

A computer-generated score of Schumann's "Träumerei" is shown in Fig. 1. The layout on the page illuminates the phrase structure, which is discussed in more detail in Repp (1992). Metrical positions in the music will be

referred to by bar number, beat number, and half-beat number; thus "13-3-2" refers to the second eighth note of the third beat in bar 13.

B. The pianists

Ten pianists participated in the study as paid volunteers. Nine of them were graduate students of piano performance at the Yale School of Music; five were in their first year, one was in her second year, and three were third-year (artist's diploma) students. The tenth pianist was about to graduate from college and had been accepted into the piano graduate program for the coming academic year. The pianists' age range was 21 to 29, and they had started to play the piano between the ages of 4 and 8. Seven were female, three male. They will be identified by numbers prefaced by the letter P (for pianist).

C. Recording procedure

The recording session took place in a fairly quiet room housing an upright Yamaha MX100A Disclavier connected to a Macintosh computer. The music to be played included four pieces, one of which was "Träumerei." The pianist was given the music and asked to rehearse it at the Yamaha for an hour. Subsequently, the four pieces were recorded once, in whichever order the pianist preferred, and then the cycle was repeated twice. If something went seriously wrong in a performance, it was repeated immediately. One pianist, P4, as a result of several retakes and a computer problem, was able to record only two performances of each piece; all others recorded three, as planned. All performances were fluent and expressive. At the end of the session, each pianist filled out a questionnaire and was paid \$50.

The questionnaire asked the pianists, among other things, to rate the adequacy of the Yamaha Disclavier in terms of sound quality and responsiveness. From among the five categories provided, nobody chose "excellent;" the choices were "very good" (2), "good" (4), "adequate" (2), and "poor" (2). The questionnaire also asked in some detail how well the pianists knew each of the pieces. Schumann's "Träumerei" had been previously studied by three (P5, P7, P8) and played informally by two; the rest knew it from listening only. The pianists were also asked to indicate how satisfied they were with their performances, choosing from the categories "best effort," "good effort," "average," "below average," and "poor." The distribution of their responses for "Träumerei" was 0/4/5/1/0.

D. Data analysis

The MIDI data were imported as text files into a Macintosh spreadsheet and graphics program (Deltagraph Professional), where the note onsets were separated from the other events (note offsets and pedal actions) and labeled with reference to a numerical (MIDI pitch) transcription of the score. In that process, wrong pitches (substitutions) were identified and corrected, omitted notes were supplied, and added notes (intrusions) were removed. An analysis of these errors is presented in Repp (submitted). Of the four pieces,

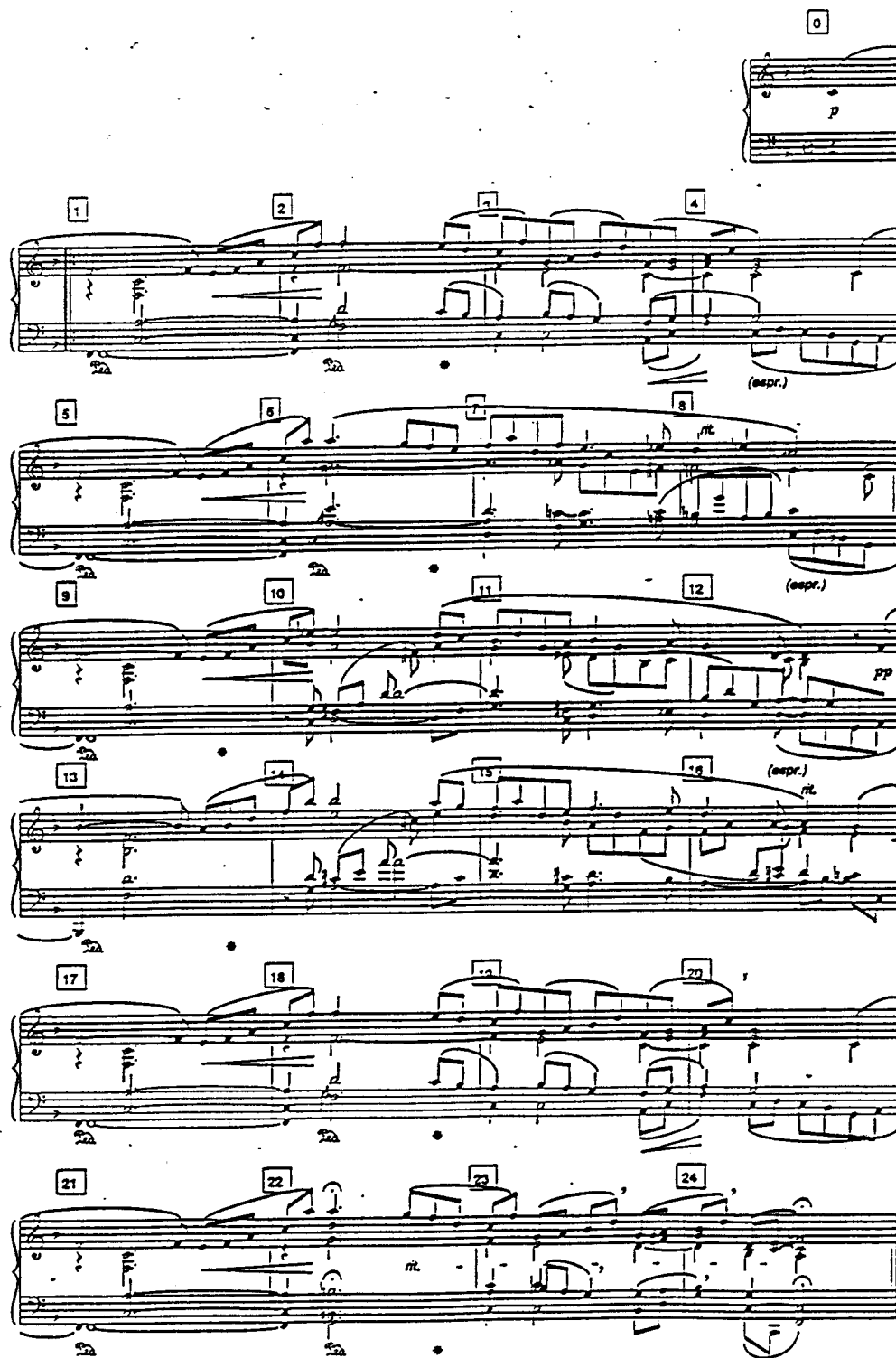


FIG. 1. The score of Robert Schumann's "Träumerei," op. 15, No. 7.

"Träumerei" had the smallest number of errors. The 29 performances contained a total of 101 omissions (0.79%), most of which were inner notes of chords.

The parameter of interest in this study was MIDI velocity, or velocity for short, which has a theoretical range from 0 to 127. Its value increases monotonically with hammer velocity, which is picked up by two sensors in the Disclavier, but the precise functional relationship is not known. The re-

lationship between MIDI velocity and peak rms sound level on the Yamaha Disclavier was determined in the course of previous research (Repp, 1993a): For any given pitch, the relationship was nearly linear over the range examined (20–100), though somewhat negatively accelerated toward the higher velocities, with between 3 and 4 velocity units corresponding to 1 dB of sound level for a given pitch.² Outside this range nonlinearities do occur. *Fortissimo* notes did not

occur in the present performances, but *pianissimo* notes did. Most notes examined, however, had velocities between 20 and 80 and thus were free from any irregularities.

After the initial stages of data analysis, the velocity data for the three performances of each pianist (two for P4) were lined up and averaged to yield an average velocity (or dynamic) profile. From the resulting ten individual average profiles a grand average profile was then computed, which captures the features shared by most of the performances. This overall profile was then parsed horizontally into "melodic gestures" (Repp, 1992) and vertically into voices. Because of the polyphonic construction of the piece, four voices (soprano, alto, tenor, and bass) can be distinguished throughout, with only a few "secondary" notes not fitting into this scheme. The 457 notes in the piece were assigned to voices as follows: 179 soprano notes (166 primary, 13 secondary), 79 alto notes (67 primary, 12 secondary), 106 tenor notes (89 primary, 17 secondary), and 93 bass notes. A note was considered secondary when it accompanied a seemingly more important (primary) note in the same voice. The principal melody is in the soprano most of the time, but it cascades down through the other voices in bars 7–8, 11–12, and 15–16.

E. Use of the soft pedal

It is common practice among pianists to depress the soft pedal during quiet passages, even though Banowetz (1985) cautions against using the pedal as a substitute for true *piano* playing. Individual differences in the use of the soft pedal were considerable. Five of the pianists (P1, P5, P7, P9, P10) used it almost continuously, with only occasional brief releases that tended to occur in the same places, probably to highlight a chord or brief passage. P6 used the soft pedal in bars 1–8 and from bar 13 on, but not during the repeat of bars 1–8 and during bars 9–12 (except in her first performance). P2 and P3 used the soft pedal only from bar 13 on (where a *pp* appears in the score), and intermittently after bar 18. In the first performance, P3 did not use the soft pedal at all. P4, too, did not make use of the soft pedal, except for a brief episode at the beginning of the first performance. The most unusual case was P8 who used the soft pedal frequently, but only for brief periods, so that it was off most of the time. This strategy was just the opposite of that of the five pianists who depressed the pedal most of the time. Apparently, P8 used the soft pedal to color specific chords or passages, such as the downbeat and the following chord in bars 1, 5, 9, 13, 17, and 21.

Given these patterns of soft pedal use, the question naturally arose whether the velocity data should be "corrected" to take into account the reduction in sound level caused by the soft pedal. To determine the magnitude of this effect, isolated tones ranging from C2 to C6 in three-semitone steps, with a fixed arbitrary MIDI velocity of 60, were produced on the Disclavier under MIDI control and were recorded with a microphone, with the soft pedal either raised or depressed. The sounds were digitized, and their peak rms sound levels were measured from their amplitude envelopes. Surprisingly, there was no systematic effect of pedaling, and pairs of tones

TABLE I. Average correlations among the MIDI velocities in three performances: (a) all notes ($n=457$), (b) right-hand notes only ($n=257$), (c) left-hand notes only ($n=200$), (d) soprano voice only ($n=179$), (e) all notes, bars 1–8 only ($n=113$). Also, (f) average within-performance correlations for the two renditions of bars 1–8 ($n=112$).

Pianist	(a)	(b)	(c)	(d)	(e)	(f)
P1	0.853	0.837	0.710	0.788	0.877	0.869
P2	0.888	0.903	0.747	0.887	0.832	0.857
P3	0.797	0.792	0.754	0.736	0.873	0.755
P4	0.865	0.838	0.740	0.878	0.872	0.855
P5	0.825	0.847	0.551	0.810	0.824	0.809
P6	0.850	0.843	0.722	0.803	0.869	0.831
P7	0.825	0.793	0.696	0.740	0.870	0.816
P8	0.897	0.911	0.782	0.845	0.874	0.916
P9	0.831	0.813	0.793	0.791	0.850	0.869
P10	0.732	0.702	0.655	0.599	0.835	0.836
Mean	0.836	0.833	0.715	0.788	0.863	0.841

differed by less than 1 dB.³ Therefore, no correction was applied for the pianists' use of the soft pedal.

II. RESULTS AND DISCUSSION

A. Reliability

We begin with a consideration of the replicability of dynamic microstructure across repeated performances of the same music. Each performance contained maximally 457 velocity values. (The velocities of omitted notes were left unspecified.) The three performances of each pianist yielded three between-performance correlations whose average was subsequently computed. (For P4 there was only a single correlation.) These average correlations are listed in the first column of Table I. They ranged from 0.732 to 0.897, with a mean of 0.836. Thus they were distinctly lower than the same pianists' between-performance correlations of timing microstructure, whose average was 0.947 (Repp, 1995).⁴ Interestingly, there was a high correlation ($r=0.826$, $p<0.01$) between the two sets of reliabilities: Consistency in timing went hand in hand with consistency in dynamics. The pianist with the highest reliabilities in both domains, P8, happened to be the one who played at the slowest tempo. Overall, however, there was no significant relationship between average tempo and dynamic consistency ($r=-0.153$). Also, there seemed to be no relationship of reliability to familiarity: P5 and P7, who, like P8, had studied "Träumerei" at some time in the past, did not show equally high reliabilities.

Table I also shows the reliabilities for the right and left hands separately. The correlations for the right hand (which played the soprano and alto voices) were as high as those for all notes combined, but those for the left hand (which played the tenor and bass voices) were a good deal lower. This could reflect poorer dynamic control in the left than in the right hand, but it could also be due to the lesser importance of the lower voices and/or to a more restricted dynamic range for them (see below). Next, Table I lists the reliabilities for the most important voice, the soprano voice, alone. These were somewhat lower than those for the whole right hand, which may again be due to a restriction of the dynamic range, as the pianists surely gave special attention to the soprano voice. The dynamic reliabilities for the soprano

TABLE II. Mean velocities and standard deviations (in parentheses) for all notes and for each voice separately.

Pianist	All voices	Soprano	Alto	Tenor	Bass
P1	30.4(11.5)	39.3 (9.3)	27.4 (11.0)	22.3 (7.3)	25.0 (6.7)
P2	40.5(14.1)	51.5 (11.7)	34.1 (12.6)	32.4 (10.0)	34.1 (8.8)
P3	33.2(14.0)	41.8 (13.2)	26.4 (12.3)	24.9 (9.1)	31.9 (12.3)
P4	36.6(12.8)	45.8 (10.8)	30.8 (11.6)	29.3 (9.6)	32.0 (9.9)
P5	32.6(11.5)	41.3 (10.0)	30.3 (10.4)	25.5 (7.2)	25.6 (6.7)
P6	34.3(12.6)	43.8 (9.6)	31.8 (11.4)	26.0 (9.0)	27.6 (9.9)
P7	29.0(13.0)	40.0 (9.7)	24.2 (8.8)	19.0 (6.8)	23.2 (11.2)
P8	47.0(14.7)	58.2 (10.0)	41.9 (16.3)	38.5 (12.3)	39.3 (8.8)
P9	33.9(13.8)	43.2 (11.1)	28.7 (14.1)	26.9 (11.6)	28.5 (10.5)
P10	29.9(13.1)	39.6 (10.7)	23.9 (10.9)	22.4 (9.1)	24.9 (11.5)
Mean	34.7(13.1)	44.5 (10.6)	30.0 (11.9)	26.7 (9.3)	29.2 (9.6)

voice are more directly comparable to the timing reliabilities, which likewise were based only on the highest notes in each chord (Repp, 1995a), but the conclusion remains the same: The dynamic pattern was less reproducible than the timing pattern.

There could be yet another reason for this difference, which is the presence of large *ritardandi* at phrase endings, which inflate the reliabilities for timing. A fair comparison would consider bars 1–8 only, which do not show such extreme timing deviations and therefore have lower timing reliabilities. Therefore, Table I also lists the dynamic reliabilities for all notes in these initial bars (not including their repeat). They are somewhat higher than the dynamic reliabilities for the piece as a whole, but they are still lower (average of 0.863) than the timing reliabilities for bars 1–8 (average of 0.907). Finally, Table I shows the average dynamic *within*-performance reliabilities for bars 1–8. They are somewhat lower than the between-performance reliabilities, suggesting that at least some pianists (most notably P3) intended to play the repeat differently. A similar decrease in reliability was observed for timing (Repp, 1995), but again the average within-performance timing reliability (0.899) was greater than the average within-performance dynamic reliability (0.841). A similar difference was reported by Palmer (in press) for the repeat in a Mozart Sonata, played by a well-known concert pianist.

B. Dynamic range

We turn now to the dynamic levels and ranges of the performances, both overall and for the individual voices. From this point on, we will no longer consider the three individual performances of each pianist but only their average. The relevant data are shown in Table II. The first two columns show the mean velocities of all notes and their standard deviations. It is evident that two pianists (P8, P2) played a good deal louder than the others, who played in what seems an appropriate range for the *piano* prescribed in the score. The overall dynamic ranges of the pianists were fairly similar and in the vicinity of 13 dB.⁵

The means and standard deviations for the separate voices are shown in the remaining columns of the table. Not surprisingly, all pianists played the soprano voice more strongly than the other voices; the difference from the alto voice was 14.5 velocity units on the average, or about 3.6

dB. Likewise, all pianists played the alto voice (right hand) more strongly than the tenor voice (left hand), although the average difference was small, only 3.3 velocity units or about 0.8 dB. Finally, all pianists played the bass voice somewhat more strongly than the tenor voice, the average difference being 2.5 velocity units or about 0.6 dB. The alto and bass voices were of similar average intensity. As to dynamic range, the alto voice actually exceeded the soprano voice, which in turn had a wider range than the two lower voices.

C. Dynamic level and pitch

Given the relative prominence of the soprano voice, there was clearly an overall relationship between pitch height and dynamic level. This was confirmed by computing correlations between MIDI pitch and velocity for each individual average performance. These correlations ranged from 0.40 to 0.65, with an average of 0.57. However, this relationship could have been due to the pianists' intention to emphasize the principal melody over the other voices. The relation between pitch and dynamic level is therefore better investigated within each voice.

There were indeed positive within-voice correlations between pitch height and velocity, both in the soprano voice (average 0.42, range 0.26 to 0.58) and in the alto voice (average 0.51, range 0.40 to 0.69). In the tenor voice, the correlation was weak (average 0.29, range 0.12 to 0.45), and in the bass it was absent (average -0.04 , range -0.10 to 0.31). It may be concluded, therefore, that there is an increase in dynamics with pitch, but the relationship is not very strong. Moreover, it seems to hold only in the higher voices, which play a more important melodic role. The present correlations are substantially higher, however, than the one obtained by Palmer (in press) for the melody voice in a Mozart Sonata, which suggests that style and/or piece-specific structure play a role.

D. Intercorrelations

The reliabilities, which were computed between and within each individual pianist's performances, may now be compared with the intercorrelations among different pianists' average performances. Across all notes ($n=457$), these intercorrelations ranged from 0.614 to 0.847, with a mean of 0.748. Thus all pianists' performances were similar to each other in terms of dynamics, but they were less similar than each pianists' multiple performances were to each other (average correlation of 0.836). This was so despite the fact that random variability was reduced in the average performances. Only one pianist (P10) showed higher correlations with several other pianists' average performances than among his own individual performances. Three other pianists showed a higher correlation than their own reliability with just one other pianist, P1 in each case.

The intercorrelations for the soprano voice only ($n=179$) yielded a similar picture. They ranged from 0.452 to 0.814, with an average of 0.644. The average individual soprano voice reliability, by comparison, was 0.788. Three pianists (P3, P7, P10) correlated more highly with other pianists

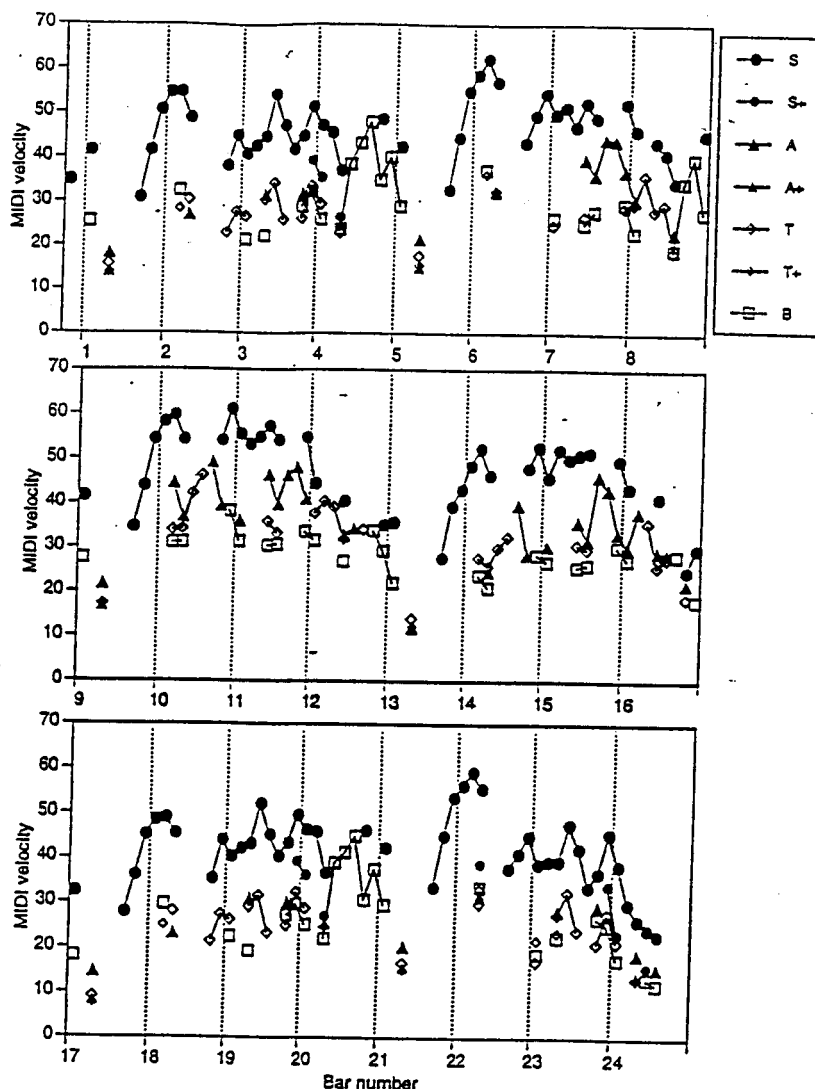


FIG. 2. The grand average dynamic profile for four voices (S=soprano, A=alto, T=tenor, B=bass) and subsidiary notes (S+, A+, T+).

than within themselves. Although the uniqueness of each individual performer's expressive profile had been more striking in the timing domain (Repp, 1995), there is evidence for individuality of dynamic profiles as well.

Both sets of intercorrelations (overall and soprano only) were subjected to principal components analysis and, as for timing (Repp, 1995), only a single significant component emerged in each case. This indicates that there was a single underlying dynamic pattern that all performances had in common, and that individual profiles represented variations around this common standard. Therefore, the grand average dynamic profile is representative of the group of pianists as a whole.

E. The grand average dynamic profile

This profile was obtained by averaging the velocities across the individual average performances of the ten pianists. Furthermore, the velocities for the two renditions of bars 1-8 were averaged. The grand average dynamic profile is shown in Fig. 2. The different voices are represented by different symbols. Contiguous eighth notes in the same voice

are connected. The close similarity of the patterns in bars 1-4 and 17-20 should be noted; they represent identical passages.

The initial quarter-note upbeat (0-4-1) was played softer than the following downbeat. However, when the quarter-note upbeat recurred, overlapping the momentarily prominent bass voice (4-4-1, 20-4-1), it was played more strongly than the following downbeat. This was also true for the eighth-note upbeat at 8-4-2. At the point of modulation to B-flat major (12-4-2, 13-1-1), where a *pp* is indicated in the score, both upbeat and downbeat were played at a similar, lower intensity, though still more strongly than the bass voice. The grace-note upbeat (16-4-2) was also dynamically close to the following upbeat (17-1-1), which initiates the reprise and was played much more softly than its analogues at 1-1-1 and 9-1-1. The dynamic relationship of phrase-initial upbeat and downbeat thus was sensitive to context.

The bass note accompanying the downbeat was always considerably softer. The following 4-note chord was even softer. Its constituent tones were not much differentiated, though the highest of them (assigned here to the alto voice)

was usually the strongest. The exception was the chord in B-flat major (13-2-1), in which all tones were about equally strong.⁶

The following 6-note ascent to the melodic peak in the soprano voice (across bar lines 2, 6, 10, 14, 18, and 22) had a characteristic dynamic shape that was repeated in each of the six phrases. There was a strong *crescendo* over the initial three tones, followed by a much smaller increase to the (unaccented) downbeat and the subsequent pitch peak. The final tone, which repeats and prolongs the highest pitch, was played slightly softer than the preceding tone. The melodies in bars 5-6 and 21-22, which ascend to A₅, were played more loudly and covered a slightly greater dynamic range than those in bars 1-2 and 17-18, which only reach F₅. The one in bars 13-14, which is marked *pp*, was played more softly than that in bars 9-10. It also exhibits a slightly different dynamic shape, probably due to the modulation to B-flat, which necessitated a change in fingering.

Two things are noteworthy about this melodic gesture. First, although intensity increased with pitch, the largest dynamic increase occurred at the beginning, where there was the smallest change in pitch; the largest pitch change (immediately following the downbeat) was accompanied by only a minute dynamic change. This illustrates the loose connection between pitch and dynamics. Second, this phraselet exhibited a pronounced *ritardando*, which culminated at the pitch peak (Repp, 1992, 1995). The simultaneous *crescendo* is contrary to Todd's (1992) observation of a positive covariation between dynamics and tempo. However, the reduction and slight inversion of the *crescendo* toward the end of the *ritardando* could be due to an underlying trend that counteracts the *crescendo*.

The lower tones accompanying the peak of the phrase were all much softer and not greatly differentiated. In bars 10 and 14, an imitative motive begins in the tenor voice and transfers to the alto. Starting softly, it reached its dynamic peak at the point of transfer (a single note played by the right hand, 10-3-2 and 14-3-2) and then dropped to a lower level for the final tone, which coincides with the resumption of the soprano voice (10-4-1, 14-4-1). The dynamic shape of this imitation motive was quite different from that of its soprano model, especially in that it lacked an initial *crescendo*.

Consider now the soprano voice from 2-4-1 to 4-2-1 and from 18-4-1 to 20-2-1, as well as the nearly identical passage from 22-3-2 to 24-1-2. These are the descents from the melodic peak in the phrases that in previous studies were dubbed type A (Repp, 1992, 1995). What is noteworthy here are the dynamic peaks or accents in metrically weak positions immediately preceding strong beats (2-4-2, 3-2-2, 3-4-2, and analogous positions elsewhere). The tones in these positions are harmonically and melodically unstable and move strongly towards the following, more stable pitches; they, not the stable and metrically strong tones, were emphasized by the pianists. Parallel patterns at a lower intensity can be seen in the accompanying tenor voice and, in bars 23-24, also in the supplementary soprano voice and in the bass.

At the end of the type A phrase in bars 4 and 20, the bass voice takes over. Its soft initial tone coincides with the

end of the soprano melody, but the following tones were almost at the dynamic level of the soprano. The dynamic peaks again fell on the less stable tones (4-3-2, 4-4-2, 20-3-2, 20-4-2). However, the lower intensity of the bass notes at 4-4-1 and 20-4-1 could also be due to their coincidence with the soprano upbeat to the next phrase; and their low intensity at 5-1-1 and 21-1-1 could be due to their making way for the soprano note on the downbeat. The final soprano tones in bar 24 and their accompaniments trail off toward the final chord, accompanying the extreme *ritardando* at this point.

The phrases of type B are characterized by an overlapping, cascading descent of the melody through the four voices, from soprano to bass. The soprano line (6-3-2 to 7-3-1, 10-4-1 to 11-3-1, 14-4-1 to 15-3-1) was relatively steady but again exhibited a slight accent on the unstable tone preceding the downbeat (6-4-2, 10-4-2, 14-4-2). A similar accent in positions 7-2-2, 11-2-2, and 15-2-2 was merely hinted at, but it emerged more clearly in the alto voice, which takes over at this point. The alto voice then showed a dynamic increase in positions 7-3-2 and 7-4-1 and their later analogues, where it is not accompanied by any other voice. A strong emphasis on the unstable tone at 7-4-2 followed, seen primarily in the soprano voice, which here continues its melodic line while the tenor picks up the cascading motive. The remainder of the phrases in bars 12 and 16 was mainly a *diminuendo* toward the beginning of the next phrase, whereas in bar 8 the bass voice achieved brief prominence (8-3-2, 8-4-1), probably due to the temporary inactivity of the other voices.

F. Dynamics and timing

Todd (1992) postulated a coupling between timing and dynamics, such that the slower the local tempo, the softer the dynamics. This implies a negative covariation between inter-onset interval (IOI) duration and MIDI velocity. (For a graph of the grand average timing profile, see Repp, 1995, Fig. 2.) Although a local violation of this relationship in the ascent to the melodic peak (bars 1-2, 5-6, 9-10, 13-14, 17-18, and 21-22) has been noted, the overall correlation predicted by Todd nevertheless was confirmed for each individual voice. The coefficients were -0.39 (soprano), -0.52 (alto), -0.49 (tenor), and -0.33 (bass), all significant at $p < 0.01$. Palmer (in press) recently found a similar relationship for the melody notes of a Mozart Sonata, as played by an excellent pianist.

An additional correlational analysis examined whether individual differences in dynamics might be related to individual differences in timing. The grand average timing profile was subtracted from each pianist's individual average timing profile, and likewise the grand average dynamic profile was subtracted from each pianist's individual average dynamic profile, for the soprano voice only. Correlations were then computed between these residuals. The coefficients were negative for eight of the ten pianists and reached significance in six instances, though they were small. P8 showed the highest correlation (-0.40), P5 the second highest (-0.26). Thus there was a slight tendency for pianists to play relatively softly when they played relatively slow (relative to grand average expressive dynamics and timing).

Todd's model also predicts that velocity during a *crecendo* or *decrecendo* changes as a linear function of metrical distance (see also Shaffer, 1981). Clearly, there are some instances in the present data where linearity does not hold, particularly in the "ascent to the melodic peak" (bars 1–2, 5–6, 9–10, 13–14, 17–18, and 21–22). The dynamic change may be linear over the first three notes (except in bar 13), but then the velocity increases only by very small amounts (see Fig. 2). A similar observation may be made about the final *decrecendo* in bars 23–24. There are other places, however, where a linear model would seem to capture dynamic changes in the soprano voice quite well, especially across longer IOIs. The dynamic changes in bars 2, 6, 10, 14, 18, and 22 fit a V-shaped pattern that, in bars 10 and 14, includes the tenor and alto voices while the soprano note is sustained. The three-note sequences crossing bar lines 5 and 21, the five-note sequences in bars 12–13 and from bar 7 into bar 8 are other instances where linearity seems to hold. However, a precise evaluation of the model is difficult in the present context because it is not clear whether all velocity minima and maxima should be taken at face value, and whether and how the influence of other factors (pitch, local accents, presence of other voices) should be taken into account.

III. SUMMARY AND CONCLUSIONS

In the Introduction, five hypotheses were deduced from the literature on expressive dynamics. We may summarize now how the present data bear on these hypotheses.

(1) *Repeated performances of the same music generally have highly similar dynamic patterns.* The present results confirm that the dynamic profile is a stable and replicable characteristic of expressive performance, though its reliability is not quite as high as that of the expressive timing pattern. While timing profiles seem to be unique to individual performers (Repp, 1992, 1995), there were a few instances here in which a pianist's dynamic profile resembled that of another pianist more than it resembled his or her own profile across repeated performances.

An interesting new finding was that the reliabilities of timing and dynamics are correlated. This may indicate a coupling between these two parameters, or it may be due to individual differences in consistency that are reflected in both primary dimensions of expressive microstructure. Also, the dynamics of the left hand were less reliable than those of the right hand. In part this may have been due to the somewhat narrower dynamic range of the left-hand part, though the difference in range was rather small. A genuine difference in dynamic control between the hands would not be surprising in view of the fact that, in most piano music, the right hand is assigned musically more interesting and technically more challenging material than the left hand. Greater attention to the right hand may enhance the difference. Handedness probably did not play a role here: Two pianists (P5 and P7) were left-handed but showed some of the lowest left-hand reliabilities.

(2) *Like timing, dynamic microstructure seems to reflect the hierarchic grouping structure of the music, with crescendo–decrecendo patterns within phrases.* This obser-

vation is also supported by the present findings. Within each of the 4-bar phrases, the intensity of the melody rose steeply during the first half (the antecedent part) and fell gradually during the second half (the consequent part). However, this asymmetry parallels the pitch contour of the principal melody and thus may have been due in part to a covariation of dynamics with pitch (see below).

(3) *There seems to be a coupling of timing and dynamics, which is most evident in phrase-final ritardandi/decrecendi.* This hypothesis was supported in an overall analysis. However, although phrase-final *ritardandi* were indeed accompanied by *decrecendi* (bars 12, 16, and 24), the phrase-medial *ritardandi* were coupled with *crecendi*. This departure from the default pattern may have been due to the steeply rising pitch, whose effect on dynamics overrode that of the coupling to timing. During the second half of each phrase, there was a tendency to stress metrically weak but harmonically dissonant upbeat more than the following downbeats, which parallels a tendency to lengthen these notes (Repp, 1995).

(4) *The change of successive hammer velocities during a crescendo or decrecendo may be a linear function of metrical time.* The present analysis differs considerably from that of Todd (1992) who computed average intensities of all notes in a beat and modeled them by the superposition of several underlying linear functions, according to the hierarchical phrase structure of the music. Here, expressive dynamics were examined at a finer level of detail, and therefore the results may not bear directly on Todd's model. The hypothesis considered here was that successive individual notes exhibit linear increases and decreases in MIDI velocity, because of an aesthetic preference for this manner of change. The evidence regarding this hypothesis is mixed, probably due to a multiplicity of factors that govern expressive dynamics.

(5) *Hammer velocity may increase with pitch.* Among the four voices, the soprano voice was clearly the most prominent. This was true even in passages where other voices had the principal melody (bars 8, 12, and 16). The lower voices were not much differentiated; only when one of them assumed melodic significance did it rise above the others. The prominence of the soprano voice was probably due to deliberate emphasis, not just to its high pitch: The other voices also differed in pitch register, but barely in average dynamic level. More unambiguous evidence for a relationship between intensity and pitch was obtained within each of the two right hand voices, suggesting that higher notes indeed tended to be played louder. The correlation was not very strong, however, suggesting that pitch is by no means the only influence on dynamics. The pitch-dynamics relationship may reflect an expressive convention (Friberg, 1991), but in part it may also be a compensation for a decrease in the perceived loudness of piano tones with increases in pitch.

Individual differences in expressive dynamics were not considered here in great detail. According to principal components analyses, there was only a single underlying pattern, so that the individual dynamic profiles can be regarded as variations around a common standard. A similar finding was

obtained in an earlier analysis of these pianists' timing profiles (Repp, 1995), which contrasted with the diverse timing patterns observed in a group of famous concert artists (Repp, 1992). Unfortunately, it is impossible to recover hammer velocities from acoustic recordings, so a future comparison of student and expert dynamics will have to be conducted on the basis of acoustic energy measures or will have to await the availability of a sufficient number of MIDI recordings by experts. It is not known at present whether famous artists also show greater diversity than students in dynamic patterns. Certainly, they may be expected to show more finely differentiated patterns and more precise control over dynamic contours and balances than young pianists, but these differences may be more quantitative than qualitative in nature. In other words, it seems unlikely that individual dynamic patterns will differ as radically as do some individual timing profiles.

Finally, it should be emphasized once more that the present analysis concerned only the level of MIDI velocities, which are a reflection of the forces applied by pianists' fingers to the keys. While pianists' intentions may be formulated at the level of action, they are also informed by auditory feedback and thus take into account acoustic and perceptual factors. The relationship between the hammer velocities of individual notes and the resulting sound structure is not simple, as the latter includes effects of instrument characteristics, sound transmission, and the interaction of simultaneous tones. Perception, in addition, introduces phenomena such as masking, fusion, and stream segregation. A multilevel study of expressive dynamics—including kinematics, acoustics, and perception—remains a project for the future.

ACKNOWLEDGMENTS

This research was supported by NIH grant MH-51230. I am grateful to Charles Nichols for assistance, to Jonathan Berger for permission to use the Yamaha Disclavier at Yale University, to the pianists for their participation, and to Janet Hander-Powers, Nigel Nettheim, and Neil Todd for helpful comments on the manuscript.

¹The tempo measure was obtained by taking the inverse beat duration. The intensity measure was inverse hammer flight time, as in Shaffer (1981). The correlation was computed over 128 beats, though data points were spaced two beats apart; the reason for this is not clear.

²Across different pitches, the relationship between velocity and sound level is greatly perturbed by factors such as soundboard resonance and room acoustics (see Repp, 1993a).

³This was perhaps to be expected, given that the soft pedal of an upright piano only moves the hammers closer to the strings. At the same time, the acoustic analysis confirmed the presence of large differences in peak rms sound level (here up to 13 dB) between tones of different pitch, as observed previously by Repp (1993a) on the same instrument as well as on a well-maintained Bösendorfer Imperial concert grand. This alarmingly large variation was not "corrected" for because it may depend on microphone position and because it is unlikely that the pianists adjusted their playing in response to it, given their limited experience with the instrument. The MIDI velocities are almost certainly a better measure of the pianists' intentions than is the radiated sound.

This may in part be due to the dynamic rate increases, timing reliability may decrease more rapidly than dynamic reliability. However, this remains to be investigated. The dynamic range may be taken to be four times the standard deviation. Since about four velocity units correspond to 1 dB, the standard deviations can thus be interpreted roughly as ranges in dB. The terms "note" and "tone" are used interchangeably here, following common usage. Strictly speaking, however, notes are graphic symbols that do not have intensities, and the intensities of the tones are inferred from the measured velocities.

- Banowetz, J. (1985). *The Pianist's Guide to Pedaling* (Indiana U.P. Bloomington).
- Clynes, M. (1987). Computerized system for imparting an expressive microstructure to succession of notes in a musical score (U.S. Patent No. 4,704,682).
- Friberg, A. (1991). "Generative rules for music performance: A formal description of a rule system," *Compt. Music J.* 15, 56-71.
- Gabrielsson, A. (1987). "Once again: The theme from Mozart's Piano Sonata in A major (K.331)," in *Action and Perception in Rhythm and Music*, edited by A. Gabrielsson (Royal Swedish Academy of Music Publication No. 55, Stockholm), pp. 81-103.
- Henderson, M. T. (1936). "Rhythmic organization in artistic piano performance," in *Objective Analysis of Musical Performance*, edited by C. E. Seashore (University of Iowa Studies in the Psychology of Music, Vol. IV, The University Press, Iowa City), pp. 281-305.
- Kendall, R. A., and Carterette, E. C. (1990). "The communication of musical expression," *Music Percept.* 8, 129-164.
- Lerdahl, F., and Jackendoff, R. (1983). *A Generative Theory of Tonal Music* (MIT Press, Cambridge, MA).
- Nakamura, T. (1987). "The communication of dynamics between musicians and listeners through musical performance," *Percept. Psychophys.* 41, 525-533.
- Palmer, C. (in press). "Anatomy of a performance: Sources of musical expression," *Music Percept.*
- Palmer, C., and Brown, J. C. (1991). "Investigations in the amplitude of sounded piano tones," *J. Acoust. Soc. Am.* 90, 60-66.
- Repp, B. H. (1992). "Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's 'Träumerei,'" *J. Acoust. Soc. Am.* 92, 2546-2568.
- Repp, B. H. (1993a). "Some empirical observations on sound level properties of recorded piano tones," *J. Acoust. Soc. Am.* 93, 1136-1144.
- Repp, B. H. (1993b). "Music as motion: A synopsis of Alexander Truslit's 'Gestaltung und Bewegung in der Musik,'" *Psychol. Music* 21, 48-72.
- Repp, B. H. (1994). "Relational invariance of expressive microstructure across global tempo changes in music performance: An exploratory study," *Psychol. Res.* 56, 269-284.
- Repp, B. H. (1995). "Expressive timing in Schumann's 'Träumerei': An analysis of performances by graduate student pianists," *J. Acoust. Soc. Am.* 98, 2413-2427.
- Repp, B. H. (submitted). "The art of inaccuracy: Why pianists' pitch errors are difficult to hear."
- Seashore, C. E. (1938/1967). *Psychology of Music* (McGraw-Hill/Dover, New York).
- Shaffer, L. H. (1981). "Performances of Chopin, Bach, and Bartók: Studies in motor programming," *Cog. Psychol.* 13, 326-376.
- Sundberg, J. (1988). "Computer synthesis of music performance," in *Generative Processes in Music*, edited by J. A. Sloboda (Clarendon, Oxford, UK), pp. 52-69.
- Thurmond, J. M. (1983). *Note Grouping: A Method for Achieving Expression and Style in Musical Performance* (JMT Publications, Camp Hill, PA).
- Todd, N. (1985). "A model of expressive timing in tonal music," *Music Percept.* 3, 33-58.
- Todd, N. (1989). "A computational model of tempo," *Contemp. Music Rev.* 4, 405-416.
- Todd, N. P. (1992). "The dynamics of dynamics: A model of musical expression," *J. Acoust. Soc. Am.* 91, 3540-3550.
- Truslit, A. (1938). *Gestaltung und Bewegung in der Musik* (Chr. Friedr. Vieweg, Berlin-Lichterfelde).