The Use of Rules for Expression in the Performance of Melodies

W. F. THOMPSON
Department of Psychology, Queen’s University, Kingston, Ontario, Canada, K7L 3N6

J. SUNDBERG and A. FRIBERG
Department of Speech Communication and Music Acoustics, KTH, Stockholm, Sweden

L. FRYDÉN
Edsberg Conservatory, Stockholm, and Department of Speech Communication and Music Acoustics, KTH, Stockholm

Starting from a text-to-speech conversion programme (Carlson and Granström, 1975), a note-to-tone conversion programme has been developed (Sundberg and Frydén, 1985). It works with a set of ordered rules affecting the performance of melodies written into the computer. Depending on the musical context, each of these rules manipulates various tone parameters, such as intensity level, fundamental frequency, and duration. In the present study the musical effect of nine rules is tested. Ten melodies were played under several rule-implementation conditions, and musically trained listeners rated the musical quality of each performance. The results support the assumption that the musical quality of performances is improved by applying rules.

Introduction

A great part of musical enjoyment is provided by the contributions of the performer. A performer does not merely provide a simple realisation of musical notation into sound, but rather, goes beyond what is indicated in the score, and attempts to convey feeling and expression appropriate to his or her musical interpretation of the piece. Even for a musically gifted student, it may take 15 years or more to acquire the ability to give sensitive, musical performances. What has the music teacher passed on to the student?

Part of the process of teaching music may involve giving technical advice, such as the appropriate fingering for difficult passages, various exercises, proper hand and arm positioning, and encouragement. But there are also other essential aspects of teaching. Through demonstration and verbal explanation, the teacher passes on numerous expressive devices. Often, performers have difficulty describing clearly the expressive devices that they employ. Indeed, precise descriptions of how best to perform a piece of music are deliberately avoided by some music teachers, so as to encourage the student to develop imagination and individuality.
Nonetheless, previous research has demonstrated that, beyond what is indicated by the score, general patterns of timing and loudness do exist in musical performances (e.g. Clarke, 1982; Gabrielson, 1974; Shaffer, 1981; Seashore, 1938; Sloboda, 1983; Todd, 1985). It is likely that these patterns of timing and loudness represent an aspect of music that is very important to the music listener.

The classification of such general patterns is a difficult task, however, since the use of an expressive device is usually dependent on a number of factors. As outlined by Clarke (1985), the use of an expressive device may depend on the expressive capabilities of the instrument, the overall style of the music, the speed and/or structural complexity of the particular passage, and the extent to which the performer wishes to emphasise musical structure. In Clynes' (e.g. 1987, 1983) view, expression is also dependent on the particular composer of the piece to be performed. Clynes has attempted to identify specific patterns of timing and loudness associated with various composers of Western tonal music. Such patterns, described as the composer's pulse, can be likened to a personal signature of the composer, and without the composer's characteristic pulse, it is thought that a performance will sound inauthentic and unconvincing.

In a review of the issues, Gabrielson (1985) has suggested that an exhaustive analysis of musical performance should involve first taking measurements on actual performances, then generating hypotheses based upon these measurements, and finally testing these hypotheses by means of psychological tests in which listeners evaluate systematically varied synthetic performances. While not in conflict with this general plan, the present research project is based on the belief that hypotheses based upon the musical intuition of expert musicians and music teachers may also be a rewarding topic for scientific research.

There are several reasons to support this belief. First, the method of analysis-by-measurement has certain limitations. For example, using measurements of actual performances alone, it is difficult to separate the effects of combined, though psychologically distinct, expressive actions. Similarly, analysis-by-measurement does not enable a clear separation between regularities that are implemented for the purpose of musical expression, and those that are merely a function of physical constraints. Second, following instructions from a professional musician in an analysis-by-synthesis approach offers certain advantages. Expert musicians are likely to focus on each of the psychologically distinct regularities that are specifically intended to convey musical style and expression. An expert musician may also provide musically meaningful hypotheses, and convey intuitive insights into the origins of expressive actions. Such information may ultimately be relevant to research concerned with the cognitive structure and encoding of such actions.

Finally, statistical data cannot readily account for the fact that alternative performances of the same piece of music can be equally musical. In contrast, the current approach allows for expressive uniqueness (German "Einmaligkeit"), since the expressive devices outlined by expert musicians may be understood as a set of possibilities from which a performer can make choices.
Outline of the Performance System

The present project was started in the late seventies, and other reviews of this work can be found elsewhere (Sundberg and Frydén, 1987, 1985; Sundberg, Frydén and Askenfelt, 1983; Friberg and Sundberg, 1986; Sundberg, 1988). The project concerns the clarification and classification of regularities associated with musical expression in performance, and our approach has been to appeal to the intuitions of a highly experienced music teacher to establish a rule system that enables a computer to generate musical performances. The development of the rule system can be likened to a typical music lesson situation, where a teacher tells the student how to improve his or her performance. In this case, a professional musician and teacher of music (co-author, LF) has acted as the computer’s music teacher. Thus, we have exploited the professional expertise of a musician to help a music student improve the musical quality of performances of melodies. There are only two alien elements in this “teaching” situation. One is that the student was a computer system rather than an understanding human. The other was that, in formulating each rule, the precise quantity with which the rule should affect the notes had to be specified; this quantity was tuned according to the reactions of the authors carefully listening to a number of different music excerpts.

The software used in this research was originally adapted from a text-to-speech conversion programme devised by Carlson and Granström (1975). In our case, the input is the musical notation, and the output is the control signals to a synthesiser. The performance system that has been developed is Le-Lisp and MIDI based, and allows users to define and apply various rules that alter the performance of a given piece of music (Friberg and Sundberg, 1986).

A musical score is input to the computer either from a MIDI keyboard, or from a file of note names, or by pointing at note symbols with a “mouse”. The resulting musical input is represented in a form similar to conventional notation. The internal organisation of the music is in voices, of which there may be many. Any of these voices can be selected to be affected by a set of rules, and then played. A set of rules may contain zero or any number of rules.

The performance rules act upon various levels of the musical structure of melodies: some rules consider only the duration or pitch height of individual tones, others consider pitch and temporal relations between adjacent tones, and still others rely on knowledge of the function of tones and chords within the prevailing key. Thus, in order to implement many of the rules, the computer must be provided with some basic information about musical structure, such as phrase boundaries, implied chord changes, and the overall key. Each of the numerous performance rules can be applied to a musical input, either separately or in combination, and the resultant performance can then be judged by musically trained listeners. In this way, the importance of the performance rules can be evaluated.

Other features of the performance system include menus and dialogue boxes for convenient usage, macrofunctions for inserting new rules, and
functions for ensuring that the average amplitude and duration of a given test melody are not altered by the application of a rule or set of rules.

**Predicting and Evaluating the Effect of the Rules**

The present investigation examined nine performance rules. These rules are a sub-set of the group of rules proposed in an earlier report (Sundberg, 1988). The evaluation was realised in two experiments: four rules were tested in Experiment 1 and five other rules were tested in Experiment 2.

Each of the rules is thought to have a musical effect when implemented in a performance. The experiments reported here tested this basic prediction. We did not, however, examine whether particular rules are compulsory in a musical performance. There may be numerous ways to perform musically, and, moreover, there may be no simple criteria for separating highly musical performances from merely adequate performances. Highly musical performances may belong to the class of *family resemblance* concepts, as described by Rosch and her associates (e.g. Rosch, 1975; Rosch and Mervis, 1975). Further investigation is needed to determine whether certain rules are compulsory or optimal, or whether the rules merely describe a set of expressive devices from which a performer is free to make choices.

Because each rule acts upon features of musical structure, the effect of an individual rule may be greater in some pieces than in others, and in some sections of a piece than in others. Therefore, employing a few very effective rules may result in a performance which is as musical as employing a great number of less effective rules.

**Experiment 1**

*Performance Rules Tested*

Four performance rules were tested in Experiment 1. General descriptions and tentative interpretations of these rules are given below. The rules described in Experiment 1 seem to generally serve the purpose of articulating and enhancing the structure of the music performed.

**Rule 1.** Marking phrase boundaries. Micropauses are inserted between phrases. Appropriate phrasing was decided upon by the authors. Micropauses were 80 msec in duration, and were implemented by starting the decay of the last tone in the phrase 80 msec earlier. As tones were normally overlapped by 60 msec, micropauses had the effect of eliminating the overlap and adding a silent interval of 20 msec.

The use of micropauses to mark phrase boundaries may be understood as a performer's illustration of grouping principles. The grouping of tones into phrase structures involves numerous cognitive factors, including Gestalt principles, and harmonic and rhythmic factors. This rule may serve to facilitate this basic perceptual process.

**Rule 2.** Marking leap boundaries. A micropause similar to the type just described is inserted between adjacent notes that form a melodic leap of greater than two semitones, provided the durations of these notes exceed 250 msec. The duration of these micropauses in msec was four times the
magnitude of the leap in semitones. A ceiling was set at nine semitones. As tones were normally overlapping by 60 msec (Fig. 1), the rule merely decreased this overlap time, thus creating a perceptual effect of a slight pause. The rule was not applied across phrase boundaries.

![Diagram of micropauses](image)

**Fig. 1**

Illustration of the strategy used for creating micropauses, for the case that first and second notes constitute a major second interval, and the second and third notes constitute a fifth interval.

This rule may also be understood as a performer’s illustration of grouping principles, but at a level of musical structure sub-ordinate to the level of phrase grouping (Sundberg and Lindblom, 1976; Lehrdahl and Jackendoff, 1983). The use of micropauses between tones separated by leaps illustrates the musical significance of sub-groups determined by the principle of pitch proximity.

**Rule 3.** The marking of *melodic charge.* Loudness and duration are increased in proportion to the “remarkableness” of the current note with respect to the prevailing chord. For instance, if the current chord is C major, the note C sharp is considered very remarkable, while the note G is considered relatively unremarkable. According to the rule system, the absolute value of the melodic charge of a note reflects its remarkableness.

Fig. 2 defines, by means of the circle of fifths, the melodic charge of the various notes for the case that the current chord is a C major triad. The melodic charge of a note is computed in relation to the root of the prevailing chord. A note’s sound level in dB is increased by a factor of 0.19 times the absolute value of its melodic charge.

As the circle of fifths is the basic model from which melodic charge is calculated, this rule may be understood as performers’ way of signalling their sensitivity to the relationship between the individual note and the prevailing chord.

**Rule 4.** The marking of *harmonic charge.* Loudness and duration are increased in proportion to the remarkableness of the current chord with respect to the prevailing key. However, changes in loudness and duration due to harmonic charge are achieved gradually, and thus involve crescendos and decrescendos.
The harmonic charge, $C_H$, is derived as a weighted sum of the melodic charges of the chord notes: $\text{Cm(i)}, \text{Cm(iii)},$ and $\text{Cm(v)}$.

$$C_H = \frac{(C_{M(1)})}{2} + \frac{(C_{M(3)})}{3} + \frac{(C_{M(5)})}{6} - 1.5$$

The harmonic charge of some chords in a C major tonality is shown in Fig. 3.

The sound level of the first note after each chord change is increased, by a quantity of $\Delta L$, in proportion to the square root of the new chord’s $C_H$:

$$\Delta L = 1.5 \times (C_H)^{1/2} \text{ (dB)}$$

The intermediate notes are then given intermediate sound levels so that crescendos and decrescendos are created. However, by starting crescendos no earlier than 1.9 sec before a chord change, crescendos which would be too gradual are avoided. Decrescendos, on the other hand, start immediately after the chord change.

The duration of each note in a crescendo or decrescendo is extended by a factor $C_{DR}$ proportional to the increase in sound level $\Delta L$:

$$C_{DR} = (1+0.018\Delta L)^{1/2}$$
Harmonic charge values of various chords. For a C major tonality the chords would be:

- T  C major
- D  G major
- S  F major
- SS B flat major
- DD D major
- DSR A major
- DTR E major
- SR d minor
- TR a minor
- DR e minor

Chord changes implied by each test melody were decided upon by the authors. The principles for creating crescendos and decrescendos are illustrated in Fig. 4.

As with the previous rule, the marking of harmonic charge may be understood as a performer’s way of signalling her or his sensitivity to key relationships. In this case, structure at the level of keys is reflected by expressive actions guided by the musical function of individual chords within keys.

Assessing the Musical Significance of the Rules

Each of the rules was evaluated by examining the musical effect of each rule in a melody that had a musical structure well suited to the function of the rule.
Four possibilities were assessed regarding the musical significance of the performance rules. These possibilities were:

1. The performance sounds more musical if one performance rule is applied than if no performance rule is applied.
2. The performance sounds more musical if four performance rules are applied than if just three performance rules are applied.
3. The performance sounds more musical if three or four rules are applied than if no or one rule is applied; and
4. The performance does not sound more musical if an arbitrary or counterintuitive rule is applied (i.e. an inappropriate transformation of the rule).

Support for the first two possibilities would strongly suggest that these rules are very important to a musical performance. Support for the third possibility, but not the first two possibilities, would represent modest and more general support for the importance of these performance rules. Support for the fourth possibility would indicate that listeners did not give high ratings to performances that were merely less tedious than the mechanical version. Rather, for expressive actions to be effective, they must make musical sense.

Method

Subjects. Twenty musically trained listeners participated in the experiment. Eighteen of these listeners were students at the Conservatory of the Swedish Radio, Edsberg, Sweden. The other listeners were at a comparable level of musical training. Listeners received 100 SEK for their participation in the experiment. Listeners were aware of the purposes of the experiment, but were not aware of the performance rules being tested.
Apparatus and Stimuli. Tones were produced by a DX21 Yamaha synthesiser, controlled by a Macintosh microcomputer via a MIDI interface. A timbre with a small peak at five times the fundamental frequency was created by combining three oscillators. This timbre, which had a dull oboe-like sound, was used for all test melodies. Tones had a rise time of approximately 10 msec, and a decay time of approximately 60 msec. Consecutive tones were overlapped such that the start of the decay of one tone coincided with the onset of the next tone. Sequences were played to the listeners through Sennheiser headphones. Listeners could adjust the overall loudness to a comfortable level at the start of the experiment.

Procedure and Conditions. Five different test melodies were used to examine the four performance rules. These melodies are listed in Table I. Each melody was chosen for its appropriateness in illustrating the importance of a particular rule. The marking of phrases, leaps, and harmonic charge were tested with one melody each. Two melodies were used to test the marking of melodic charge, since it was the opinion of the authors that this performance rule was particularly subtle.

Table I
Melodies used in Experiment 1

<table>
<thead>
<tr>
<th>Rule</th>
<th>Melody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Marking leaps</td>
<td>J. S. Bach, fugue theme from first Kyrie in B minor mass (BWV 232).</td>
</tr>
<tr>
<td>2. Marking phrase boundaries</td>
<td>F. Chopin, theme from Waltz, E flat major, op. 18.</td>
</tr>
<tr>
<td>3. Marking melodic charge</td>
<td>F. Chopin, theme from Mazurka No. 5 in B flat major, op. 7:1.</td>
</tr>
<tr>
<td></td>
<td>W. A. Mozart, second theme from first movement of the String quartet in d minor (K 421).</td>
</tr>
</tbody>
</table>

Each melody was played under five conditions:
1. In a strictly mechanical way (i.e. not applying any rule).
2. Applying only the rule being tested.
3. Applying the opposite (if applicable) or an inappropriate transformation of the rule being tested.
4. Applying three rules (all but the rule being tested).
5. Applying all four rules.

Presentations were blocked by melody. The five rule conditions of each test melody were presented in a random order within blocks. The order in which blocks were presented was also random. Prior to each presentation, listeners were provided with written instructions, in which they were told to rate the musical quality of the performance on a scale of 1–9. Listeners were
encouraged to compare the task to the teaching of musical performance to children. Listeners were told that there were no right or wrong answers, and that they should use the full range of the rating scale.

**Results and Discussion**

For each test melody, mean ratings for each rule condition are listed in the Appendix and displayed in Fig. 5. The data from all five melodies were analysed by an analysis of variance, and several planned comparisons were examined. Across the five melodies, when melodies were presented with just one rule applied, ratings of the performances were significantly higher than when no rules were applied, $F(1, 19) = 25.92, p < .001$. This rating difference occurred for all melodies with no significant interaction of the effect with the melody presented. The finding indicates that each of the performance rules had a musical effect when implemented as a single expressive device.

![Graph showing mean ratings for various rule conditions](image)

**Fig. 5**

Average ratings for the various performance rules tested in Experiment 1. For each rule tested, the five bars, from left to right, represent the following conditions: inappropriate rule; no rules; one rule; three rules; four rules.

Across the five melodies, when all four rules were applied, ratings of musical performances were significantly higher overall than when just three rules were applied in the presentations, $F(1, 19) = 8.51, p < .01$. Again, this rating difference occurred for all melodies, with no significant interaction of the effect with the melody presented. This finding demonstrates that each of the four rules are effective not only when they are implemented as a single expressive action, but also when they are implemented in addition to other performance rules.
Melodies presented with an inappropriate transformation of a rule were given lower ratings overall than melodies presented with the correct application of the rule and melodies presented with no rules applied, $F(1, 19) = 7.87$, $p < .05$. This finding indicates that the performance rules being tested did not act merely to make the performance less tedious than the mechanical version. For if that were the case, then arbitrary or inappropriate rules should have been as effective as appropriate rules. That the inappropriate rules yielded low ratings suggests that for a rule to be effective in performance, it must make musical sense.

There was one trend that was an exception to this latter finding. The inverse of melodic charge applied to one of the two pieces used to test this rule yielded quite high ratings. Although not statistically significant, applying the inverse of melodic charge in the Chopin Mazurka yielded higher average ratings than applying the correct version of the rule or applying no rules in this piece.

While surprising at first sight, it should be stressed that this inappropriate rule condition was the opposite of melodic charge. The high ratings are perhaps understandable if one assumes that under some conditions, listeners may accept a tradeoff between notated duration and the desired magnitude of the rule. In the Chopin Mazurka, but not in the Mozart Quartet, tones with a strong melodic charge were often notated with a much greater duration than were other tones in the melody. Thus, the melodic charge in the Mazurka was, in effect, already well marked by the composer. Applying the opposite of the melodic charge rule may have been heard as a musically appropriate way of making the already obvious melodic charge marks more subtle. However, when increases in melodic charge were not correlated with changes in duration, as in the Mozart Quartet, applying the opposite of melodic charge yielded low ratings of the musical performance, while applying melodic charge in the appropriate direction yielded high ratings of the musical performance.

In general, ratings were higher when three or four rules were applied than in other conditions, $F(1, 19) = 21.09$, $p < .001$. However, in three of the five melodies, the application of just one rule yielded ratings as high or higher than the application of the three other rules together. Critical to this finding is the fact that each melody was chosen because it was well suited for illustrating the importance of a particular performance rule. When one rule alone was applied in a performance, that rule was always a very effective expressive device for the piece being performed. Thus, the finding shows that applying one performance rule that is very effective can be as beneficial to the performance as applying three less effective rules. As previously suggested, such a finding is consistent with the notion that the effectiveness of a rule depends on the music being performed. Also, that ratings were similarly high for these two performance conditions (the one and three rule conditions) is consistent with the view that equally musical performances may be achieved in different ways.

The average data provide strong support for the validity of the rules examined in this experiment. It should be noted, however, that when data for melodies were analysed separately, only those concerning harmonic charge
yielded all four of the significant differences reported for the overall data (i.e. no rule v. one rule; three rules v. four rules; three or four rules v. other conditions; and an inappropriate transformation of the rule v. no or one rule). The analysis of individual data from other melodies typically yielded significance in one or two of these four comparisons. Obviously, the rules investigated are quite subtle in their individual effects. Nonetheless, the overall judgement data showed a systematic relationship between the listeners’ evaluation of the musical performances and the presence or absence of the rules.

**Experiment 2**

**Performance Rules Tested**

Experiment 2 was conducted to examine the musical importance of the following five rules:

**Rule 5.** The higher, the louder. The sound level of each tone is raised in proportion to its fundamental frequency, by an amount of 3 dB per octave.

This rule is thought to have a basis in the physical characteristics of many instruments, and most notably the human voice. An increase in voice fundamental frequency is associated with an increase in sound intensity both in normal speech (Fant, 1968) and in singing (Bloothooft, 1985). Thus, the rule may fulfil a general expectation which has developed through repeated exposure to this relationship between frequency and intensity.

**Rule 6.** Higher tempo in ascending melodic motion. The tempo is increased during passages in which ascending melodic motion occurs. This effect is achieved by shortening all members of the ascending figure by 5 msec. Ascending melodic motion was defined as two or more rising intervals in a row.

When ascending melodic passages are played without this rule, one easily gets the idea of laborious climbing, while the melody appears to point to the end of the uphill movement—when the rule is applied. In addition to this effect, the rule may act to complement Rule 5. Increasing the tempo of an uphill sequence also increases the rate of change in loudness imposed by Rule 5, and hence enhances its effect.

**Rule 7.** Higher notes in leaps are lengthened. When adjacent notes form an interval of greater than two semitones, the higher note is slightly lengthened. For an interval of 1 semitones width, the higher note in the leap was lengthened by $\Delta DR$:

$$\Delta DR = 5 \times I^{1/2} \text{ (msec)}, \ I > 2 \text{ (semitones)}.$$

When the leap was part of an alternating pattern of high and low notes, the rule was applied only to the first leap. Such alternating patterns are generally heard as two interleaved melodies, while this rule was intended for leaps that are part of one single melodic line. This rule may reflect the fact that in tonal music, scale tone steps are more frequent than leaps. In some sense, then, a melodic leap is more remarkable than a scale tone step, and might profit from some emphasis.

Adding duration where melodic leaps occur may also function to strengthen the perceived connection between the two notes. As demonstrated by van
Noorden (1975), although a wide pitch separation may weaken the perceived connection between two tones, the streaming of tones into two parts becomes less likely to occur as the rate of alternation between tones is reduced. Thus, where a melodic leap is meant to be part of the same grouping structure, lengthening the higher note may help to convey this grouping.

**Rule 8.** Short notes are made shorter. The duration of quarter, eighth, and sixteenth notes are reduced by three, eight, and two msec respectively.

This rule has the general effect of sharpening durational contrasts. Contrast enhancement is well known as a fundamental principle in the visual system (e.g. Hartline, 1949; Hartline and Ratliff, 1957; Ratliff, 1965), and is also implicated in the system that handles incoming pitch information (Carterette, Friedman and Lovell, 1969; Houtgast, 1972; Deutsch, 1973). The enhancing of durational contrasts in a musical performance is consistent with this standard perceptual principal.

**Rule 9.** Short notes are made softer. The sound level of a note is reduced by 80 dB/DR, where DR is the duration of the note in msec. As an example, the sound level of a sixteenth note, which is typically about 80 msec in duration, would be reduced by 1 dB.

The main purpose of making short notes softer was to avoid what we have termed the “machine gun” effect, which occurs if a string of short notes is played without care to reduce the overall loudness. Note sequences for which this rule was primarily intended are ornamental and usually characterised by good continuation. In addition, because the notes of such sequences are short, they are proximal in time to the other notes in the sequence. Considering these factors, there is a strong tendency toward perceptual grouping of the note sequence. However, if the individual tones within such a group are not softened, they tend to maintain their individual character, and the result is a conflict between hearing the sequence as a group, and hearing the sequence as a set of individual tones. Possibly, it is this conflict which is heard as the “machine gun” effect.

As with the rules tested in Experiment 1, the above rules were evaluated by examining the musical effect of each rule in a melody that had a musical structure well suited to demonstrate the effect of the rule. This experiment used the same design as Experiment 1, and allowed an assessment of the same four possibilities outlined earlier regarding the musical significance of the performance rules.

**Method**

**Subjects.** Twenty musically trained listeners participated in the experiment. Seventeen were students at the Conservatory of the Swedish Radio, Edsberg, Sweden. The three other listeners were at a comparable level of musical training. Many of the listeners had also participated in Experiment 1. Listeners from the Conservatory received 100 SEK for their participation in the experiment. The other listeners received an extra 100 SEK for travel time. Listeners were not informed of the performance rules being tested.

**Apparatus and Stimuli.** The apparatus and stimuli were identical to those used in Experiment 1.
Procedure and Conditions. Five melodies were used to test the five performance rules. These melodies are listed in Table II. Each melody was chosen for its appropriateness in illustrating the usefulness of one of the rules.

**Table II**
Melodies used in Experiment 2

<table>
<thead>
<tr>
<th>Rule</th>
<th>Melody</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The higher the louder</td>
<td>G. F. Händel: Sonata in E major for violin and continuo, op. 1:15.</td>
</tr>
<tr>
<td>3. Lengthening upper note in leap</td>
<td>F. Schubert: First phrase of Frühlingstraum, song nr 11 from Winterreise, op. 89 (D 911).</td>
</tr>
<tr>
<td>4. Shortening of short notes</td>
<td>F. Schubert: Theme from Trois Marches Militaires, op. 51:1 for piano, four hands (D 733).</td>
</tr>
</tbody>
</table>

The presentation conditions, instructions, and judgement task were the same as in Experiment 1. For all rules, condition 3 (implementing the inverse or inappropriate version of the rule) used the inverse of the rule being tested.

Results and Discussion

For each melody, mean ratings for each rule condition are listed in the Appendix and displayed in Fig. 6. Across melodies, there was an overall effect of Rule condition, \( F(4, 76) = 2.94, p < .05 \). Overall, implementing a single performance rule did not significantly influence ratings. However, the results do provide modest support for the effectiveness of the rules: ratings were significantly higher when four or five rules were implemented than when just one rule or no rule was implemented, \( F(1, 19) = 7.24, p < .05 \).

To a greater degree than in Experiment 1, ratings from Experiment 2 were dependent on the melody presented. This effect was indicated by an overall significant interaction between Melody and Rule condition, \( F(16, 304) = 2.32, p < .005 \). Analysing the data for each melody individually, it was found that only one rule appeared to significantly influence ratings on its own. An unplanned comparison in the analysis of ratings given to Handel’s sonata showed that ratings were higher when the rule making higher notes louder was implemented (one-rule and five-rule condition) than when that rule was not implemented (no-rule and four-rule condition), \( F(1, 19) = 6.27, p < .05 \).

Further analysis of data for individual melodies showed that no other rules yielded significant effects on their own. However, for three of the other melodies, ratings were higher when either four or five rules were implemented than when just one or no rule was implemented. This effect was statistically significant for ratings of the Bach Bourree, \( F(1, 19) = 14.89, p < .05 \).
Average ratings for the various performance rules tested in Experiment 2. For each rule tested, the five bars, from left to right, represent the following conditions: reversed rule; no rules; one rule; four rules; five rules.

$p < .005$, and showed a very strong trend in ratings of Schubert’s Marche Militaire, $F(1, 19) = 4.26, p < .051$. Data from the Roman theme showed a weaker trend in this direction, $F(1, 19) = 3.54, p = .07$.

It is, of course, possible that many of the latter effects were due to the presence of the rule making higher notes louder. This rule was indeed able to significantly influence ratings on its own, and it was present in the four and five rule conditions in melodies used to test other rules. However, this possibility was not supported by an informal analysis of the extent to which each rule affected the performances of the examples. First, except for ratings of Handel’s sonata, there was no obvious relationship between performance ratings and the amount of amplitude change introduced by making higher notes louder. Second, there did appear to be some relationship between the extent to which the other rules affected the performance of the examples and whether or not ratings were significantly higher in the four and five rule condition. Thus, it is likely that the higher ratings seen in the four and five rule conditions were due to the combination of rules applied, rather than to the presence of one particular rule.

The difficulty in verifying the importance of four out of the five rules examined in Experiment 2 was somewhat surprising, but nonetheless presents an important finding in the development of the rule system. There are several possible reasons for such weak effects.

One possibility is that the four performance rules yielding weak effects do not describe expressive actions in actual performances at all, and hence do
not relate well to the musical quality of performances. This extreme interpretation seems unlikely, however, since the rules were defined by a highly qualified professional music teacher and musician, and are in fact used in the teaching of music students.

A second possibility is that the expressive devices to which some of the rules refer are more context dependent than the rule descriptions suggest, and such rules need further refinement in order to precisely characterise their usefulness in musical performances.

A third possibility is that many of the rules have an important cumulative effect, but are too subtle on their own to influence ratings. Many of the findings in Experiment 2 were consistent with this latter view.

A final possible reason for the lack of clear effects relates to the judgement task itself. Because listeners were asked to rate the musical quality of performances, and were not informed of what aspects of expression they should focus on, they may have listened for general qualities of the performances. But since the differences between performances were often very specific, listeners may not have noticed the subtle effects of some individual rules. In our experience from tuning rules, knowing what to expect greatly facilitated detecting the effect of a rule. If listeners had been informed of the expressive effects that they should expect, it is possible that judgements of whether such effects were musically appropriate in a performance would have been easier.

**General Discussion**

The aim of this ongoing project has been to develop a formalised performance rule system that acts upon a musical score, given a general knowledge of musical structure. Numerous rules other than those tested in the present investigation have been proposed (e.g. Sundberg et al., 1983), and have yet to be fully tested.

The present approach may be a useful complement to research involving measurements of actual performances. For instance, it allows a separation of psychologically distinct expressive actions that combine to form complicated patterns of timing and loudness. Using analysis-by-synthesis in the study of musical performance avoids certain difficulties encountered by measurement approaches. For instance, although some research has attempted to measure loudness in performances (e.g. Seashore, 1938; Shaffer, 1981; Sloboda, 1983), precise measurements are difficult because loudness is a complex function of the acoustic properties of both instrument and environment (Shaffer, 1981). The analysis-by-measurement approach consequently has tended to focus on properties of timing (e.g. Clarke, 1982, 1985; Gabrielsson, 1974, 1985; Todd, 1985; Shaffer, 1981; Povel, 1977).

Finally, a more general consideration is that the current project provides a unique study of the question as to whether, in fact, an expert musician is able to make explicit the expressive devices that he or she uses in practice.

The findings point to several factors that appear to contribute to the musical quality of a performance. Of the nine rules examined in the present investigation, five significantly increased the musical quality of computer-realised performances when they were implemented.
Importantly, the research also reveals some of the difficulties that an expert musician may have in attempting to identify the individual factors involved in a musical performance. In particular, the weak effects found in Experiment 2 may suggest that some rules which are intuitively important to an expert musician may have a limited effect on their own.

Support for the significance of performance rules has come not only through the listening tasks reported here. As discussed in Sundberg and Frydén (1987), the basic concepts of melodic and harmonic charge clearly correlate with results of listening tasks carried out by Krumhansl and her associates (Krumhansl et al., 1982). Furthermore, Thompson and Cuddy (1987) have reported an asymmetry of key relationships depending on the direction of key change analogous to the asymmetry assumed in the rules for marking harmonic and melodic charge. Finally, it has been found that note durations computed using the rule system are similar to those observed in an actual performance: a particular cyclic pattern of bar durations, measured in an actual performance and reported by Todd (1985), was found to be reproduced by the application of our complete rule system (for details, see Sundberg et al., forthcoming).

Many of the rules proposed can be tentatively explained in terms of current knowledge of the cognitive organisation of musical structure. As we have argued, one important function of performance expression may be to articulate and enhance the structure in the music that is being performed. While many of the rules act upon surface structure, such as the duration or pitch of individual tones, the significant effects of marking melodic and harmonic charge indicate that other aspects of performance may depend on knowledge of high levels of musical structure.

It has yet to be determined whether certain expressive devices or rules are compulsory for a performance to sound highly musical. As previously discussed, there may not be simple criteria for distinguishing between merely adequate and highly musical performances. Highly musical performances may be characterised only to the extent that they all involve a well executed use of several expressive devices selected from a large set of possibilities. Again, analysis-by-synthesis may prove to be a useful tool for revealing the implications of “well executed”.

Certain fundamental questions and issues are raised by the present approach. For instance, we have argued that the rules may be understood as a set of expressive possibilities, from which a performer is free to make choices regarding the quantity by which a rule should affect the performance. The ability to account for expressive freedom is one important feature of the rule system. Future work may examine the extent to which expressive freedom itself contributes to the aesthetic value of music and musical performance. A second question concerns the musical knowledge that must be invoked in a good performance. In the present investigation, higher ratings were given to performances in which a greater number of performance rules had been applied. Since each rule acts upon some aspect of the musical structure, this finding suggests that better performances will require a greater amount of musical knowledge. However, it is possible that there is a limit after which further implementation of expressive actions causes a performance to sound too complex or pedantic, and hence relatively unmusical.
Considerations for Future Work

Some remarks should be made on the organisation of the rule system. When several rules influence a performance at the same time, the strength of each rule must be very small. Otherwise, the cumulative effect of several rules would be too much, and may even distort the perceived rhythmic structure of the music. In the present study, the strength of each rule tested was initially set according to how strong each rule could be if all other rules were being applied at the same time. Therefore, only a small quantity of each rule was implemented in this investigation.

Future research may examine the effect of rules implemented in greater quantities than were used in this investigation. Such research would be useful, since it is possible that the quantity with which an expressive device is applied varies considerably in actual performances. Treating the quantity of each rule as a free variable has certain implications for the performance system. It would suggest, for instance, that even when the musical structure is suitable for many kinds of expressive actions, certain expressive actions may be emphasised, others may be implemented in small quantities, and some may not be implemented at all.

We may find, for example, that a note is emphasised either by lengthening its duration, or by increasing its vibrato, or by slightly raising its pitch, or by inserting a micropause in front of it, or by playing it louder. In other words, there may be many synonyms for various aspects of musical expression, and a rule may simply represent one out of several possibilities. Incorporating this flexibility into the theoretical basis of the performance system would avoid the need to consider cumulative effects of the rules, and would allow individual rules to be implemented in greater quantities.

As the rule system is further developed and tested, it may become clear how expressive actions interact with each other, and under which precise conditions their implementation is most effective. Of course, while our research has focused on the performance of melodies, it will also be important to consider expression in music containing several voices. According to Shaffer (1981), pianists do have the ability to control hands independently, though the right hand sometimes has a more prominent role in building crescendos and decrescendos. However, more work is needed to establish the extent to which an expressive device affecting the melody will influence expression in other voices (Sundberg et al., forthcoming).

The performance of music presents a highly complex set of issues to be investigated. It is probable that a full understanding of this challenging area can only be reached through the convergence of data from several different approaches.

Acknowledgements

This investigation was completed while Dr. Thompson spent four months as a guest researcher at the Department of Speech Communication and Music Acoustics. This visit was supported by the Swedish Institute. The research project was supported by the National Bank of Sweden Tercentenary Foundation.
References


**Appendix**

**Mean ratings for each condition of each sequence (standard deviation given below each mean)**

<table>
<thead>
<tr>
<th></th>
<th>inverse</th>
<th>0 rules</th>
<th>1 rule</th>
<th>3 rules</th>
<th>4 rules</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozart</td>
<td>4·70</td>
<td>4·85</td>
<td>6·25</td>
<td>6·20</td>
<td>6·95</td>
</tr>
<tr>
<td></td>
<td>2·00</td>
<td>2·52</td>
<td>2·34</td>
<td>2·12</td>
<td>1·64</td>
</tr>
<tr>
<td>Chopin Mazurka</td>
<td>5·55</td>
<td>4·55</td>
<td>4·85</td>
<td>5·35</td>
<td>5·80</td>
</tr>
<tr>
<td></td>
<td>2·14</td>
<td>2·06</td>
<td>2·16</td>
<td>2·56</td>
<td>2·31</td>
</tr>
<tr>
<td>Bach Kyrie</td>
<td>3·95</td>
<td>4·05</td>
<td>4·60</td>
<td>5·75</td>
<td>6·40</td>
</tr>
<tr>
<td></td>
<td>1·85</td>
<td>2·26</td>
<td>1·85</td>
<td>2·12</td>
<td>2·06</td>
</tr>
<tr>
<td>Chopin Waltz</td>
<td>3·00</td>
<td>4·00</td>
<td>5·80</td>
<td>4·50</td>
<td>5·20</td>
</tr>
<tr>
<td></td>
<td>1·41</td>
<td>2·10</td>
<td>2·02</td>
<td>2·19</td>
<td>2·24</td>
</tr>
<tr>
<td>Schubert Symphony</td>
<td>1·95</td>
<td>2·95</td>
<td>4·85</td>
<td>4·75</td>
<td>5·75</td>
</tr>
<tr>
<td></td>
<td>1·28</td>
<td>1·47</td>
<td>2·70</td>
<td>2·25</td>
<td>1·83</td>
</tr>
<tr>
<td><strong>Experiment 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Handel Sonata</td>
<td>4·45</td>
<td>5·35</td>
<td>6·40</td>
<td>4·60</td>
<td>5·90</td>
</tr>
<tr>
<td></td>
<td>1·82</td>
<td>2·43</td>
<td>1·85</td>
<td>1·67</td>
<td>2·05</td>
</tr>
<tr>
<td>Schubert Marche</td>
<td>4·80</td>
<td>4·70</td>
<td>4·05</td>
<td>5·25</td>
<td>5·45</td>
</tr>
<tr>
<td></td>
<td>1·77</td>
<td>1·95</td>
<td>1·47</td>
<td>2·12</td>
<td>2·42</td>
</tr>
<tr>
<td>Bach Bourree</td>
<td>4·25</td>
<td>4·00</td>
<td>3·90</td>
<td>5·95</td>
<td>5·55</td>
</tr>
<tr>
<td></td>
<td>1·83</td>
<td>2·22</td>
<td>1·92</td>
<td>2·39</td>
<td>2·06</td>
</tr>
<tr>
<td>Roman Drottningholmsmusique</td>
<td>4·10</td>
<td>4·30</td>
<td>4·20</td>
<td>5·00</td>
<td>4·90</td>
</tr>
<tr>
<td></td>
<td>1·62</td>
<td>2·00</td>
<td>1·70</td>
<td>2·27</td>
<td>1·83</td>
</tr>
<tr>
<td>Schubert Fruehlingstraum</td>
<td>4·95</td>
<td>5·10</td>
<td>5·00</td>
<td>4·80</td>
<td>5·30</td>
</tr>
<tr>
<td></td>
<td>1·93</td>
<td>2·10</td>
<td>1·92</td>
<td>2·26</td>
<td>2·25</td>
</tr>
</tbody>
</table>