Music communication as studied by means of performance

Sundberg, J. and Frydén, L. and Friberg, A.

journal: STL-QPSR
volume: 32
number: 1
year: 1991
pages: 065-083

http://www.speech.kth.se/qpsr
MUSIC COMMUNICATION AS STUDIED BY MEANS OF PERFORMANCE

Johan Sundberg, Lars Frydén, & Anders Friberg

Abstract

This article presents an overview of a long-term research work with a rule system for the automatic performance of music. The performance rules produce deviations from the durations, sound levels, and pitches nominally specified in the music score. They can be classified according to their apparent musical function: to help the listener (1) in the differentiation of different pitch and duration categories and (2) in the grouping of the tones. Apart from this, some rules serve the purpose of organizing tuning and synchronization in ensemble performance. The rules reveal striking similarities between music performance and speech; for instance final lengthening occur in both and the acoustic code used for marking of emphasis are similar.

INTRODUCTION

In our computer age, it is neither difficult nor rare to have computers perform music. What distinguishes these performances from man-made performances is a lack of musical expressivity. Computer-generated performances sound musically pathological. Physical correlates of musical expressivity have been identified in terms of numerous minor and major deviations from the pitches, durations, and sound levels nominally assigned to the different tones by the score. These deviations are often referred to as expressive deviations.

The unmusical quality of computer-generated performances is interesting, as it seems to point to a very significant aspect of music communication. Ten years ago we started to investigate this area. The aim was to describe the principles according to which musicians distribute expressive deviations. Results from our research have been presented in detail elsewhere (Friberg, 1989; Sundberg, 1988; Sundberg, 1989; Sundberg, Friberg, & Frydén, 1989).

Once the expressive deviations have been described, the principles underlying them can be submitted to a scientific inquiry. Here, we will concentrate on two questions: what do the musicians add which is not written in the score, and why do they add this?

METHOD

We have used analysis-by-synthesis as our main method. The music score is written into a Macintosh microcomputer which controls a synthesizer through MIDI. Expressive deviations are generated by a set of ordered, context-dependent performance rules written in a program called RULLE. These rules explicitly and exhaustively describe all the expressive deviations that appear in the resulting computer-generated performances. In other words, we use rules for defining the principles which musicians apply when they add expressive deviations in music performances.

The rules originate from the professional competence of a string-quartet violinist and conservatory music teacher. After listening to computer generated performances, he suggested how the performance could be improved. His suggestions have been translated into performance rules. Thus, his professional musical competence, being a well-documented fact, is our prime research object.

The exact formulation of each rule is the result of careful listening to musical examples to which a rule has been applied in isolation or together with other rules. As the analysis-by-synthesis method used is comparative in its nature, it can only demonstrate to what extent the presence of a rule improved the performance, as compared with the performance generated without the same rule.

The rules have also been tested in formal listening experiments with expert listeners, and some rules have been assessed in threshold and preference tests. Also, comparisons between rule-generated and actual performances have been carried out in some cases. Next, some of the rules will
be described, and the result of the tests will be summarized. Finally, the function and origin of the rules will be discussed.

**EXAMPLES OF RULES**
Table I shows an overview of all rules as yet formulated. They affect duration, sound level, vibrato amplitude, frequency, and fine tuning. They are organized into a program which automatically generates a performance of the music score given as input. Note that there are many rules affecting the various parameters. Thus, according to our rule system, there are several reasons to shorten a note, for example.

**RULE SOUND PARAMETERS**
**NUMBER NAME CONCERNED**
**SEGMENTATION**
Duration categories
DDC 1A. Durational Contrast A DR
DDC 1B. Durational Contrast B L
DDC 2A. Accents L envelope
DDC 2B. Double Duration DR
Pitch categories
DPC 1A. High Sharp F
DPC 1B. High Loud L
DPC 2A. Melodic Charge DR, L, VA
DPC 2B. Melodic Intonation F
**GROUPING**
Micro level
GMI 1A. Leap Articulation L envelope
GMI 1A'. Leap Articulation (alt.) L envelope
GMI 1B. Leap Tone Duration DRO
GMI 1C. Faster Uphill DR
GMI 2. Amplitude Smoothing DR
GMI 3. Inegalles L envelope
GMI 4. Articulation in Repetition L envelope
GMI 4'. Articulation in Repetition (alt.) DRO
Macro level
GMA 1. Phrase DR, DRO
GMA 2A. Harmonic Charge DR, L, VF
GMA 2B. Chromatic Charge* DR, L
GMA 3. Final Retard DR
**ENSEMBLE PERFORMANCE**
ENS 1. Mixed Intonation F envelope
ENS 2. Melodic Synchronization DR
ENS 3. Bar Synchronization DR

Table I. Overview of the performance rules. The rules affect duration (DR), sound level (L), fine tuning (F), vibrato amplitude (VA), and vibrato frequency (VF). Some rules insert a micropause at the end of the note in terms of the timing characteristic DRO for the decay. Some rules have two alternative (alt.) formulations. In the present article, the rules for ensemble performance are not discussed.

A rule has two important aspects with regard to function: its target notes, i.e., the notes which it modifies, and the degree or quantity of the effect produced by the rule. In the program, the quantity of a rule can be conveniently varied by a multiplication factor which is given the default value of 1.0 in the program. This value gives a very small effect, appropriate for the case that all rules are applied at the same time.
The rules seem to serve two main purposes in music communication. In music, pitch and duration are categorically perceived (Burns & Ward 1978; Clarke 1987). One purpose seems to be to help the listener to identify pitch and duration categories. Thus, some rules increase the differences between pitch categories, and some increase contrasts between duration categories. We call these rules differentiation rules, because they assist the listener in differentiating the elements in the acoustic signal flow. Another group of rules helps the listener to identify tone groups, i.e., musical gestalts of various lengths. We call these rules grouping rules.

In this section, some examples of rules will be presented in the order reflected in Table I, giving the rule numbers in parentheses. For a full technical specification, the reader is referred elsewhere (Friberg, 1989).

Rule Durational Contrast A (DDC 1A) shortens shorter notes and lengthens longer notes. It thus increases the durational contrast between short and long notes.

Rule Double Duration (DDC 2B) lengthens the shorter note and shortens the preceding longer note in 2:1::1 durational context. As will be explained later, it increases the difference between certain durational configurations. Fig. 1 shows how it affects a simple melody in 3/4 time.

![Dashed curve: duration deviations from nominal produced by rule Double Duration (DDC 2b) affecting tone durations in a ...2:1::1... context. The example is a folk tune "Sorgerliga saker hända" (Sad things happen). Solid curve: measurements from a sung performance of the same song according to Bengtsson & Gabrielson (1983).](image)

Rule High Sharp (DPC 1A) increases the fundamental frequency of notes: the higher the note, the sharper it is played.

Rule Melodic Charge (DPC 2A) attempts to reflect quantitatively the fact that, given the harmonic and melodic context, some notes are more remarkable than others. The melodic charge of a scale tone is defined using the circle of fifths with the root of the prevailing chord as the reference, as shown in Fig. 2. The rule adds emphasis to the notes by increasing amplitude, duration, and vibrato depth depending on their melodic charge, as illustrated in Fig. 3a.

Rule Melodic Intonation (DPC 2B) adjusts the fine tuning of the scale tones depending on their melodic charge, see Fig. 3b.

Rule Leap Articulation (GMI 1A) inserts a micropause between two tones forming a leap. The micropause is generated by letting the final part of the note's amplitude decrease quickly before the
onset of the following note arrives. The duration of the pause, measured as the duration of the decay (DRO), see Fig. 4, depends on the size of the leap up to 9 st and remains constant for wider leaps. The figure also shows how the rule affects a music example.

Rule Faster Uphill (GMI 1C) moves tones in ascending melodic figures closer in time by decreasing their duration by 2 ms shorter, if preceded by a lower and followed by a higher note.

Rule Amplitude Smoothing (GMI 2) eliminates steps in amplitude between adjacent tones. It does not apply across phrase and subphrase boundaries.

Rule Inegalles (GMI 3) lengthens the stressed note by 22% of its duration and shortens the following unstressed note by the same number of ms. It is applied in sequences of equally long notes.

Rule Phrase Marking (GMA 1) is a macro-grouping rule operating on phrase and subphrase boundary signs which are added to the input notation. Fig. 5 shows an example. It inserts an 80 ms micropause to the final note of subphrases and lengthens phrase final notes by 40 ms.

Rule Harmonic Charge (GMA 2A) represents an attempt to reflect in a quantitative form the fact that not only scale tones, but also chords, form a hierarchical system of remarkableness in traditional tonal music. It equals a weighted sum of the chord notes' melodic charges computed with the root of the tonic as the reference (see Fig. 6). The macro-grouping rule harmonic charge creates crescendos when a chord of higher harmonic charge is approaching and decrescendos in the opposite case, and these level variations are accompanied by proportional tempo and vibrato variations, as illustrated in Fig. 7.

Rule Chromatic Charge (GMA 2B) can be regarded as an equivalent of Harmonic Charge applicable to atonal music. After transposition of all tones of a voice into the same octave, the chromatic charge is computed as the average of the inverted pitch distance in semitones of five adjacent tones. Thus, it increases when adjacent tones approach each other in pitch, and decreases when they are widely scattered. The effect, therefore, is that crescendos are generated when adjacent tones are close in pitch and diminuendos are produced in the opposite case.
DIFFERENTIATION OF PITCH CATEGORIES:

MARKING MELODIC CHARGE BY
SOUND LEVEL, DURATION, AND VIBRATO

Fig. 3a. Example of the effects of rule Melodic Charge (DPC 2A) on the theme of the first b minor KYRIE in the b minor mass by J.S. Bach (BWV 232). The chords are given in the top line in terms of the interval in st between the root of the chord and the root of the tonic, minus sign denoting minor chords.

FINE TUNING

Fig. 3b. Deviations from the equally tempered tuning (E. T.) produced by rule Melodic Intonation (DPC 2B) which tunes the scale tones according to their melodic charge.
Fig. 4. Illustration of rule Leap Articulation (GMI 1A) which inserts a micropause depending on the pitch separation between adjacent tones. The ordinate shows the DRO parameter, defined in the insert figure, and the ordinate is the serial position in the melody.

Fig. 5. Illustration of the effect of rule Phrase (GMA 1) which introduces subphrase (SP) markers in terms of an 80 ms micropause at the end of subphrases, and phrase (P) markers by duration increases (DR) of 40 ms for phrase final notes; another 40 ms duration increase is made for the last note in the melody.

VERIFICATION

Listening Panels
In an explicit, quantitative form, the rules reflect the professional expertise of our consultant musician. As such, they represent psychologically interesting information.

However, the rules must also possess some degree of generality, because, obviously, a successful professional musician adheres, so to speak by definition, to certain common principles of music performance. Therefore, it is interesting to show to what extent these rules have a more general applicability. Different approaches have been tried for testing this.
Harmonic charge for some chord functions: \(T\)=tonic, \(D\)=dominant, \(S\)=subdominant, \(DD\)=double dominant, \(SS\)=double subdominant, \(R\)=relative. For example in a C major tonality the chords would be \(SS=Bb\) major, \(S=F\) major, \(T=C\) major \(D=G\) major, \(DD=D\) major, \(DSR=A\) major, \(DTR=E\) major, \(SR=d\) minor, \(TR=a\) minor, \(DR=e\) minor.

Fig. 7. Example of the effect of rule Harmonic Charge (GMA 2A) which marks changes in harmonic charge by means of long-term changes in sound level, tempo, and vibrato. The chords are given in terms of the interval in st between the root of the chord and the root of the tonic.

One approach has been to ask professional musicians or top-level music students to rate the musical quality of different rule-generated performances or to choose the "best" performance in a pair of performances of the same excerpt, one with, the other without, a particular rule applied. For a detailed account of these experiments, including statistical data, the reader is referred to Friberg, Frydén, Bodin, & Sundberg (1987a); Friberg, Sundberg, & Frydén (1987b; 1990); Frydén, Sundberg, & Askenfelt (1989); Kronman & Sundberg (1987); Sundberg & al. (1988;1989); Thompson, Sundberg, & Frydén, (1989).

Table II shows a schematized overview of results from all listening experiments carried out so far with different music excerpts. As can be seen in the table, the results have been by and large encouraging, except for rules Durational Contrast (DDC 1 A and B).
Table II. Summary of results from listening tests. Expert listeners listened to melodies where a rule had been applied either alone (ONE) or in combination with other rules (MANY). The subjects' task was either to give a preference in a comparison for a deadpan or a rule generated performance of the same melody (PREFERENCE), or to adjust the quantity of a rule (ADJUSTMENT). Each column refers to one test; - and + refers to negative and positive test results, and 0 means that the rule was not tested. The rules marked ATONAL were tested in combination with other rules for atonal keyboard music.*

Preferred Quantity
The magnitudes of the effects generated by the rules represent an interesting aspect for several reasons. The program produced very small effects because of the low default values used for the multiplication factors which produced the effect of the particular rule, as mentioned. Thus, using the default value of 1.0 for rule quantity many of the rules added only a few ms to the duration of a note, while the differential threshold is much higher, about 10 ms or higher (van Noorden, 1975). However, the effects of different rules often accumulate on certain notes.

We do not believe that the default quantities are universally applicable. A piece of music can be performed in many different ways which all are musically acceptable. Much of this variability can probably be accounted for in terms of different rule quantities. For example, a performance considered exciting at the beginning of this century may appear quite exaggerated and tasteless to listeners in our time.

In the listening tests described above, the small default rule quantities were used; often in the comparison test, the subjects apparently failed to focus their attention on the effect generated by the rule. In two other listening tests, the quantity was varied by manipulating the default value for a rule. A change in this default value changed the quantity of the effect produced. For example, a

* As these listening tests have been carried out over a long period of time, during which work has been continuously spent on the rules, some rules have been slightly modified between tests. Details are given elsewhere (Frydén & al., 1989, Thompson & al., 1989, Friberg & al, 1990, and Sundberg & al, 1988, and Sundberg forthcoming).
doubling of the default value for a rule lengthening certain notes produced a performance in which this lengthening was twice as great. One of these tests aimed at determining subjects' preferred quantity, another at determining the threshold quantity. Detailed accounts of these experiments are given elsewhere (Friberg & al., 1990; Sundberg & al., 1988).

The preferred quantity in a music example was determined for each of six rules (see Table II) in a production test. Six professional musicians were asked to adjust, by means of a slide ruler on the computer screen, the quantity of each rule so that the best performance of a certain one-voice music excerpt was obtained. The available variation range was wide, including zero and, if possible, also negative quantities.

Fig. 8. Average and 95% confidence interval bars for preferred quantities of the rules tested in the production experiment: Durational Contrast A and B combined (DDC 1A), Hi h Sharp (DPC 1A), Melodic Charge (DPC 2A), Leap Tone Duration (GMI 1B), Phrase (GMA 1), and Harmonic Charge (GMA 2A). Professional musicians' task was to adjust the rule quantity to their own satisfaction. Quantity 1 is the default value in the performance program.

THRESHOLD TEST

SERIES 1

RULE 1: The higher, the higher

Fig. 9. Principle used for stimulus presentation in the threshold experiment. Each rule was tested by presenting a series of different performances of the same music example. There was one series for each rule, and within each series the different performance versions were arranged in pairs. In each pair of versions, the first version represented a quantity which varied between 0 and a very great maximum value (Max) between the pairs, while the second version always presented a deadpan standard with Q=0 (std). The figure shows the first three pairs of a series.

Fig. 8 shows averages and 95% confidence intervals for the preferred quantities. It can be seen that all these rules except one (Durational Contrast being a combination of rules DDC 1A and B in this test) obtained a preferred mean quantity above zero. This result shows that, with this exception, the effects of all rules were considered musically desirable by these musicians.

The generality of these findings is hard to judge. It is possible that the preferred quantity varies to some extent depending on both the excerpt chosen and on the musical taste of the listener.
Threshold Quantity
How large must an effect be in order to be perceptible? The threshold quantities of seven rules were estimated in an experiment with two subject groups: 10 top-level music students and 12 musically untrained. Subjects were asked to decide if pairwise-presented, rule-generated performances of the same one-voice excerpt were identical or different. Performance differences within pairs ranged from huge to nil.

Figure 9 illustrates the principle used for stimulus presentation. The pairs were presented in series, one series for each rule, and there was one excerpt for each rule. In each pair, the second version was always the deadpan standard, and the task was to judge, for each pair, if the pair consisted of the same version played twice or two different versions. To direct the listener's attention to the aspect which the rule affected, pair number 1 in each series presented a huge difference. Each series also contained one pair representing the case of zero difference, i.e., consisting of two deadpan versions, and two replicated pairs.

For the seven cases of zero difference, the musicians gave an average of 63% "Same" answers only, while the corresponding value for the nonmusicians was 80%. This means that false alarms were about 17% more common in the musicians' group, probably demonstrating an eagerness among the music students to show an excellent musical ear.

Figure 10 shows the results in terms of the percentage of "Same" answers as function of the magnitude of performance differences greater than zero. It can be seen that this percentage is great for small performance differences and tends to approach zero, as the difference becomes greater. The nonmusicians' curves are generally well above those of the musicians, generally much more than the 17% difference in bias toward false alarms. This indicates that the musicians could detect smaller performance differences than the nonmusicians.

In the figure, preferred quantity has been plotted after normalization with regard to the default quantities, so that a quantity value of 1.0 always corresponds to the default quantity. It can be noted that for many of the rules, the main shift in the response curves towards few "Same" answers occurred close to this quantity of 1.0. Thus, for these rules, the default value produced effects which were perceptible for most musicians.

This result shows that the threshold quantity for effects generated by these rules differs between musicians and nonmusicians. This seems to implicitly suggest that musicians are familiar with these effects: the effects belonged to their domain of expertise. If so, sensitizing people to such effects would belong to music education, and sensitivity to these effects would belong to and, thus, reveal musicality.

In the same figure, the preferred quantities as established in the production test described previously are also shown (arrows). For most rules, the preferred quantity corresponded to a quantity which was so great that about 80% of the subjects could hear the performance difference. This relation between the preference and the threshold tests is encouraging as it suggests that those quantities were preferred which produced clearly audible effects for most musicians.

Predicted and Observed Performance Data
Duration
The rules have also been tested by comparisons between actual and predicted performance, as produced by the performance program using the default quantities. Fig. 2 showed such a comparison for a folk-tune lullaby demonstrating a high degree of similarity. Fig. 11 shows more examples in terms of two performances of Mozart's A major sonata for piano, as measured by Gabrielson (1987) and by Palmer (1989). With regard to the long-term pattern encompassing a phrase, Player A shows a pattern not predicted by our program. With regard to the short-term pattern embracing a few adjacent notes, there is a good qualitative agreement between observation and prediction for both players. This agreement is largely due to rule Double Duration (DDC 2B).

This and other comparisons between actual and rule-generated performances show that there is often a clear similarity. Nevertheless, our rules rarely produce all the effects observed.
Fig. 10. Percentage of "Same" answers received from musicians (filled circles) and nonmusicians (open circles) in the listening test as function of the physical difference between the two performances compared. The difference is given in terms of the quantity parameter $Q$, normalized with respect to the defaults value used in the program. Solid and dashed curves show results for musicians and nonmusicians, respectively. The arrows show the preferred quantities according to the preference test.
Fig. 11. Deviations from nominal note values performances of W.A. Mozart's Piano Sonata in A major (K 331). Observed data (solid curve) were taken from Gabrielsson (1987) (upper graph) and from Palmer (1989) (lower graph). Predicted data (dashed curve) were obtained from the performance program using the default quantity parameter values.

Fine tuning

According to rule-melodic intonation (DPC 2B), the scale tones on the dominant side of the circle of fifths are tuned sharp while the others are tuned flat, the extent being determined by the melodic charge. Thus, this rule predicts small deviations from the equally tempered tuning depending on the melodic charge of the note.

Fig. 12 compares such predictions with a set of measurements published by Garbuzov (1948) from a David Oistrach, a Mischa Elman, and an Efrain Zimbalist performance of the solo part of the Air from J.S. Bach's C major Suite for orchestra. The measurements of Oistrach's performance were taken from a gramophone record which we could retrieve; we checked Garbuzov's data on this record and found results differing from Garbuzov's by less than 10 cent in most cases.

The data in the figure were taken from the three players' two versions of the first six bars of the piece. In the figure, dots and bars show averages and standard deviations. The mean deviations from the equally tempered tuning is about twice as large as the predicted deviations. A more detailed analysis of these data is deferred to future writings. Suffice now to state that there is a significant correlation ($r=0.572$) indicating that the predictions show a qualitative agreement with the fine tuning of these three violinists.
Fig. 12. Comparison between observed and predicted deviations from equally tempered (E.T.) tuning in both reprises of the first six bars in three piano accompanied solo violin performances of the Air from J.S. Bach's Suite for orchestra in C major (BWV 347) as played by violinists D. Oistrach, M. Elman and E. Zimbalista according to measurements published by N. Garbuzov (1948). The filled circles show the mean, the bars +/- one standard deviation, the numbers under the bars show the number of observations, and the line is the best linear regression fit.

Atonal Music

A preference test was carried out with atonal keyboard music (Friberg et al., 1987a). A total of six performance rules, including the chromatic charge rule, were applicable; these rules are marked in Table II. The subjects were asked to make preference judgments between pairwise presented performances where one was always deadpan, and the other was generated by applying all rules. The examples were realized on a sampler synthesizer set to piano timbre for piano music excerpts by Xennakis, Boulez, Webern and a random algorithm, and to a timbre similar to electric guitar for four one-voice examples composed by random algorithms. The subjects were highly specialized in contemporary music.

The results are shown in Fig. 13. The rule-generated performances were clearly preferred to deadpan performances for all examples. This result supports the idea that several of our performance rules apply also to contemporary atonal music, even though they were formulated and tuned to fit traditional tonal music.

Independent Evidence

Support for the notions of melodic and harmonic charge has emerged from independent experiments by Krumhansl and collaborators (see Krumhansl, 1987; Krumhansl & Kessler, 1982; Krumhansl, Bharucha, & Kessler, 1982). In these experiments, subjects judged how well a probe tone or probe chord served as a continuation of a preceding standard scale and standard cadence, respectively. Fig. 14 shows the relation between this probe tone and these probe chord ratings and melodic and harmonic charge, respectively.

There is a significant correlation in both cases. As one might expect, an excellent continuation (high rating) corresponds to a low melodic and harmonic charge, and vice versa. These results
suggest that, while playing, musicians take into account listeners' ratings of how well the note or the chord served as a continuation. Incidentally, this marking of melodic charge seems essential. If the melodic charge is not reflected in the performance when a melodically highly charged note appears, it sounds as if the player is playing the wrong note by mistake.

**Fig. 13.** Number of preference judgments from a panel of five expert listeners on contemporary atonal music comparing pairwise presented deadpan (hatched) and performance rule generated (brickwall pattern) performances of atonal keyboard music. The music was composed by the indicated composers or by random algorithms and performed on a computer controlled sampler producing a piano sound (a) or an electric guitar like sound (b). (For more details see Friberg & al., 1987a.)

**Fig. 14.** Relationship between melodic and harmonic charge on the one hand, and probe tone (a) and probe chord (b) ratings according to experiments by Krumhansl & Kessler (1982) and Krumhansl, Bharucha & Kessler (1982), in which subjects rated how well a probe tone or chord served as a continuation of a scale and a cadence, respectively.
Fig. 15. Illustration of the effect of rule Double Duration (DDC 2B). The solid time markers to the left show the nominal durational realization of the note patterns to the right. In the middle duration context the rule moves some time from the half note to the subsequent quarter note, as shown by the dashed time marker; this moves the corresponding durational pattern to an orientation more midway between the durational patterns represented in the top and bottom lines.

Fig. 16. Filled circles show the mean slowing of the tempo in final retards according to Sundberg & Verrillo (1980); the bars show +/- one standard deviation. The solid line indicates the inverse of the time intervals between footsteps during an idealized stopping of running according to Kronman & Sundberg (1987).

DISCUSSION

We are convinced that our rule system is far from being complete. For instance, in the near future we hope to complement our performance rule system by a set of micro-grouping rules which insert micropauses at the end of small note groups. Also, some rules missing in our system have already been identified in analyses of music performances. One example is the effect that a melody voice may be leading by some 30 ms over the accompaniment tones in chords, as observed by Rash (1979) in ensemble performances and by Palmer (1989) in piano performances. Another effect observed by Clarke (1988) is that the arrival of an emphasized note may be delayed by a short
pause just before the note. Thus, our rule system is not complete.

According to Todd (1985), phrases are marked by a special pattern characterized by an accelerando at the beginning and decelerando toward the end of a phrase. In our program, the only phrase marker is a lengthening of the final note. This suggests that different codes, or "synonyms", may be available for announcing the beginning and termination of a phrase.

It was mentioned before that the rules can be divided according to the specific purposes which they apparently serve in the music-communication process: differentiation and grouping. These purposes deserve some further comments.

The differentiation rules enhancing differences between pitch and duration categories seem to use two different principles. One is to simply increase the physical separation between the categories. Examples are rules Durational Contrast A (DDC 1A) and High Sharp (DPC 1A) which shorten short notes or sharpen high notes. In a less obvious way, rule Double Duration (DDC 2B) offers another example. By transporting some duration from the long note to the short note in this particular durational contexts, the half-note – quarter-note – half-note pattern becomes less similar to the otherwise rather close pattern of dotted half-note – quarter-note – dotted half-note and more halfway to the half-note – half-note – half-note pattern, as illustrated in Fig. 15.

Another principle used for enhancing the contrasts between pitch and duration categories seems to be to add, to an extent dependent on the difference in the primary property pitch or duration category, a secondary property. Rule High Loud (DPC 1B) is a good example. It adds a difference in a secondary property (sound level) to an extent depending on the difference in the primary property (pitch). Thus, high pitches are not only tuned sharp but also slightly larger in amplitude than low notes. Another, less obvious example is rule DPC 2 (melodic charge). Melodic charge differs between the scale tones. Rule Melodic Charge (DPC 2A) apparently adds emphasis to an extent reflecting the melodic charge. Emphasis corresponds to several secondary properties: sound level, duration, vibrato depth, and fine tuning.

As regards grouping, our rules work at both a lower micro level and a higher macro level. At the micro level, the context consists of pitch intervals. Examples are all the GMI rules, inserting micropauses in leaps, moving tones belonging together closer in time, such as Faster Uphill and Inegalles (GMI 1C and 3), or smoothing the amplitude differences between them, such as Amplitude Smoothing (GMI 2).

The two principles discussed which seem to help the listener to differentiate pitch and duration categories and to group elements belonging together are wellknown in speech. Speech sounds are also categorically perceived and the difference between speech sound categories is often enhanced (Carlson, Friberg, Frydén, Granström, & Sundberg, 1989). For example, short vowels are not only shorter than long vowels, they also have different formant frequencies in many languages. Other instances of this idea can be found outside the area of acoustic communication. Architects often choose one color for all windows, another color for the walls, and a third for the roof. And females and males are certainly prone to add to the anatomical differences between the sexes by means of clothes, hair style, etc. Categorization seems essential in communication.

By marking remarkableness, quantified in our program in terms of melodic and harmonic charge, the musician takes the hierarchically organized system of scale tones and chords into account. The means, thereby, used seem to be emphasis; melodically charged notes sound emphasized in our computer generated performances.

Emphasis would be related to predictability. The purpose may be to alert the listener so that communication may be maintained even at turns in the signal flow which the listener is not likely to predict. The same need for emphasizing unexpected elements can be observed also in speech. Predictability seems to be another important component in communication.

The performance rules define in a tangible way how musicians modify the sound signals nominally described by the music score. Therefore, the rules offer an opportunity to examine the acoustic code which the musician chooses to help the listener with the grouping task. Acquaintance with this code must be relevant to musicality; if listeners do not understand this code, much of the charm in listening to music may evade them.
The origin of the code is an interesting issue. Is the code used only in music, or is it more widely used? In the former case, music is an exclusive form of art, demanding special knowledge, while in the latter case, music would be available to almost anyone.

We just saw that the phenomenon of differentiation occurs both in music performance and in speech. It seems that, in many cases, there are striking parallels also with regard to the acoustic code used in music performance and speech (Carlson & al., 1989). Actually, the code is sometimes even identical. For instance, as pointed out by Todd (1985), final lengthening is used for marking the ending of sentences and smaller structural units also in speech. Further, emphasis is marked by increased syllable duration. A tempting conclusion would be that speech is the origin of the code. However, the code used in speech varies to some extent between languages while musicianship seems largely independent of the player's mother tongue. This suggests that the origin of the code used in music performance originates somewhere else.

One possible origin is locomotion. The average retard pattern revealed by measurements on 24 recordings of misc from the Baroque or Classical eras is shown in Fig. 16. It was found to be nearly identical with the curve showing the decrease of foot-step rate during stopping of locomotion in running, provided that neither step length, nor braking force was changed during the stopping process, i.e., that the stopping was perfectly preplanned (Kronman & Sundberg, 1987). This suggests that listeners understand the meaning of a retard because it is iconic, it alludes to the listener's experience of stopping locomotion in running.

There are more examples of iconic rules. Rule Faster Uphill (GMI 1C) moves tones belonging together closer in time; also, it can be viewed as a negation of the final lengthening used for marking the termination of tone groups. Thus, lengthening notes means "end" and shortening notes means "continuation". Another example of an iconic rule is offered by rule Harmonic Charge (GMA 2A); it produces crescendos leading up to a new, harmonically more charged chord or a decrescendo to a harmonically less charged chord. The underlying purpose would be to keep all the notes together. This may be accomplished by letting all those satellite notes participate in one and the same long-term event, namely, the increase or decrease of sound level.

The acoustic code is not always iconic. Rule Amplitude Smoothing (GMI 2) eliminates amplitude differences between notes belonging together. Discontinuity is a common method of announcing group boundaries, perhaps alluding to our auditory experience; things that sound differently often originate from different sound sources.

The physiology of the perceptual system may be another factor of relevance to the choice of acoustic code. We may speculate that the code used for emphasis offers more time for the processing of an unexpected event. The grouping rules pack notes belonging together closer in time and lengthen the final element. In this way, the perceiving system is allowed some more time to process a note group, once it has been completed.

CONCLUSIONS

In the introduction we raised two questions: What do musicians add to the music described in the score, and why do they add this? Our analysis-by-synthesis of music performance allows us to propose hypothetical answers to these questions.

The musicians add and subtract duration, sound level of the notes, and also vary fine tuning and vibrato characteristics, if possible. The purpose of the resulting expressive deviations seems to be to help the listener to process the flow of acoustic signals. By means of the performance, the musician helps the listener to correctly categorize the sound events, to correctly identify structural boundaries, and to raise the listener's attention when remarkable events occur. Thus, the performance seems relevant to the listener's parsing work.

Our analysis also shows striking similarities between music performance and speech. Also speakers use prosody to assist the listeners to correctly categorize the speech signals, to group them, and to add emphasis on remarkable events. There are striking similarities also regarding the acoustic means used for conveying this information; shortening and lengthening seem to be used
with the same meaning, and the means to signal emphasis are largely identical.

The performance rules revealed by our investigation seem to constitute a lexicon which allows the musician to translate into acoustic signals his interpretation of the music. There would be synonyms in this lexicon, i.e., different acoustic codes may be used for conveying the same information. Also, there must a number of rules in music performance which are not yet represented in our rule system. A further exploration of this lexicon seems a rewarding task for future research.

ACKNOWLEDGMENTS
The author is indebted to Dr. Jaan Ross, Tallinn, Estonia, for showing him the cited data published by Garbuzov. This text is an expanded revision of the author's contribution to the First International Conference on Music Perception and Cognition, Kyoto, October 1989. The research program was supported by The Bank of Sweden Tercentenary Foundation.

REFERENCES


