What tells you the player is musical?

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

journal: STL-QPSR
volume: 23
number: 4
year: 1982
pages: 135-148

http://www.speech.kth.se/qpsr
C. WHAT TELLS YOU THE PLAYER IS MUSICAL?

A study of music performance by means of analysis-by-synthesis

J. Sundberg, L. Fryden*, and A. Askenfelt

Abstract

When computers are used for converting strings of note signs into the corresponding played melodies, the result is mostly disastrous from a musical point of view. As there is apparently a consensus among musically experienced subjects about this, one may postulate the existence of rules for musical performance. The present article describes an attempt to formulate such rules for the case of playing one part melodies. The procedure used is analysis-by-synthesis. The musical notation is fed into the computer, which converts it into a sequence of tones using a modified version of the Carlson & Granström program for text-to-speech conversion. The sounds are generated by a hardware synthesizer. Some rules operate with a small time window and are triggered by certain constellations of pitch or duration relationships between two or three adjacent notes. Other rules are triggered by the changes in harmony or by the phrase boundaries, which are marked in the input musical notation by means of specific symbols. An attempt is made to speculate about the psychological purpose of the rules.

Introduction

In musical performance, a string of note signs is generally converted into a sequence of tones. If this conversion is made in a simple-minded way, e.g., by applying a one-to-one relationship between the notation and the sound sequence by means of a computer, the result is musically a disaster. This fact raises a question: what are the ideal or acceptable relationships in music practice between the note signs and their corresponding acoustic signals? The present paper is an attempt to contribute to the answer to that question by formulating rules for the conversion of note signs into tone sequences.

Method

Two strategies are available in an investigation of this kind: measurements and/or synthesis-by-rule. Previous studies of musical performance have revealed an almost overwhelming complexity (see, e.g., Seashore, 1930; Bengtsson & Gabrielsson, 1975), as expected. This suggests that the rules relating notation to sounding music can be assumed to be numerous and interacting. This implies the need for a large amount of measurements before any conclusions can be drawn regarding the underlying rules. Moreover, the pedagogical and artistic experience of one of the authors (LF) had generated a number of hypotheses regarding such rules. For these reasons the synthesis-by-rule strategy was preferred.

* Conservatory of the Swedish Radio, Edsberg, Stockholm.
As in a previous study of musical performance focusing on singing (Sundberg, 1978), a computer-controlled vowel synthesizer was used ("MUSSE", see Larsson, 1977). The synthesizer can generate one part only. In the present experiment the signal characteristics were adjusted so as to be similar to the timbre of a wind instrument and also a very small vibrato was used. The fundamental frequencies were in accordance with the equally tempered scale, but could be modified in steps of 7 cents, which is close to one difference limen for frequency in normal subjects (Rakowski, 1971). The amplitude was controlled in steps of 1/4 dB. The duration was controlled in steps of a time unit corresponding to .8 to 1.2 csec depending on the tempo, which is accurate enough according to measurements by van Noorden (1975). The computer programs used for controlling the synthesizer were (1) a notation program (Askenfelt, 1979) by means of which the melody can be written in ordinary music notation on the computer terminal display; and (2) a text-to-speech program written by Carlson and Granström (1975). The information encoded in the notation is translated into "vowel sounds", which possess, among other characteristics, duration, fundamental frequency, and amplitude, which can be manipulated by the rules. Apart from the information normally included in a music score, special signs were used to mark (a) changes in harmony and (b) boundaries between phrases and sub- phases.

Rules

THE HIGHER THE LOUDER

The amplitude is increased as function of fundamental frequency. The amount is 4 dB/octave rise in the fundamental frequency, see Fig. 1.

SHARPENING DURATIONAL CONTRASTS

Short note values are shortened according to the values shown in Table I. No compensation is made for the resulting perturbation of the mechanical meter.

Table I. Shortening of the durations for different note values.

<table>
<thead>
<tr>
<th>Note value</th>
<th>Shortening csec</th>
<th>Shortening %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sixteenth</td>
<td>-0.5</td>
<td>-3.1</td>
</tr>
<tr>
<td>Eighth</td>
<td>-1.0</td>
<td>-3.1</td>
</tr>
<tr>
<td>Quarter</td>
<td>-1.0</td>
<td>-1.6</td>
</tr>
<tr>
<td>Half</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Whole</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
SHORTENING LOWER NOTE IN RISING INTERVAL

The duration of a note initiating a rising interval is decreased by 1 csec, see Fig. 2. In sequences of rising intervals the rule has the effect of raising the tempo somewhat. As before, no compensation is made for the resulting perturbations of the meter.

LENGTHENING TARGET NOTE IN LEAPS

The duration of tones terminating a single melodic leap of a given direction is shortened. The amount is $0.6 \times 1^{1/2}$ (csek), where $I$ is the size of the interval in semitones, see Fig. 3.

PAUSES IN LEAPS

The amplitude of the final portion of the tone is decreased at a constant rate. The start of this amplitude decrease occurs at $0.5 \times I$ csec from the end of the note, and the final amplitude value is $3 \times I$ dB below the initial amplitude, where $I$ is the interval size in semitones. The effect of this rule is negligible for narrow intervals but is quite noticeable for wide intervals, see Fig. 4.

ACCENTS MARKING CERTAIN DURATIONAL CONTRASTS

Accents are distributed to certain notes which constitute a contrast in duration between adjacent notes. The accent corresponds to an amplitude event with the following characteristics. The amplitude is increased by $6/DR$ dB, where $DR$ is the duration of the note in csec; this amplitude value is reached after $1 \times DR + 3$ csec and is kept for 7 csecs. Thereafter the amplitude falls off towards the initial value, which is reached at the end of the note, see Fig. 5. The rule adds such an accent to two types of notes. One is a short note surrounded by longer notes. The other case is a note terminating a specific pattern of changes in duration: a decrease followed by an increase. This rule has a clear effect particularly on the short notes after a dotted note.

AMPLITUDE SMOOTHING

Large stepwise changes of amplitude between adjacent tones are eliminated by adding to or subtracting from the last amplitude value of each tone a constant corresponding to 80% (in log. terms) of the amplitude difference with the following note, see Fig. 6.
Fig. 1. Schematical illustration of the rule increasing the amplitude of a note as a function of its pitch.

\[ \Delta d r = -1 \quad 0 \quad 0 \quad -1 \quad -1 \quad 0 \quad (\text{csec}) \]

Fig. 2. Illustration of the effect of the rule which shortens the durations of notes initiating a rising melodic leap by 1 csec.

\[ \Delta d r = 0.6 \cdot 1^{1/2} \]

\[ \Delta d r = k \cdot 7^{1/2} \quad 0 \quad k \cdot 5^{1/2} \quad 0 \quad k \cdot 4^{1/2} \quad 0 \]

Fig. 3. Illustration of the effect of the rule which lengthens the duration of a note terminating a single melodic leap in a given direction.
Fig. 4. Schematic illustration of the rule which inserts amplitude decreases towards the end of a note. The decrease has a constant rate. The onset time of the decrease is determined by the interval between the note and the following note.

\[ t = 0.5 \cdot I \text{ (csek)} \]
\[ a = 3 \cdot I \text{ (dB)} \]

Fig. 5. Schematic illustration of the accents marking certain contrasts in note value.

\[ \Delta_1 = 40 / DR_2 \text{ (dB)} \]
\[ \Delta_2 = 60 / DR_3 \text{ (dB)} \]
Fig. 6. Schematic illustration of the rule decreasing the amplitude contrasts between adjacent notes. The last amplitude reading of a note is corrected by 80% of the amplitude difference with the following note.
AMPLITUDE MARKING HARMONIC DISTANCE

This rule generates crescendos and decrescendos. The technicalities are as follows. In the input notation the chord changes are marked and each chord is specified. The rule assigns to each chord a value corresponding to its harmonic distance (see Fig. 7). The harmonic distance of a chord equals the number of chord changes required in traditional harmony in order to reach the tonic again (see Sundberg & Lindblom, 1976). By multiplying the square root of this harmonic distance by 16, an amplitude value is obtained, which is added to the first tone appearing over the chord considered. The same procedure is repeated for each chord. Finally, the amplitudes of the remaining tones are linearly interpolated on a dB-scale. Thereby, too slow crescendos are avoided by letting the amplitude remain constant until 1.6 sec before the chord change. The end result is crescendos culminating on harmonic changes introducing a chord which is more remote from the tonic than the preceding chord and decrescendos ending on chords which are less remote from the tonic than the preceding chord. The upper part of Fig. 8 demonstrates how the rule operates on a typical cadence.

MARKING TONIC DISTANCE

The distance along the circle of fifths from the root of the tonic is determined for each note. This distance is multiplied by 1.5 for scale tones located on the subdominant side of the circle. We have called the resulting values tonic distances, see Fig. 9. By multiplying this tonic distance by .5, this tonic distance is converted into an amplitude value which is added to each note. For example, in C major tonality the pitch of F sharp receives a rather large amplitude increase, while the pitch of G receives a negligible increase. The rule is illustrated in the lower half of Fig 8.

LENGTHENING FIRST NOTE AFTER CHORD CHANGE

The first note appearing over a new harmony is lengthened by the number of csec which corresponds to the harmonic distance of the underlying chord.

FINAL LENGTHENING

In the input notation phrases and subphrases are marked by specific signs. The rule adds 4 csec to the note terminating a phrase. For notes terminating subphrases the last 1 csec is used for a pause.
Fig. 7. Definition of harmonic distance by means of the circle of fifths. For a given chord this distance equals the number of chord changes which is required in traditional harmony in order to reach the tonic chord again.
Fig. 6. Schematic illustration of the various steps which generate crescendos and diminuendos. First, the chords are identified, and the chord changes are marked. Second, the harmonic distance is determined for each note in accordance with Fig. 7. Last, the harmonic distance is converted into an amplitude increment which is applied to the first note appearing over a new harmony. Finally, the amplitude of the remaining notes are linearly interpolated on a logarithmic scale.
Fig. 9. Definition of tonic distance of the various scale tones. For tones lying on the dominant side of the tonic on the circle of fifths, the tonic distance is the number of fifths separating the note from the root of the tonic. For tones on the subdominant side of the tonic, the tonic distance is the same number multiplied by 1.5.
Discussion

Of course, more performance rules than those presented above exist, even if we stay within the realm of one-part music. Also, we do not pretend that the present formulation of the rules is definite nor that all of them possess generality. What we do pretend is that the rules improve the musical acceptability of the performance of the melodies on which they have been tested.

We have not tried to model the multitude of choices which is available to the living musician and which allows him to play the same piece in many different ways, all of which are equally acceptable from a musical point of view. On the other hand, there would be possibilities to include such a liberty in the future. For instance, different interpretations of where the boundaries are between phrases and subphrases will result in different performances. It also may be that the quantity of the effect of the rules may vary from one performance to the other. We also would like to declare that we do not believe that our performance rules must always be obeyed in a good performance. On the contrary, such a performance may be boring in the long run. We believe that musicians should violate one or more of the performance rules as soon as they want to communicate something in particular or excite the audience by means of a surprise.

In our work with formulating the rules we have had the typical experience that the quantity of amplitude or duration is highly critical, by which a rule is affecting a specific note. Even when the effect is only slightly exaggerated, it becomes over-explicit and embarrassing. The musically useful quantity appears to be one, which is just noticeable but not identifiable. Perhaps this is something which is essential for art in general: we do not want to be disturbed by information about the technical means producing the piece of art, we just want to experience the piece of art as such.

The rules presented above have an ad-hoc character. However, several of them seem related to psychoacoustic or other types of already existing evidence. The principle of lengthening the note which terminates a leap might reflect certain effects studied by van Noorden (1975); the disruption of a melodic line, which is often the result of a wide melodic leap, can be eliminated by slowing the tempo. A similar effect might be obtained if the tone terminating the leap is lengthened.

Some rules appear to have other types of origin in that they seem to serve various purposes associated with the psychological process
involved in musical communication. One group of rules seem to apply to notes which in some respects are more or less surprising to the listener. It is as if the player wanted to point out to the listener that he/she makes no mistake in playing this unexpected note. Examples of such "no-mistake"-rules may be the lengthening of the target note in leaps, the amplitude increases on notes with high tonic distance, the crescendos towards harmonically remote harmonies, the accents on short notes surrounded by long notes, and the lengthening of the first tone appearing after a change in harmony.

The rule that generates harmony-dependent crescendos and decrescendos produces a crescendos ending on chords with the function of a dominant and decrescendos ending on the subsequent "rest chords". As such sequences of dominant chords - rest chords are used - among other things to mark the termination of phrases and subphrases in simple tunes, this rule tends to have the effect of marking the melody structure (cf., Sundberg & Lindblom, 1976).

A second group of rules appears to serve the purpose of bringing the synthesis in reasonably close agreement with the sound of music as played on a traditional instrument (including the human voice). The rule increasing the amplitude with pitch is a good example, as this is a characteristic shared by most traditional instruments. It is also possible that the rule inserting pauses in wide leaps refers to a general characteristic of many instruments; for instance, the fingers have to be moved a large distance on the fingerboard of the stringed instruments, when wide leaps are being performed and this would give rise to small pauses.

A third group of rules appears to take into consideration different kinds of associations, which a listener is likely to experience by listening to the music. Even though not included in the present study, the final retard represents a classical example of this (see Sundberg & Verrillo, 1980). Apparently, it would be perfectly possible to end a piece of motor music without a retard. However, if the music reminds the listener of physical motion, a retard is needed, because we know from experience with, e.g., running that physical motion is preferably slowed down before it stops. It is possible that the accelerando in sequences of rising intervals is another example; if pitch increases tend to cause associations with upward physical motion, and vice versa, then a tempo increase in "uphill" motion might inform the listener that the motion is performed without strain.
Every music listener would be biased from his/her acquaintance with speech. In other words, in trying to decode the sequence of acoustic signals, a music listener is likely to make use of his/her skill in doing the same thing in a speech listening situation. Both speech and music represent systematized interhuman communication by means of acoustic signals. This fact suggests that the codes might be partly similar. In speech the final lengthening seems to be a principle of great generality (Lindblom, 1978). This means that most listeners are inclined to interpret lengthening as a sign of termination. This might be the background of the rule which lengthens the last tone in phrases.

It is interesting to note that many rules appear to have an extramusical background. This suggests that music communication requires both intra- and extramusical experience from the part of the listener. We believe that further research on this background will be interesting and rewarding.

In summary, then, it seems possible to imagine plausible backgrounds and purposes behind several of our rules, in spite of the ad-hoc-origin of these rules.

Conclusions

From the above we conclude the following:

(1) It is possible to improve the musical acceptability of a performance by applying a limited set of "pronunciation" rules.
(2) Such rules can be discovered by means of an analysis-by-synthesis approach.
(3) Such an approach enables us to formulate new hypotheses as to how the rules proposed should be complemented and thus to contribute to knowledge about and scientific understanding of music.
(4) Some of the rules seem to reveal basic requirements for musical communication.

Acknowledgments

This is a status report of a project in progress; after the above text was written the formulation of some rules has been improved, particularly regarding the crescendo/diminuendo rule.
References


