Importance of Note-Level Control in Automatic Music Performance

Roberto Bresin
Department of Speech, Music and Hearing
Royal Institute of Technology - KTH, Stockholm
email: Roberto.Bresin@speech.kth.se

Abstract
A summary is presented of recent studies on articulation in piano music performance together with the applications they originated. Emphasis is given to legato and staccato articulation. Results from measurements of performances recorded with MIDIfied grand pianos are summarized. Some rules for the simulation of articulation are presented and their application in renderings of emotional expressive performances is discussed. These rules can produce effects that reflect tempo indications in the score as well as the expressive intentions of the player. Also discussed is how the articulation rules can be applied to the control of sound synthesis algorithms and why music performance research is important to a wider range of applications connected to perception and human behavior.

1 Introduction
Since the beginning of the computer era, researchers have tried to replicate human behavior with machines. Humanoids that can walk (Pandy and Anderson, 2000) and dance (Lim, Ishii and Takanishi, 1999) have been designed as well as those that can play the piano, talk, listen and answer (Kato, 1987). Still, these machines lack the ability to understand and process the emotional states of real humans and to develop and synthesize an emotional state and personality of their own. To overcome this limitation research on music performance seems particularly promising, since music is a universal communication medium, at least within a given cultural context. Research on music performance represents a promising starting point for understanding human behavior.

What scientific research on music performance mainly seeks is the understanding of underlying principles accounting for variations from score indications as applied by musicians in their performances. In recent years, research on piano music performance has devoted an increased attention to the study of duration variation at note-level such as in articulation (Repp, 1995; Repp, 1997; Woody, 1997; Repp, 1998). In legato and staccato articulation, these deviations are responsible of the acoustical overlap or gap between adjacent tones. It has been found that players vary their articulation strategies when rendering different expressive intentions (Bresin and Battel, 2000; Bresin and Widmer, 2000); staccato articulation will be more pronounced in an allegro tempo than in an adagio tempo. Recently it has been demonstrated that articulation is an important cue in the emotional coloring of performances, and therefore it is important to discriminate between different levels of legato and staccato articulation. Staccato articulation, in the performance of a score that did not include staccato marks, helps in communicating happiness to listeners. A performance with exaggerated legato articulation can be perceived as sad (Gabrielsson and Juslin, 1996). The importance of articulation is also demonstrated by the care dedicated to the technique of articulating notes in the practicing of musical instruments. Investigations in this direction have been demonstrated to be useful in a wide range of applications including automatic music performance and models for the control of sound synthesis algorithms. The problem of articulating sounds produced with synthesis algorithms is well known: these algorithms are usually very good in synthesizing one isolated tone but insufficient to interconnect sounds in an acoustical and musical realistic way (Mathews, 1975).

In the following paragraphs, an overview of recent research results related to articulation in piano music performance is presented. Some applications, and future directions are also discussed.

2 Legato and Staccato Articulation
Before proceeding to the presentation of recent research results on articulation, explanation of some terms is needed.

To measure the degree of legato articulation Repp (1995) introduced the definition Key Overlap Time (KOT) for adjacent tones. It is defined as the time interval between the key depression for the following tone and the key release for the current tone (see Figure 1). Bresin introduced later the term Key Overlap Ratio (KOR), defined as the ration between KOT and inter-onset-interval (IOI) for the n-th note. In analogy with KOT and KDT for legato articulation, Bresin introduced the terms Key Detached Time (KDT) and Key Detached Ratio (KDR) for staccato articulation (see Figure 1) (Bresin and Battel, 2000).
In the following paragraphs the main results from two studies on expressive articulation are presented. In both experiments players performed on computer-controlled pianos.

In a first experiment differences in the amount of legato and staccato articulation as performed by five pianists were analyzed (Bresin and Battel, 2000). The pianists, two female and three male, were Italian students in their diploma and pre-diploma year of piano classes at the Conservatory of Music “Benedetto Marcello” in Venice. They performed nine times the first sixteen bars of the Andante movement of W A Mozart’s Piano Sonata in G major, K 545. Each time the pianists adopted a different expressive intention: natural, brilliant, dark, heavy, light, hard, soft, passionate, and flat, natural representing the players’ preferred rendering in the absence of a specific emotional coloring, i.e. musically natural. The pianists performed on a Yamaha Disklavier II grand piano, C3 model. The articulation performed by the right hand was analyzed only for those notes that were marked staccato and legato in the original score by Mozart. Interesting results emerged from the analysis of the recorded data. Even if there were some large differences between the five pianists’ performances, they generally used similar strategies in their different renderings. In particular, legato was played with a KOT that depended on the IOI of the first of two overlapping tones: longer notes were performed with relatively shorter KOT than shorter notes. Performances intended as flat had the lowest average KOR while performances communicating heaviness were performed with the average largest KOR (see Figure 2). The natural performance (white bar in Figure 2) was played by the five pianists with an average KOR of about 15%, which gave it the middle position in a rank order according to KOR. Staccato tones had an overall average duration of about 40% of the IOI. Staccato ranged from mezzo-staccato in the heavy version, with a KOT of about 70% of the IOI, to staccatissimo in the brilliant and light versions, with a KOT of about 25% of the IOI (see Figure 3). The staccato in natural performances was produced with average duration about 66% of the IOI. Repeated notes articulation was also analyzed. The first note in a pair of repeated ones was performed with average duration of about 60 % of the IOI, i.e. in the range of mezzo-staccato articulation.

A second study focused on the analysis of staccato articulation (Bresin and Widmer, 2000). In this experiment a professional Viennese pianist played 13 Mozart’s piano sonatas on a Bösendorfer SE290 computer-monitored concert grand piano. The performance of notes that were marked staccato in Mozart’s original score was analyzed. The large-scale data analysis of 4082 notes revealed that the amount of staccato varied with melody contour, tempo
indications, and context. Notes were played with larger staccato in allegro tempi than in adagio tempi. The amount of staccato for the middle note in a three-notes context varied if the preceding and/or following notes were also marked staccato. For instance, notes were played more staccato if they were followed by non-staccato tones (NSN and SSN cases in Figure 4).

Even if the two studies were based on different materials, similar results were observed. It emerged that in staccato articulation the relative amount of staccato for one tone is independent of its IOI, as observed also by Repp (1998). Another important result was that articulation depends also on the melodic direction. In legato articulation, notes initiating ascending intervals were played with shorter duration and KOT than notes initiating descending intervals. In staccato articulation, notes initiating ascending intervals were played more staccato, i.e. with shorter duration, than notes initiating descending intervals. This dependence of articulation on the melodic shape is in accordance with previous findings. The <Faster uphill> rule proposed by Lars Frydén and implemented in the Director Musices (DM) program, shortens the duration of notes in ascending melodies (Friberg, 1991; Friberg, Colombo, Frydén and Sundberg, 2000). Furthermore, the piano action is faster for keys corresponding to higher notes (Askenfelt and Jansson, 1990; Goebl and Bresin, 2001).

3 Articulation Rules for Automatic Piano Performance

The results from the two studies presented above were implemented in the DM music performance grammar in terms of a new set of articulation rules for piano music performance. They operate on notes marked legato or staccato in the score, and on repeated notes. These rules are named respectively <Score legato articulation> rule, <Score staccato articulation> rule, and <Articulation of repetition> rule. A fourth rule, <Duration contrast articulation>, controls the articulation of notes that are not covered by the three previous cases, i.e. this rule can be used to add a legato or staccato articulation to any other note in the score.

All these rules affect articulation in different ways according to expressive indications (such as brilliant or dark), tempo indications (such as adagio, andante, allegro, presto, and menuetto), and legato and staccato marks in the score. For a detailed description of these articulation rules see (Bresin, 2001). As an example, the equation for the <Score legato articulation> rule is (see also Figure 5):

\[ KOT(k) = -g(k) \cdot IOI^2 + f(k) \cdot IOI \]  

where the variable k is an emphasis parameter used to control legato articulation; the two functions g(k) and f(k) are plotted in Figure 5; IOI is the inter-onset-interval referred to the current note; KOT(k) is the resulting key-overlap-tone for the current note. Equation (1) can produce legato articulation effects for performances ranging from flat to passionate, and passing through natural.
The changes of tone duration produced by the articulation rules are equal to or larger than the just noticeable quantity necessary for perceiving legato or staccato tones according to findings by Woody (1997). Furthermore analysis-by-synthesis experiments confirmed the importance of these rules for a qualitative improvement of computer-generated performances.

In the next sections two applications of the articulation rules are presented: (1) production of emotionally expressive performances, and (2) new models for the control of sound synthesis algorithms.

4 Articulation in Emotional Expressive Music Performance

In a previous section, it was shown how articulation plays an important role in expressive performance. Gabrielsson and Juslin (1996) observed that articulation is relevant to the emotional coloring of a performance: when asked to portray sadness, solemnity or tenderness players use legato articulation, while staccato or non-legato articulation is applied in happy, scared and angry renderings. This was observed already by Carl Philippe Emanuel Bach (1753/1949) who wrote "...activity is expressed in general with staccato in Allegro and tenderness with portato and legato in Adagio...".

The articulation rules presented in the previous section were included in the design of DM macro-rules for emotional coloring of performances. A macro-rule is a collection of performance rules that are applied to a score in order to obtain a complete expressive performance. A macro-rule for sadness is presented in Table 1. The effect of this macro-rule is illustrated in the version of Carl Michael Bellman’s song Letter 48 presented in Figure 6. The corresponding sound excerpt is available on the RENCON 2002 Proceedings CD-ROM and on the Internet (see the Links section for the URL).

The articulation rules were successfully applied in macro-rules for the production of emotionally expressive performances of music scores. These performances were classified in formal listening tests as happy or sad, thus confirming hypotheses by Juslin and collaborators (Juslin, Friberg and Bresin, In press).

The use of articulation rules for the rendering of different expressive performances has a correspondence with hyper- and hypo-articulation in speech. Formants, intensity and duration of vowels and duration of consonants can vary with the speaker’s emotional state or the intended emotional communication (Lindblom, 1990). Still, as in expressive music performance, the structure of phrases and the meaning of the speech remain unchanged.

More information about emotionally colored performance can be found in a paper by Bresin and Friberg (2000).

Figure 6. Inter-onset-interval (IOI in %), offset-to-onset duration (DRO in %) and sound level (dB) deviations in the sad version of Carl Michael Bellman’s song Letter 48. A positive DRO corresponds to KOT; a negative one to KDT.
Table 1. DM macro-rule description for the sad performance of Carl Michael Bellman’s song Letter 48. In the first column are the important expressive cues identified by Gabrielsson and Juslin; in the second column are reported qualitative observations by them; in third column is described the DM rules setting for implementing the observations in column two.

<table>
<thead>
<tr>
<th>EXPRESSIVE CUE</th>
<th>GABRIELSSON &amp; JUSLIN (1996)</th>
<th>DIRECTOR MUSICES RULES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tempo</td>
<td>Slow</td>
<td>Tone Duration is shortened by 15%</td>
</tr>
<tr>
<td>Sound level</td>
<td>Moderate or Low</td>
<td>Sound Level is increased by 8 dB</td>
</tr>
<tr>
<td>Articulation</td>
<td>Legato</td>
<td>Score Legato Articulation rule (k = 2.7)</td>
</tr>
<tr>
<td>Time deviations</td>
<td></td>
<td>Duration Contrast rule (k = -2, amp = 0)</td>
</tr>
<tr>
<td></td>
<td>&amp; Sound level deviations</td>
<td>Punctuation rule (k = 2.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phrase Arch rule applied to three phrase levels (level 1, k = 2.7; level 2, k = 1.5; level 3, k = 1.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High Loud rule (k = 1)</td>
</tr>
</tbody>
</table>

5 Novel Models for Sound Control

Recently, the articulation rules described above have been applied to the design of control models for sound synthesis (Bresin, Friberg and Dahl, 2001). The aim was to provide a more natural and realistic control of synthesis algorithms that typically fail to allow sufficient control and to produce a natural acoustic behavior of transitions between adjacent tones (Dannenberg and Derenyi, 1998).

The starting point was results from previous research. Many investigations have shown that human locomotion is related to timing in music performance. For example, it has been demonstrated how final ritardando in Baroque music and stopping runners follow the same tempo curve (Friberg and Sundberg, 1999). Friberg and co-workers (Friberg, Sundberg and Frydén, 2000) studied the relationship between music and human motion also in a direct way. Two subjects simulated tired, energetic and solemn gaits on a force platform. The vertical force patterns exerted by the foot were used as sound level envelopes for tones played at different tempi. Results from listening tests indicated that each tone, corresponding to a specific gait, could clearly be classified in terms of motion.

The articulation rules have been found to cover analogies with human locomotion too. In a previous work Bresin showed how note duration in staccato and legato articulation has a correspondence to gait duration in running and in walking respectively (Bresin, 2000).

These analogies between locomotion and music performance resulted in the design of new control models for synthesizing walking sound patterns. It seems likely that similar sound control models, based on locomotional patterns, can be developed further. In particular a model for humanized walking and one for stopping runners have been implemented with promising results (Bresin, Friberg and Dahl, 2001). These models control the timing of the real sound of one step on gravel. In these models the \<Score legato articulation> rule was used for controlling the step sound of a person walking on gravel, and the \<Final ritard> rule was used for controlling the step sound of a person stopping from running on gravel. The models were validated with a listening test: subjects could discriminate between walking and running sounds that also were classified according to the corresponding types of motion produced by the control models. More information about the control models and the experiment can be found in the paper by Bresin et al. (Bresin, Friberg and Dahl, 2001).

6 Conclusions

Articulation has indeed great importance in piano performance. Measurements of performances on MIDI grand pianos have shown that pianists vary the quality and quantity of articulation when coloring renderings of a piece according to different expressive adjectives. Generally, happy performances are characterized by staccato articulation while sad ones are played applying legato. The amount of legato articulation for one note was found to be depending of the IOI of that note. On the other hand, the amount of staccato articulation was independent of the IOI.

The measurements of articulation in different expressive performances have led to the design of a set of performance rules. These rules play an important role in the automatic rendering of a score, since articulation is one of the parameters used for differentiating the expressive coloring of performances.
Finally, on the basis of strong analogies between body motion and music performance articulation rules were developed in control models for sound synthesis algorithms. This and other applications can be seen as a further indication that studies of music performance can be used also for extra-musical applications.

7 Links
The art of music performance:
http://www.speech.kth.se/music/performance
The Director Musices program:
http://www.speech.kth.se/music/performance/download/
Articulation rules and sound examples:
http://www.speech.kth.se/music/performance/articulation
Deadpan and sad versions of Carl Michael Bellman’s song Letter 48. The <score legato articulation> and the <score staccato articulation> rules were applied in the sad version:
http://www.speech.kth.se/music/performance/germ

8 Acknowledgments
I would like to thank the organizers of RENCON 2002 for inviting me and for making this paper possible. My gratitude goes to all the colleagues involved in the works mentioned in this paper.

References