ABSTRACT

As one of the main expressive feature in music, articulation affects a wide range of tone attributes. Based on experimental recordings we analyzed human articulation in the late Baroque style. The results are useful for both the understanding of historically informed performance practices and further progress in synthetic performance generation. This paper reports of our findings and the implementation in a performance system. Because of its flexibility and universality the system allows more than Baroque articulation.

Keywords
Expressive Performance, Articulation, Historically Informed Performance

1. INTRODUCTION

Humans achieve expression in music performances by several features. Whatever is additionally named by different authors[12, 15, 5], all of them conform to three expressive features, which are timing, dynamics (loudness) and articulation [10].

Today’s performance systems established articulation as tone duration [15], sometimes reduced to the span between legato and staccato [7] or non-legato [13]. In this respect expressive articulation was measured [8, 14] and also implemented into expressive performance rule systems [4].

Duration is indeed the most striking feature of articulation but is not its only one. Articulation describes the forming of a tone in all its facets. This also includes loudness, timbre, intonation, and envelope characteristics.

This paper aims at three major tasks concerning expressive articulation: First, articulation is supposed to influence duration, but all remaining tone features like loudness, timbre, intonation, and envelope characteristics as well. Section 2 introduces the whole range of these aspects. Based on the first task, Section 3 shows our method demonstrating that different articulations change these tone features. The analysis exemplifies Baroque articulation. Consequently, Section 4 describes the implementation of articulation features including the possibility to freely define further articulation styles. A conclusion follows in Section 5.

2. WHAT IS ARTICULATION?

To trace the meaning of articulation it is first necessary to be aware of the difference between the notation of music and its actual performance. Throughout this article the terms ‘note’ and ‘tone’ are strictly distinguished. A note is a symbol of a musical event. Its attributes indicate pitch, length, loudness, onset and timbre (instrumentation). A tone, by contrast, is the actual physical event in a performance. Its attributes are pitch, duration and so forth that all correspond to the referring indications of the note. Expression in music performance touches deviations from tone attributes and note indications. Both timing and dynamics shape musical structure by influencing tone onsets and loudness values, respectively. Articulation as the forming of the single tone adds further deviations (e.g., an accent on a crescendo). This concerns all tone features, which are loudness, pitch, duration, envelope, and also timbre. The following description briefly introduces the parameter space articulation is involved in.

Envelope: A tone can consist of the four parts attack, decay, sustain and release. Every part can show different proportions or even be absent. On the whole they describe the loudness progression over a single tone.

Duration: It is the time from tone-onset to its offset, either proportional to the inter onset interval (IOI) or in absolute time.

Loudness Deviations: Independently from the dynamic shape of a musical section, certain articulations influence loudness.

Timbre Deviations: Some articulation instructions particularly refer to playing techniques (pizzicato, hampered bow strokes). They affect timbre changes that are neither exclusively caused by loudness changes nor depend on instrumentation.

Intonation Deviations: Similar to loudness and timbre, different articulations may influence intonation [9].

Articulation can affect one or more of these tone features. The following Section shows the analysis of envelope, tone duration, and loudness deviations.

3. MEASURING ARTICULATION

The meaning of articulations changed with time and place. We therefore considered a stylistic homogeneity. In addition, we wanted to choose key articulations for testing. The focus on German late Baroque/early Classic music fulfilled both conditions; it reduced the range of interpretation and, on the other hand, supports articulations that are both reasonable and still valid today [1, 18].
experimental exercises. and largely in later epochs. The definition refers to the
risk that body movements influence the position of the
verberation. On the contrary, a closer position would carry
tail less loudness differences and increased recorded re-
front of the instrument. A larger distance would have en-
shown in Figure 1.

The exercises followed eighteenth century articulation prac-
tices that had been taken from two major treatises on vio-
ern and Baroque instruments or between string, brass and
woodwinds. Admittedly, modern instruments showed an
test no systematic differences were found between mod-
cern duration, loudness and envelope characteristics. In all
results taken from the experimental recordings con-
mended in absolute time (absolute condition) and propor-
tional to the IOI (proportional condition). With a com-
parison of the dispersion—represented by the interquartile
relation (IQR)—different time attributions were made: The

<table>
<thead>
<tr>
<th>articulation</th>
<th>max percentile</th>
<th>offset percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>tenuto</td>
<td>.363 .661 .780</td>
<td>.799 .893 .963</td>
</tr>
<tr>
<td>neutral</td>
<td>.339 .469 .586</td>
<td>.692 .771 .876</td>
</tr>
<tr>
<td>bow vibrato</td>
<td>.413 .495 .571</td>
<td>.693 .780 .812</td>
</tr>
<tr>
<td>portato</td>
<td>71 95 124</td>
<td>533 727 769</td>
</tr>
<tr>
<td>staccato</td>
<td>73 87 101</td>
<td>532 567 744</td>
</tr>
<tr>
<td>staccatissimo</td>
<td>60 80 102</td>
<td>466 523 624</td>
</tr>
</tbody>
</table>

Table 1: Loudness distributions show time proportions to IOI, except italic letters that are ms values.

3.1 Methodology

Figure 1 shows all articulations that were analyzed. They are
described as follows [1, 17, 18, 19]:

tenuto: Tones are to play as long as possible but clearly
separated from each other.

neutral: If notes are without any annotation it is to decide
how to articulate them. This depends on the respective
common practice. The analysis refers to a promi-
ent example in Baroque music: Eighths (quavers) in
the accompaniment are played short [17, 19].

staccato/staccatissimo: It means very short/ as short as
possible.

bow vibrato: All notes under a slur are played by a con-
tinuous increasing and decreasing bow pressure but
without any stopping of the bow stroke.

portato: All notes under a slur are played with one bow
stroke but clearly separated by stopping the bow.

Because most standard Baroque articulation instructions
emphasize duration, two special articulations with an em-
phasis on (string) playing technique were added (Figure 1g
and h).

We recorded ten professional musicians playing Baroque
and modern instruments; altogether we recorded 14 differ-
ent instruments, including strings, brass and woodwinds.
The exercises followed eighteenth century articulation prac-
tices that had been taken from two major treatises on vi-
olin playing: Johann Friedrich Reichardt [19], i.e., Reichardt
exercise, and Leopold Mozart [17], i.e. Mozart exercise, as
shown in Figure 1.

Two AKG 1000 microphones were placed two meters in
front of the instrument. A larger distance would have en-
tailed less loudness differences and increased recorded re-
verberation. On the contrary, a closer position would carry
the risk that body movements influence the position of the

1Both annotations differ in Baroque string playing [16, 17]
and largely in later epochs. The definition refers to the
experimental exercises.

d two AKG 1000 microphones were placed two meters in
front of the instrument. A larger distance would have en-
tailed less loudness differences and increased recorded re-
verberation. On the contrary, a closer position would carry
the risk that body movements influence the position of the

1Both annotations differ in Baroque string playing [16, 17]
and largely in later epochs. The definition refers to the
experimental exercises.

www.sonicvisualiser.org
3.3 Discussion

Articulations differ not only in duration, though it is the most prominent feature. Tenuto, staccato and staccatissimo were discriminated by duration. Short articulations like the staccatissimo and also the staccato did not completely turn out to be proportional to the IOI. Further analysis should focus on this problem in particular, for it is not unthinkable that staccato and staccatissimo offsets depend on absolute time, which may be additionally influenced by the absolute IOI. Consequently, the staccatissimo as an “as short as possible” articulation showed an envelope nearly reduced to its attack and release. Thus, the fastest possible tempo in a tenuto articulation can be deduced from these differences: It is reached if the notes are as short as the staccatissimo attack (it corresponds to the maximum of 80ms, see Table 1) and at the same time correspond to the tenuto proportion. This is fulfilled by playing sixteenth notes at 167 quarter notes per minute. In his masterpiece on Baroque
Articulating a note means more than changing its duration. The investigations described in the previous Section reveal, for instance, a very distinctive loudness component which did not derive from conventional dynamics or metrical accentuation.

Furthermore, all articulations are not created equal. Identi
cal instructions are rendered differently in different stylistic or expressive contexts. It can even be inhomogeneous within one and the same style. A neutral articulation in a light-footed Baroque dance is shorter and more detached than in Baroque hymns. Moreover, the articulation vocabulary is not fixed. The basic articulations, as introduced and analyzed in the previous Section, were established in the Baroque era. The Romantic period and Serialism, however, invented a lot of further articulations, mostly finer differentiations of the Baroque articulations. In addition, multiple articulation instructions can be combined. A typical example is the addition of an accent over a somehow articulated note for additional dynamic emphasis.

Our approach to articulation in synthetic performances shall provide this flexibility and extensibility. This would, for instance, facilitate an analysis-by-synthesis approach to expressive articulation and further customization.

All implementations were done in a MIDI-based music engine framework [2]. It loads the raw MIDI data and several performance styles in XML format which explicitly describe how to render expressive MIDI sequences. Such a performance style includes information on all performative aspects such as tempo, rubato (self-compensating microdeviations in timing), asynchrony, ‘human’ imprecision, dynamics, metrical emphasis, and, of course, articulation. It can be created, for instance, by a performance generation system and/or manually edited. Regarding the aspect of articulation, the following formalisms have been developed and implemented.

4.1 Concept

The basic idea is to formally separate the definition of articulations from their application to the concrete musical
context. Therefore, the formalisms articulation style and articulation map are introduced.

An articulation style is a pool of articulation definitions. One such definition describes the manipulations which the articulation makes to an input note. The developer can freely edit them and create any, even new, articulations. Several styles with possibly differing definitions can be defined.

An articulation map is a sequence of instructions to articulate a concrete musical context, in our implementation a raw MIDI sequence. Only high-level descriptors are used which act as lookup references to the definitions in an articulation style. Articulation maps can be manually edited or output of a performance generation system. For playback, the instructions are then rendered into an expressive MIDI sequence.

4.2 Articulation Style
An articulation style $S_s (s = 0...t)$ is an element of a set of styles $S = \{S_0, ..., S_t\}$. It is a set of articulation instructions.

$$S_s = \{I_0^s, ..., I_t^s\}$$

An articulation instruction defines the manipulations it applies to a raw note by a list of manipulators.

$$I_i^s = (M_0^{s,i}, ..., M_n^{s,i}) : i = 0...j, n \geq 0$$

One such manipulator is a triple which indicates mode and value of the manipulation, and the note attribute to be changed.

$$M_m^{s,i} = (\text{mode}_m^{s,i}, \text{attribute}_m^{s,i}, \text{value}_m^{s,i}) : m = 0...n$$

Possible attributes to change are the note’s duration and velocity. Three manipulation modes are distinguished

$$\text{attribute} := \{\text{value} : \text{set mode}, \text{attribute} + \text{value} : \text{add mode}, \text{attribute} \cdot \text{value} : \text{percent mode}\}$$

Negative results are restricted in either case. The implementation further rounds the results to integers for proper performance in MIDI. If the velocity attribute is modified, the result is capped if greater than 127. While the set mode sets the attribute absolutely, the add and percent mode assume that the attribute is already set at the basic value (e.g., dynamics). Articulation quasi adds finer differentiations.

For the manipulation of duration values the implementation provides a millisecond version. Thereby, it allows to set durations not just in MIDI ticks but also in milliseconds. Corresponding timing conversions are described in [3].

In the implementation, the indices $s$ (style index) and $i$ (instruction index) were replaced by unique name strings. These descriptors are used in the articulation map as lookup references. They ease the manual editing of articulation maps as they allow to call a spade a spade. This is particularly necessary for sample-based MIDI playback. The Vienna Instruments sampler [21], for instance, provides specialized sample sets named ‘staccato’, ‘legato’, ‘portamento’, ‘pizzicato’ and so forth. These have to be triggered separately by designated controller messages. If an instruction with a known name occurs, our system automatically activates the respective sample set. The current vocabulary comprises the terms ‘portamento’, ‘legato’, ‘moltoLegato’, ‘legatissimo’, ‘nonlegato’, ‘portato’, ‘marcato’, ‘tenuto’, ‘staccato’, ‘staccatissimo’, ‘spiccato’, and ‘pizzicato’.

Nonetheless, the developer is free to define and name any articulation. Only the neutral articulation plays a special role. If an articulation style defines a neutral articulation, it is applied to all notes except for those with individual instructions in the articulation map.

4.3 Articulation Map
The articulation map is a sequentially ordered list with two types of elements, style switches and articulators. A style switch is a tuple

$$\text{switch} = (d, s) : d \geq 0, S_s \in S$$

with the tempo-independent date $d$ (e.g., in MIDI ticks) from when on articulation style $S_s$ acts as the underlying style, quasi as the lookup style. Its range is terminated by the date of the next switch. The first element in an articulation map has to be a style switch.

The other elements in the map are articulators. This is a tuple

$$\text{articulator} = (\text{note}, (i_0, ..., i_k)) : I_{i_0...k}^s \in S_s, k \geq 0$$

that indicates the note note to be articulated, and the articulation instructions $I_{i_0...k}^s$ therefore. The instructions are successively applied to the note. In this way, instruction combinations are possible, like tenuto and accent, or even double accent (one accent may raise the velocity by a certain amount, double accent does it twice then).

But the note to be articulated must be located in the MIDI sequence. All necessary information therefore are given by the note term. It is a list of properties (date, pitch, duration, MIDI channel and port). Most of them are optional, only the date (in MIDI ticks) is always required. If the given information allow no clear discrimination of several notes, all contemplicable notes are articulated.

4.4 Discussion
The described system allows to model a broad variety of articulations. All basic articulations and a big spectrum of nuances are possible. The quality of the sounding results depends, of course, also largely on the underlying sound technology (quality of samples or synthesized sound). Nonetheless, the expressive intentions of the articulation are still discernible, even with low-quality sounds.

The system implements articulation manipulations in the loudness and duration domain. Even the timbre domain can be handled to a certain extent through specialized sampler control. Serious boundaries exist with regard to envelope and pitch/information. The latter can be added easily: The manipulators’ attribute domain can be extended by pitch. The respective changes can be implemented by pitch wheel controller messages.

The flexible manipulation of envelope characteristics, by contrast, necessitates more access to the sound generation process. Special playing techniques, like con sordino/muting the instrument, bowing with bridge proximity, playing and singing into the mouthpiece at the same time etc. necessitate advanced synthesis methods and full synthesizer control.

Up to now, articulations are rendered into MIDI sequences without any variance. A human performer is not able to reproduce an articulation that exactly, though. By a certain random variation further liveliness can be introduced. But which amount of variance would be reasonable, anyway? It can be more for layman musicians and less for professionals. Concrete values are still unknown and await detailed analysis and evaluation.

Similarly, our implementation uncovered a further very fundamental question. Is articulation additive? A staccato articulation may set the tone length to 170ms. Adding a marcato articulation may shorten the tone length by further twenty percent and raise its velocity to the next dynamic
level (e.g., from mf to f). Is this adequate? More meaningful combinations, like staccato+tenuto, staccato+legato or tenuto+legato are well known from music notation. On the other hand, combinations of instructions which affect the tone duration might rather averaged than added. A short articulation combined with a long may result in a medium duration (e.g., staccato + tenuto ≈ portato). But how are both articulations weighted? The generative nature of our articulation approach inspires further investigation into such effects.

5. CONCLUSION
Articulation, if not reduced to a mere tone duration feature, offers a big potential for synthetic music performances. When human performers change articulation, they alter tone features like loudness, envelope, duration, and, as presumed, timbre and pitch. Our analyses of late Baroque articulation showed that these variations do not derive from other performance features like timing and dynamics.

Although these features are systematically distinguished they are perceived as fused. As an example, we demonstrated the pseudo-crescendo by increasing duration, which is already known in theory [1] and practice. These fusion effects are worth further investigations.

Music includes a great amount of common practices, but these can disappear every time as music progresses. Research into music and its original performance accordingly should consider both structure and origin. Synthetic performance systems can contribute to this rediscovery.

This paper further described the implementation of articulation features as part of such a performance system. The implementation allows to flexibly define any articulation styles and apply them to a given raw MIDI sequence. A musicologically interesting part is the chance to easily render one and the same raw sequence in different styles, adapt the styles, and explore the effects on the performance.

Of course, were articulation added solely, no performance would sound expressive. Articulation rather supplements music expression as timing and dynamics do as well. Hence synthetic and expressive music cannot sound human-like until these three are combined.

Acknowledgement
We like to express thanks to all musicians for their participation in the recordings and for the inspiring dialogs.

6. REFERENCES