

Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance

Giovanni De Poli , Antonio Rodà and Alvisè Vidolin

depoli@dei.unipd.it, ar@csc1.unipd.it, vidolin@dei.unipd.it

CSC-DEI, University of Padova, Via Gradenigo 6a, 35131 Padova Italy

ABSTRACT

This paper describes an analysis of how the performance of a score differs when musicians are requested to play with differing expression. A professional violinist was asked to play short pieces of music in different versions expressing light, heavy, soft, hard, bright, and dark. For comparison, a normal, standard performance was recorded. Note-by-note analysis allowed us to identify the variations of the main acoustic parameters as a consequence of varying the expressive intentions. It was possible to identify two distinct expressive sources. The first refers the musical structure of the period, its division into phrases and the continuous alternation of tension and relaxation points. The second depends on the expressive intentions that the musician wants to convey to the listeners.

INTRODUCTION

Many investigations explore how musical structure influences music performance. There are a few studies on how musician's expressive intentions are reflected in the performance. In this context, expressive intention is taken to mean how a musician's inspiration varies according to certain adjectives that have been given before each performance (Gabrielsson, 1995). Recently, more attention has been paid to music performance and some interesting studies have been carried out; for an extensive review, see Gabrielsson (1997) and Palmer (1997). This field of investigation is stimulating ever greater interest not only from the scientific and cognitive point of view, but also from the applied one, both in terms of automatic music performance and, more generally, in multimedia systems. Moreover, the study of musical performance and interpretation raises problems that go beyond the specific area of musicology, running into the wider field of non-verbal expressive communication involving various aspects of human perception.

Music performance is often viewed as part of a system of communication whose leading actors are the composer, the performer and the listener (Kendall & Carterette, 1990). The performer uses the expressive means of his own musical instrument to communicate to the listener the contents of the music. Ambiguities in musical notation allow a player considerable freedom in deciding how to interpret the music's content. Thus, the same musical score is performed differently by different performers, or the same performer may perform a piece differently on separate occasions.

In Western art music, one function of interpretation is to highlight a particular structural content (Clarke, 1987). Many studies about music performance have demonstrated that musicians use small variations to communicate those aspects of the musical structure they wish the listener to perceive. These variations refer to deviations in timing, articulation, amplitude, *etc.* in relation to a literal interpretation of the score. Analyses of performances by different pianists (Repp, 1990)

show that there are differences among performances, but there are also strong commonalities (Repp, 1992) that can be related to the structure of the music. Performers emphasize the end of each phrase with a *rallentando* (Sunberg & Verrillo, 1980). Todd proposed a computational model of deviations of timing (Todd, 1985) and dynamics (Todd, 1992) to emphasize the phrase hierarchy. This model compares the variation of tempo with velocity in the equations of mechanics (Todd, 1995). Another computational model formalizes timing patterns specific of individual composers (Clynes, 1986), which seem to be validated by perceptual analysis (Clynes, 1995). Drake & Palmer (1993) found out that pianists tend to underline different accent through variations on timing, intensity and articulation.

Sundberg and co-workers (Sundberg, 1983; Friberg, 1991; Friberg *et al.*, 1991) developed a system of quantitative rules to study musical interpretation. Following an analysis-by-synthesis method, the rules describe some of the principles the musician follows to introduce deviations from the nominal values of the score. Such deviations work through manipulation of timing, intensity, intonation and vibrato in accordance with the instrument used. A fuzzy logic approach to the rules system was proposed by Bresin *et al.* (1995). Bresin (Bresin *et al.*, 1992; Bresin and Vecchio, 1995) suggested a model based on an artificial neural networks system that analyzes the parameters of the performance of a real pianist. This system showed good generalization properties, as it was possible to perform different scores, using the same neural networks, maintaining the performing style of the pianist who trained the network.

Another function of interpretation is to highlight particular content referring to moods or feelings. Such content is connected to the performer's expressive intentions and significantly affects the performance itself. Senju & Ohgushi (1987) carried out some psychological experiments on an excerpt of a Violin Concerto performed with 10 different nuances: weak, powerful, bright, sad, sophisticated, beautiful, dreamy, fashionable, simple and deep. The results revealed that the player's expressive intention and the listener's impression were in general agreement. Askenfelt (1986) asked the player to render one tender and one aggressive performance of an excerpt from Beethoven's violin concert. The aggressive version was characterized by a higher mean bow force and a more abrupt attack of the notes when compared to the tender version. Gabrielsson & Juslin (1996) asked nine professional musicians to play some brief melodies so as to communicate specific emotional characters: happy, sad, angry, fearful, tender, solemn and no expression. The comparison among perceptive tests and acoustic variables demonstrated that the performer's expressive intentions had a marked effect on all analyzed variables. In spite of some similarities among the performers, differences in emotion encoding have been discovered.

Canazza *et al.* (1997a, 1997b, 1997c, 1997d) made acoustic and perceptual analyses on some performances of a clarinet piece. The performances were inspired by a group of sensorial adjectives. The acoustic analysis reveals acoustic parameters that serve to differentiate performances of the same score. Perceptual analyses confirmed that the listener's experience and the performer's intention basically agreed. A two-dimensional semantic space was derived, and the axes turned out to be correlated with the principal acoustic parameters measured.

The results of these studies show a complex pattern. Performers may adopt many actions and expressive choices. The performance, therefore, appears to be the result of a complex interaction of different and sometimes contrasting principles. No straightforward explanation can be given in order to justify why in a certain moment the performer produces a particular accent, a *rallentando* and so on. Different expressive motivations can lead to the same acoustic effect just as different instrumental effects are available for the same expressive idea. Nevertheless, the results of the studies previously set forth prove that there is an agreement among listeners in defining as acceptable or not a certain musical performance and in recognizing a musician's

particular expressive intention. It is therefore possible to say that there are some principles that underlie plausible performances, and if these principles are ignored or broken, the performer may be perceived as not very musical (Repp, 1992).

As already stated, the work presented here aims to study the ways and means adopted by performers to communicate particular content. Such content is essentially connected to two aspects of musical communication, which we will refer to as sources of deviations. This terminology derives from the consideration that each expressive deviation introduced by the performer can be brought back to the expressive motivation (source) that caused it. The first source of deviation deals with musical structural aspects such as the phrasing, hierarchical structure of phase, harmonic structure and so on. The second involves those aspects that are indicated with the term expressive intention and which refers to the communication of moods and feelings.

This work compares the values of acoustic parameters note by note and performance by performance, and aims to demonstrate a strong correlation between the variations of these parameters and the two deviation sources. Such study allows the recognition of expressive deviations correlated to musical structure and to distinguish them from deviations related to expressive intentions. The effects of the second source will be shown in more detail. It will therefore be possible to determine some of the cues followed by the performer to communicate personal expressive intentions. The first source, instead, will be briefly presented, with the only aim to show how this source affects the second one. A detailed report of the influence of musical structure on performance, in fact, go beyond the goals of this paper.

Moreover, the expressive deviations are often different in different types of music, instruments and individual performers. Every musical instrument, in fact, has its own expressive resources (vibrato in strings, the tongue in wind instruments, etc.) used by musicians to communicate their expressive intentions. It is inevitable, therefore, that the results of any acoustic measure depend not only on the score, but also on the characteristics of the instrument used and the choices taken by the musician. Consequently, it is necessary to compare the data relative to different scores, musicians and instruments in order to identify the expressive rules that can be considered valid in a general way and which are specific cases.

At CSC of the Padua University, some acoustic analyses have been carried out on various musical pieces using different instruments and performers. Up to now, performances involving the clarinet (Canazza et al., 1997b) and the piano (Battel & Fimbianti, 1997) have been analyzed. In this paper, we will illustrate and discuss analyses of violin performances. The violin, in fact, is an instrument that offers the performer many expressive resources. The results relating to the parameters that are most important in conveying expressive intentions will be presented and compare with previous results.

This research was supported by Telecom Italia, under a research contract Cantieri Multimediali.

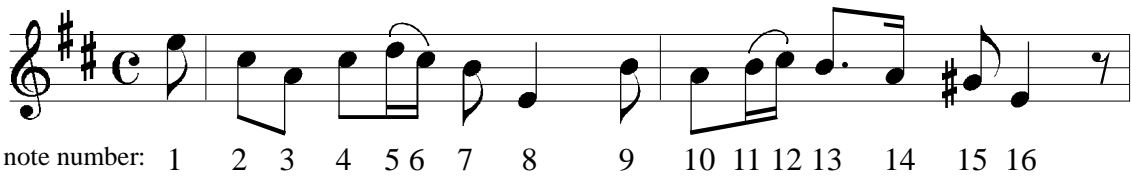


Fig. 1: Musical phrase used in the experiment.

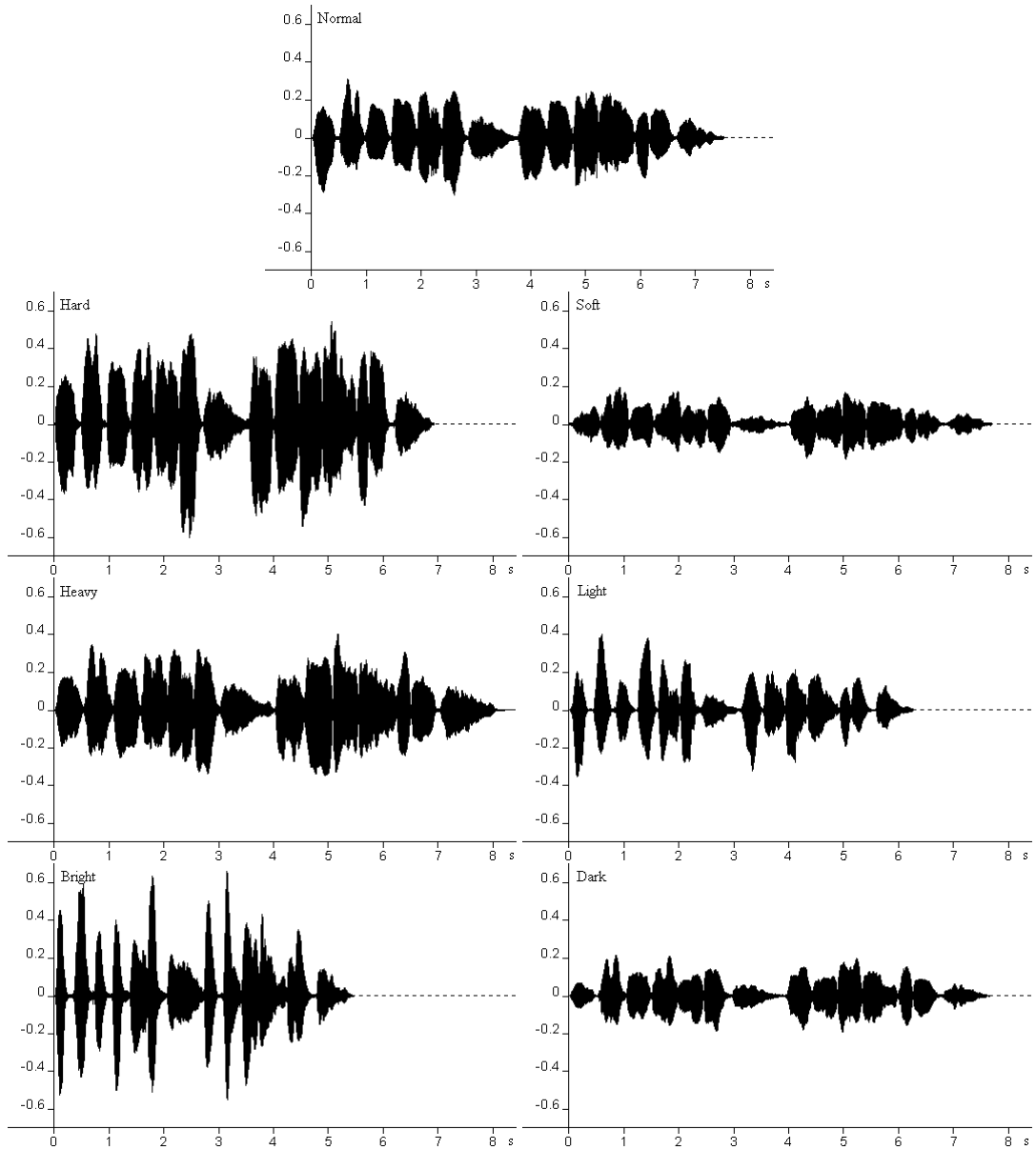


Fig. 2: Waveforms of the seven performances

ANALYSIS OF THE ACOUSTIC PARAMETERS

Experiment

A professional musician was asked to perform an excerpt from Arcangelo Corelli's Violin Sonata in A Major, V Op. (Fig. 1), inspired by the following (Italian) adjectives: light (*leggero*), heavy (*pesante*), soft (*morbido*), hard (*duro*), bright (*brillante*), and dark (*duro*). The “flat” or literal performance, described by the adjective normal (*normale*), was also added and used as a standard measure of comparison in the acoustic analysis of the various interpretations. Unconventional adjectives, in the musical field, were deliberately chosen in order to give the performer the greatest possible liberty of expression. We chose sensorial adjectives to have a compact and coherent semantic space. The recordings were carried out in three cycles, each cycle consisting of the seven different interpretations. The musician then chose the performances that, in his opinion, best corresponded to the proposed adjectives. This procedure is intended to minimize the influence that the order of execution might have on the performer. The recordings were carried out at the CSC of Padua University in the monophonic digital form at 16 bits and 44100 Hz (sound examples n° 1-7 of the electronic appendix).

Methodological Issues

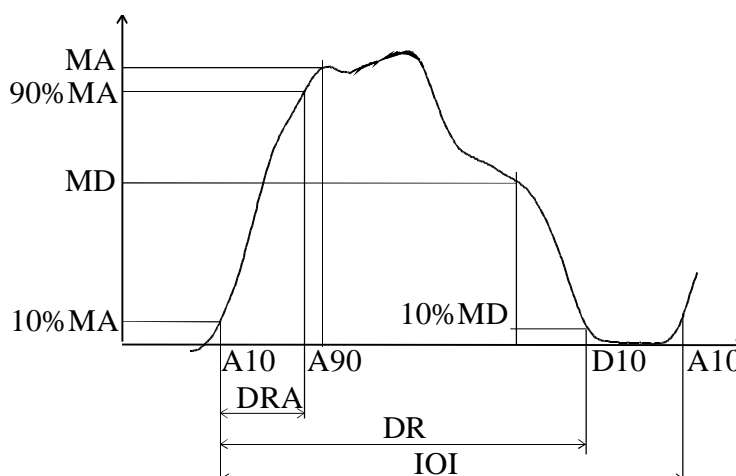


Fig 3: Definition of parameters used for time analysis

The amount of data given by the audio recordings is extremely high. The study of expressive characteristics in performances requires the definition of a series of parameters containing information at a higher level of abstraction than the acoustic waveform. Choosing which parameters to measure and how to measure them is always a very critical moment in every acoustic analysis. At the moment there is not a wide enough literature on the subject to allow a standardisation of analysis methods and therefore it is often difficult to compare the results of different works. The majority of performance studies analyze the note duration by measuring the inter-onset interval (IOI) between the beginning of two successive tones, and the duration (DR) between the beginning and the end of a tone. These parameters are particularly easy to measure when recordings are in MIDI format where the information of note-on and note-off are available. Notice that in MIDI, note-off indicates the beginning of the decay and not the effective end of the note.

Only a few works analyze how amplitude envelopes and spectral signals change depending on expressive intentions. Besides, they are carried out in a qualitative manner and subject to visual and perceptual limitations. On the other hand a quantitative analysis is rather difficult, especially because the amplitude envelopes have particularly irregular shapes (see Fig. 2). As a result, in order to define and measure the principal parameters, it is necessary to make arbitrary choices, without any certainty that the choices are in any sense “correct.” A method able to verify such choices is therefore needed. To this end, a first estimation of parameters is used to create synthesized performances. These are compared to the original performances to confirm the choices made in the previous measurement phase. This analysis-by-synthesis approach serves as a check that the data and results have meaning in human perceptual terms. This is always important in studies about performance.

Information about acoustic parameters cannot be obtained directly from the sampled signal. Some processing of the acoustic signal is in fact needed to measure the parameters useful to the analysis. The study of dynamics and timing is based on the signal amplitude envelopes. To measure these envelopes, the signal magnitude was filtered using a 4th-order Butterworth filter with a cut-off frequency $f_T=100\text{Hz}$. The study of vibrato needs information about the pitch/time profile, which was calculated using a pitch detection algorithm based on zero-crossings. In addition, a Short Time Fourier Transform (having a window of 30ms) was calculated in order to obtain information on the time evolution of the spectrum of each note.

For instruments such as the violin, timing analysis starts by measuring and modelling the amplitude envelopes of the various notes. In violin performances, the amplitude envelopes have shapes and characteristics that vary considerably (Fig. 2), both in terms of expressive intentions and among different notes within the same interpretation. Some notes have an extremely rapid onset and the amplitude reaches a peak value almost immediately. Other notes, on the contrary, have a slow, gradually rising onset. Moreover, an amplitude vibrato was present in almost all tones. There is no standard procedure for measuring the envelopes. Fig. 3 shows the parameters taken into consideration in relation to the envelopes.

The MA (end of attack) and MD (start of decay) points were determined by analyzing the envelope of each single note because of the great variability in the shape. The Inter-Onset Interval (of n -th note) has been defined as

$$\text{IOI}(n) = A10(n + 1) - A10(n),$$

while the note duration has been defined as

$$\text{DR}(n) = D10(n) - A10(n),$$

given that the use of A10 and D10 is often assumed to better reflect the perceptual onset and offset of a tone. On the basis of these choices, articulation has been expressed by means of the legato parameter defined by

$$L(n) = \text{DR}(n) / \text{IOI}(n).$$

Lastly, the attack duration (DRA) is defined as the time taken by the amplitude envelope to pass from 10 to 90% of the MA value, that is

$$\text{DRA}(n) = A90(n) - A10(n).$$

Mean tempo

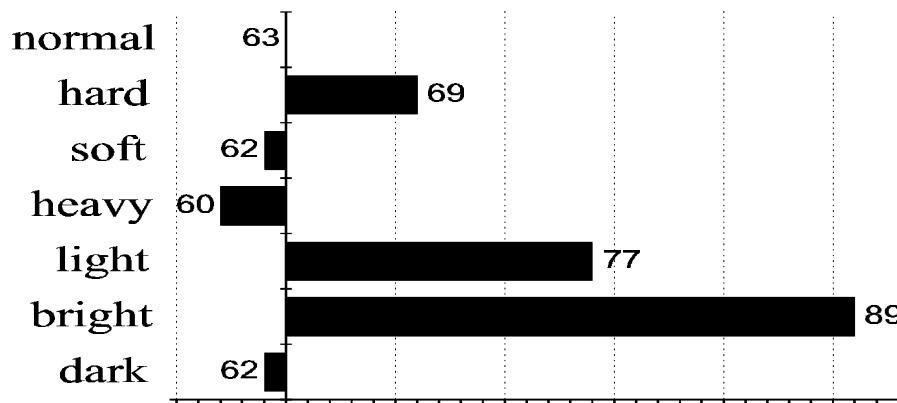


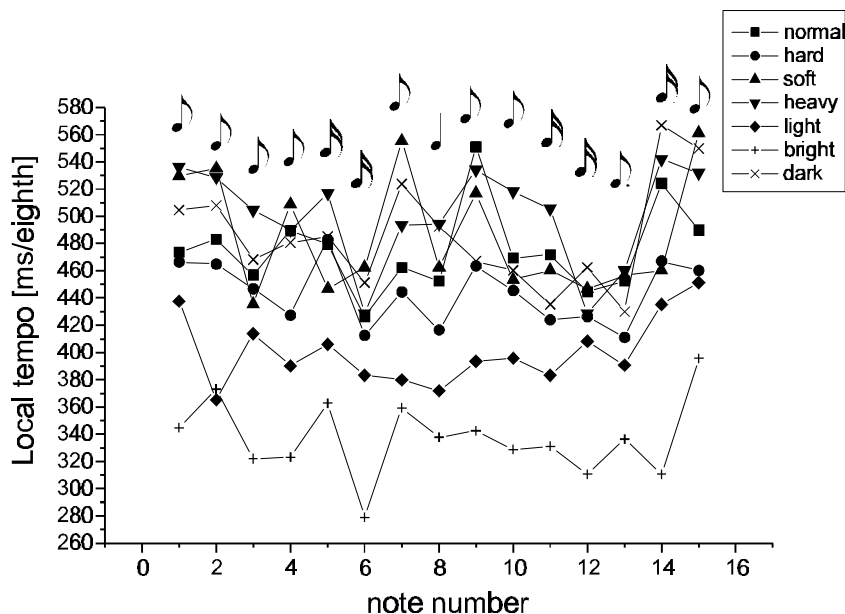
Fig. 4: Mean tempo in the seven performances

Fig. 4 shows the mean tempo (metronome MM) of the seven interpretations and it can be seen that those inspired by the adjectives bright and light were played very fast, while the heavy piece was the slowest. The violinist made great use of this expressive resource in order to differentiate his performances. It can be seen that there was, in fact, a wide range of tempi represented, with the fastest (bright) performance being played more than 40% faster than the slowest (heavy).

Local tempo

Fig. 5 shows the local tempo (Gabrielsson, 1997) in the seven performances. This was calculated by dividing $IOI(n)$ by the number of eighth-note beats in the n -th note. For example, the quarter note duration is divided by 2 and sixteenth note duration is multiplied by 2.

Fig. 5: Local tempo in the seven performances.



In the opinion of the performer, the sensorial adjectives suggested are not suitable to stimulate timing variations. It was suggested that other adjectives (for instance, passionate) should be used to give the performer wider possibilities to act on notation value. The relatively small amount of deviation from the nominal, notated values seems to corroborate this hypothesis. In all these performances, in fact, deviations are below 20%. This value, compared to the results obtained

from previous studies (Gabrielsson & Juslin, 1996) seems to be rather low, demonstrating the limited use of this kind of expressive resource.

Significant differences between pieces have nevertheless been found in the performance of some rhythmical patterns. For each pair of local tempo measurements (notes 1 and 2, 3 and 4, etc.) we define the Up-Down ratio (UDR) as the ratio between the value of the note on the upbeat and on the downbeat (e.g. IOI(1)/IOI(2), etc.). The nominal value of the UDR is 1. The variance test (ANOVA) proved a significant effect of expressive intentions on this ratio ($F(6,14) = 4.03$, $p < 0.02$). The calculated average values (light = 1.100, hard = 1.019, heavy = 1.001, normal = 0.965, dark = 0.954, bright = 0.938, soft = 0.888) showed that in the light piece notes on the upbeat last longer than those on the downbeat. In the soft piece, on the contrary, the performer emphasized notes on the downbeat.

Another particularly significant rhythmic pattern is the one formed by notes 13 and 14 (the dotted eighth note and sixteenth note). The ratio between the nominal length of these two notes is 3. The measured values (dark=2.276, heavy=2.548, normal=2.548, hard=2.641, light=2.693, soft=2.975, bright=3.248) proved that in all these pieces the dotted note was shortened. The only exception is the bright version.

Articulation

Articulation is expressed by means of the legato parameter $L(n)$ defined by the ratio between the duration and IOI of the n -th note. Fig. 6 shows the values measured for the legato parameter of each note. For each variation observed, we have worked backwards from the performances to uncover the source of the performance expression.

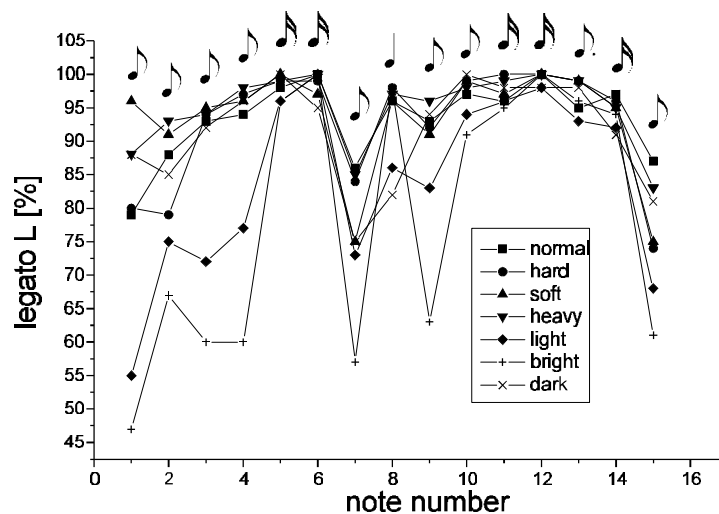


Fig. 6: value of legato in the seven performances. The value of the 16th note was not measured as the onset of the following note was not played.

Fig. 6 shows how all the profiles tend to be those of the double curve type. This trend reflects the musical structure of the phrase, or rather its division into two half-phrases (the first going from note 1 to note 8 and the second going from note 9 to note 16). Beside this common trend, quite considerable differences can be found between one performance and another. In particular, the legato values of the bright piece are smaller than those of the other pieces. This might be explained by the fact that the variations depend on the expressive intentions of the musician. By observing the data, it is possible to identify two distinct sources of deviations. The first refers the musical structure of the period, its division into phrases, half-phrases and the continuous

alternation of tension and relaxation points. The second depends on the expressive intention that the musician wants to convey to the listener. In this latter case, the choice was induced by the seven adjectives presented earlier in the experiment.

An analysis of variance (ANOVA) on the legato values by intentions (light, heavy, etc.), shows that the adjective had a considerable effect on how the piece was played ($F(6,98) = 4.5$, $p < 0.001$). The average legato value for each performance is shown in Fig. 7. It is interesting to note that, especially in the performance of the bright and light pieces, the legato values of certain notes differ markedly from the calculated average value for those notes. It is therefore necessary to analyze the legato values of each single note when the expressive intention varies. Fig. 3 shows that the 5, 6, 10, 11, 12, 13, 14 notes and, to a lesser extent note 8, have legato values which are very close to each other in all the performances. We will discuss this characteristic in the following paragraphs.

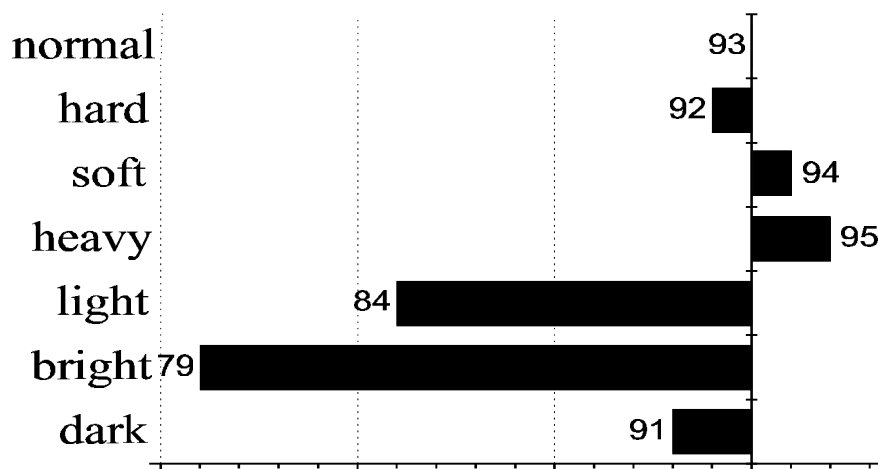


Fig. 7: Mean value of the parameter L in the seven performances.

Attack time

Fig. 8 shows the attack time values $DRA(n)$ measured for each note. All the performances had a double curve trend, reflecting the musical structure of the phrase, which can be divided into two half-phrases. In addition to this common trend, noticeable differences between individual performances can be identified. The attack times of the bright piece are generally shorter. It can be supposed that these variations depend on the expressive intentions of the musician. An analysis of variance (ANOVA) was made on the attack time values by intentions (light, heavy, etc.). It showed that these intentions had a significant effect on performance ($F(6,91) = 2.51$, $p < 0.03$).

The average attack duration for each performance was then calculated (Fig. 9). Just as we observed that some legato values were affected by expressive intention more than others, it should be noted that attack duration is relatively constant for some notes. For example, the 13th note it was played with rather short attack times in all the performances. Other notes show much greater variation.

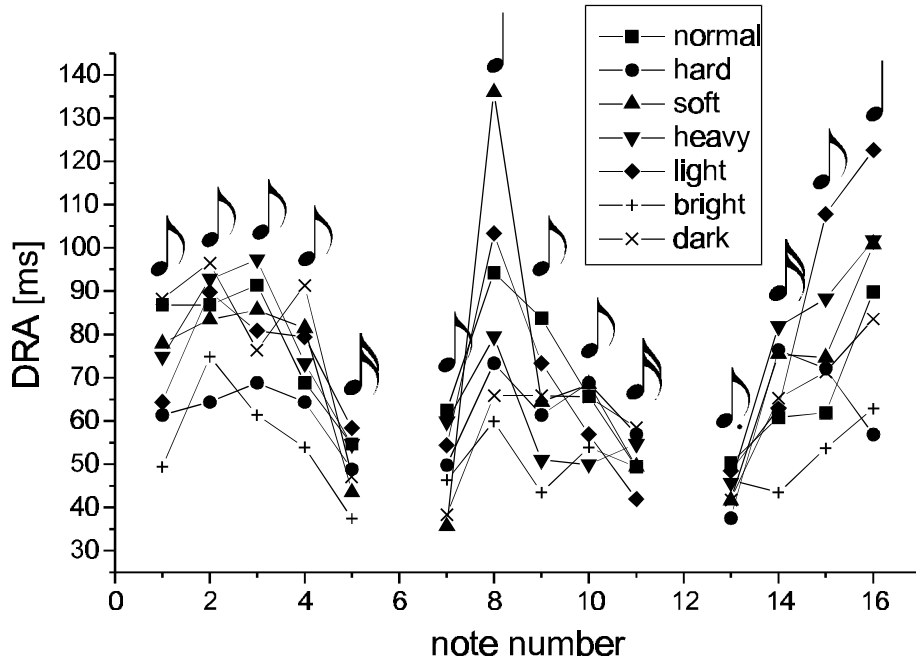


Fig. 8: Attack duration values in the seven performances. Notes 6th and 12th were played legato to the preceding note. In these cases it was not possible to determine an attack phase and therefore the DRA parameter was not measured.

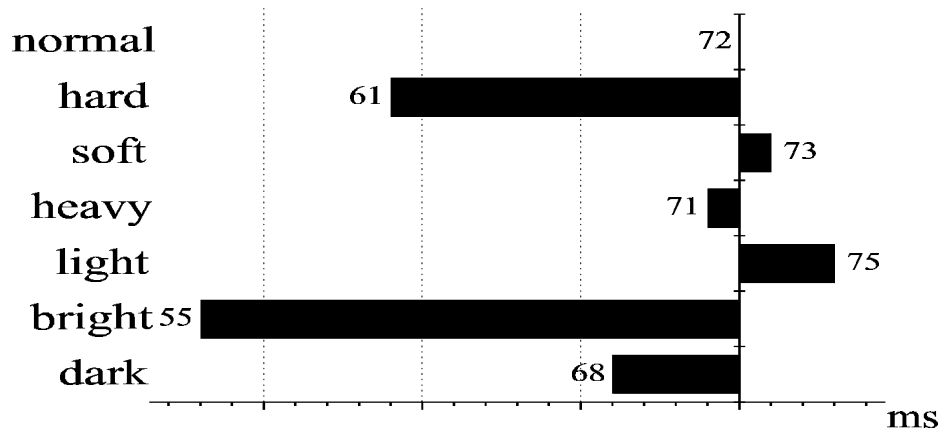


Fig. 9: Mean values of the parameter DRA in the seven performances.

Dynamics

The analysis of dynamics, based on the $MA(n)$ parameter (Fig. 3), requires some caution. While recording the pieces, the distance between microphone and musician was not kept fixed so as not to influence the performer's gestures. A change in intensity from one performance to another could therefore be due to the performer's movement. A comparison of the absolute values of $MA(n)$ in the various performances could therefore be of little significance. For this reason, the data were normalized according to the maximum amplitude of each performance. Fig 10 shows $MA(n)$ values expressed in dB. It can be seen that the intensity of every piece decreases in correspondence with notes 8 and 16, which are the final notes of the two half-phrases. This

proves that the violinist always respected the musical structure of the phrase, even when the expressive intentions varied.

The soft and dark pieces were less regular in loudness than the others, with a difference between maximum and minimum amplitude of about 20 dB. This trend is supported by the standard deviation values calculated for each performance (Fig. 11).

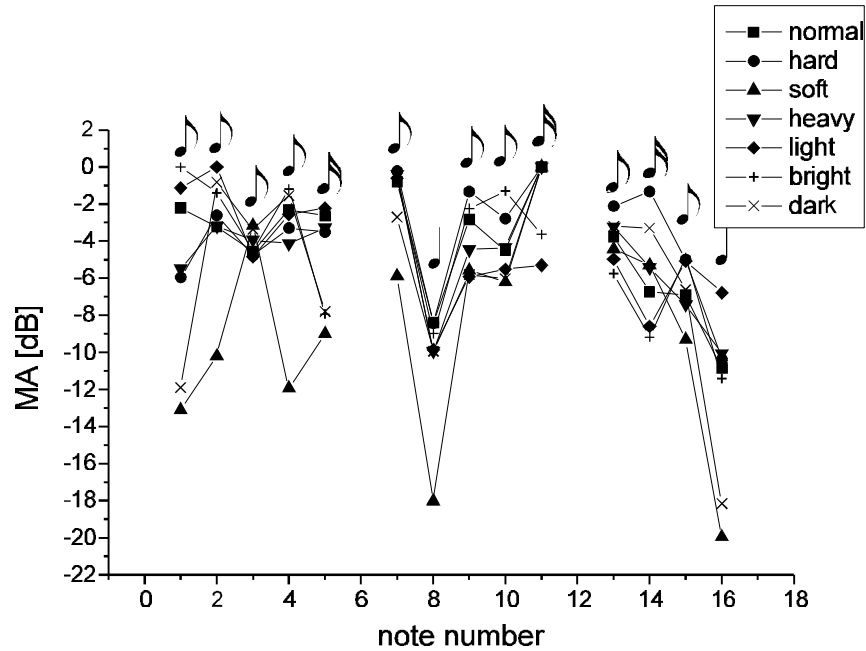


Fig. 10: $MA(n)$ in the seven performances. Notes 6th and 12th were played legato to the preceding note. In this case it was not possible to determine an attack phase and therefore the $MA(n)$ parameter was not measured.

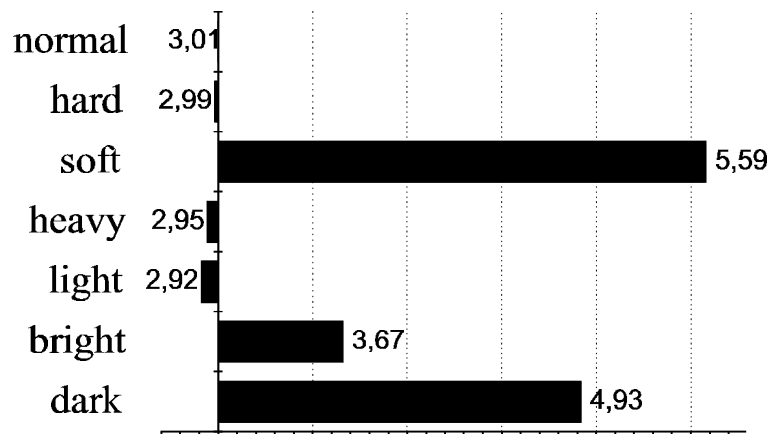


Fig. 11: Standard deviation of the parameter $MA(n)$ in the seven performances.

Sustain

Fig. 12 shows the values corresponding to the MD-MA(n) difference in dB for the seven performances, where MD is the amplitude at the beginning of the decay (Fig. 3); a positive value indicates that in the sustain there was an intensity crescendo while a negative value indicates a diminuendo. The values reproduced on Fig. 13 show that, on average, the soft performance notes are characterized by amplitude envelopes which increase in intensity during the sustain. In the other performances, on the contrary, there is an opposite trend: the note envelopes in the light version reach their maximum amplitude at the end of the attack phase and then decrease to the decay phase.

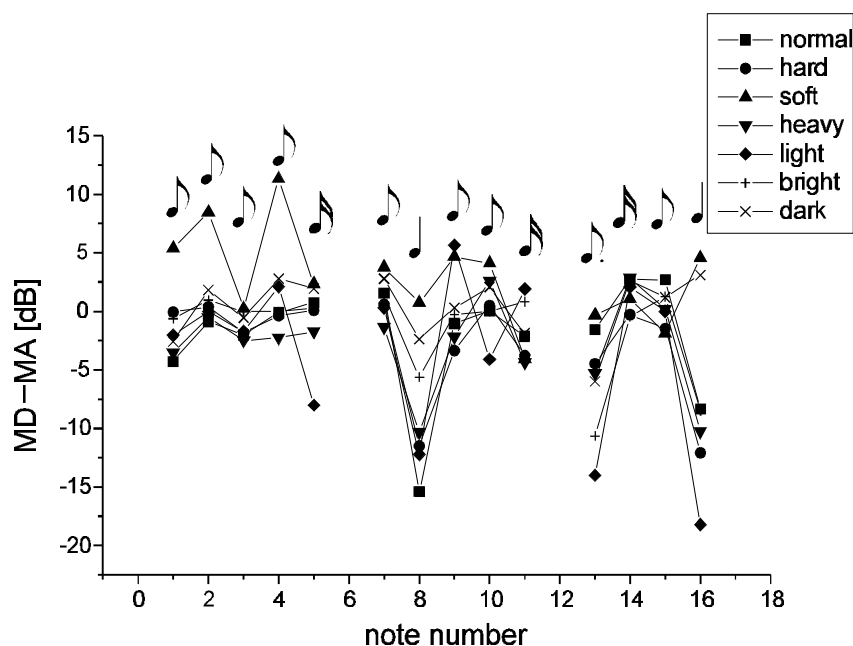


Fig. 12: MD-MA(n) in the seven performances. Notes 6th and 12th were played legato to the preceding note. In this case it was not possible to determine an attack phase and therefore the MD-MA(n) parameter was not measured.

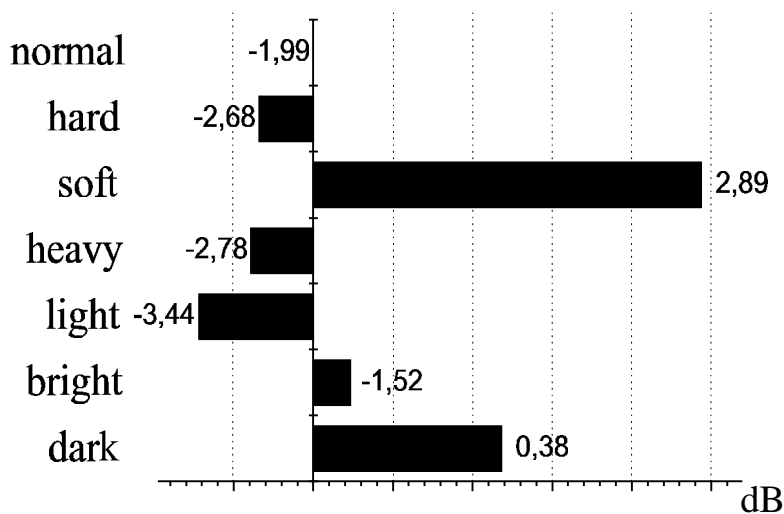


Fig. 13: Mean values of the parameter MD-MA(n) in the seven performances.

Brightness

The brightness can be quantitatively represented (Fig. 14) by the spectral centroid $BR(n)$ (Ehresman & Wessel, 1978). The analysis of variance (ANOVA) concerning the $BR(n)$ values indicates significant differences between the expressive intentions ($F(6,105) = 11.3, p < 0.001$). Fig. 15 represents BR average values measured for each performance. They demonstrate how the violinist used tone-colour to differentiate the soft, dark and light performances from the bright, hard and heavy ones. The soft and dark versions are those with the lowest values ($BR < 3$).

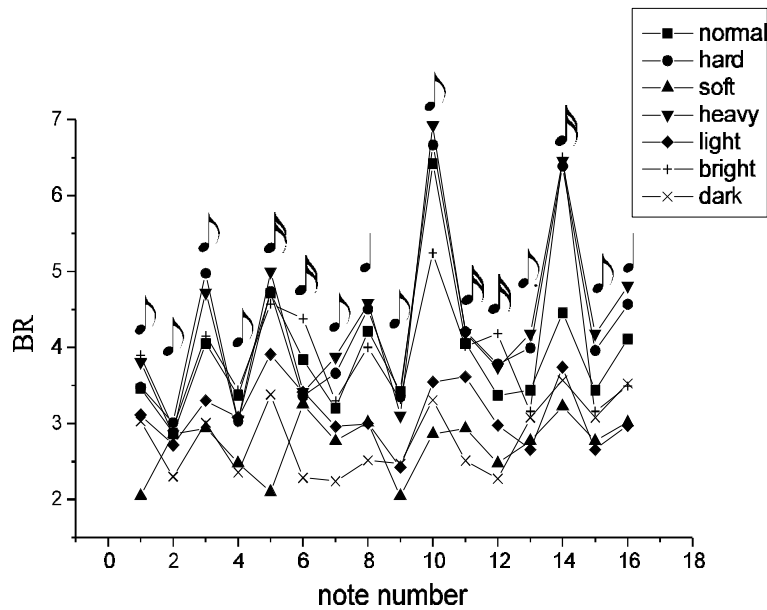


Fig. 14: Brightness in the seven performances.

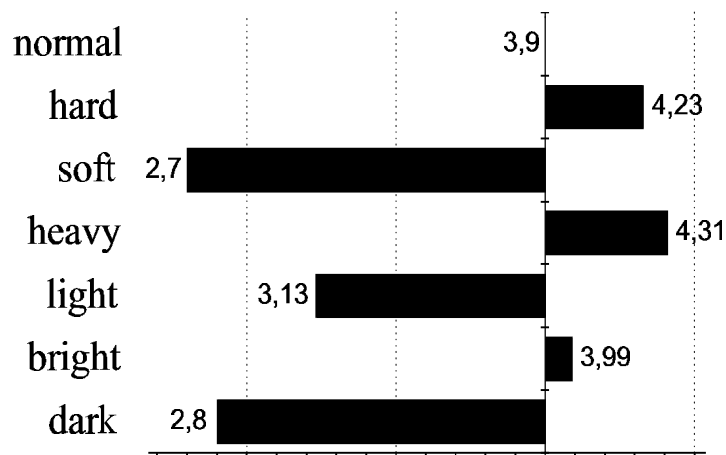


Fig. 15: Mean values of the parameter BR in the seven performances.

Apart from the differences due to expressive intentions, all performances exhibit some behaviors in common, indicating that they mainly depend upon the score. Particularly clear are the peaks in the BR parameter reached at 10 and 14 notes. Consulting Fig. 1, it can be seen that the pitch of both of these notes is A4, corresponding to the pitch of the 2nd violin string. It is therefore possible to attribute these peaks to technical reasons: the violinist played these tones on the open

string. When a string is stopped by a finger, the higher harmonics are deadened relative to an open string. Therefore, the open string has a higher value of BR.

Vibrato

There is a frequency vibrato in each of the seven performances, but not in each note. Generally, the notes with duration $DR > 250$ ms have vibrato. This limit, below which there is no vibrato, is clearly due to technical reasons: a vibrato cycle, of frequency about 5 Hz, lasts at least 200 ms. The number of vibrato notes (NV) is shown in Tab. 1. Interpretation of these data requires some observations. In the performance of the bright piece, the vibrato seldom appears (only in 4 notes). This fact does not depend on the violinist's expressive choices; it is a consequence of the fact that in the bright piece there are only five notes with duration $DR > 250$ ms. The rest are too short for vibrato. The situation is different in the other pieces where there are at least eleven notes with $DR > 250$ ms. It is interesting to notice that the normal piece contains only five notes with vibrato compared to nine in the soft and dark pieces.

In order to obtain the vibrato amplitude (VA, Fig. 16) and vibrato rate (VR, Fig. 17), the vibrato of each note was modelled with a sinusoid ($A + B \sin(\omega t)$). The parameters A, B, and ω were estimated by fitting the generated sinusoid to the measured pitch/time profile in the least squares sense. The vibrato amplitude is expressed through the ratio between $\Delta f = 2B$ and $f_0 = A$.

Tab. 1 shows the average values of vibrato rate (VR) and vibrato amplitude (VA). The averages were figured only on the notes with vibrato in every performance. It can be seen that the bright piece was played with a wide and fast vibrato. The normal piece, instead, has a small amplitude value. The ANOVA test concerning the vibrato rate indicates significant differences between the expressive intentions ($F(6,14) = 8.83, p < 0.001$), while differences in vibrato amplitude are not significant.

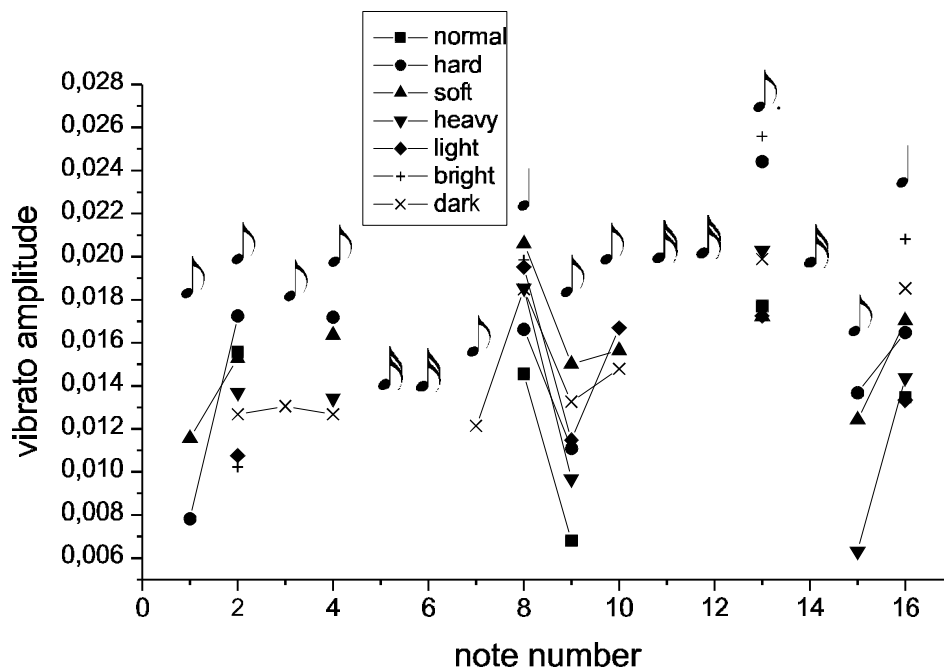


Fig. 16 vibrato amplitude in the seven performances

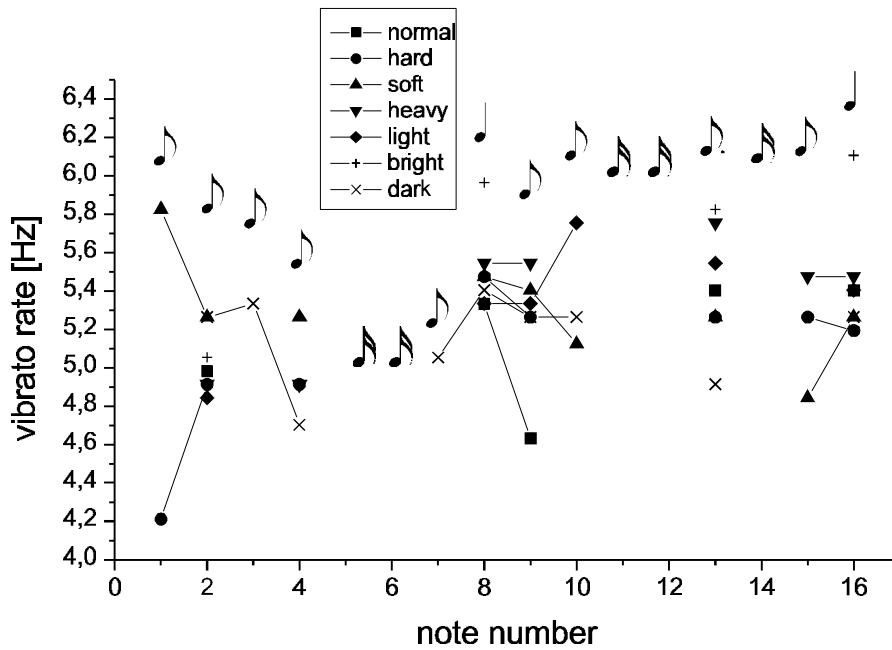


Fig. 17 vibrato rate in the seven performances

	N	Ha	S	He	L	B	D
NV	5	8	9	7	6	4	9
VR [Hz]	5.38	5.31	5.33	5.59	5.43	5.97	5.19
VA $\langle \Delta f \rangle / f_0$	0.015	0.019	0.018	0.018	0.017	0.022	0.019

Tab. 1: vibrato parameters in the seven performances.

CONTROL THROUGH SYNTHESIS

The measurement of the main acoustic parameters generated a group of characteristic values for each expressive intention. Each performance, however, is extremely rich in nuances and deviations. The choice of the parameters to be measured requires a considerable reduction of the information revealed in the acoustic signal, but presents at least two difficulties. First, it is necessary to make sure that the parameters are sufficient to represent the differences among several expressive intentions. Secondly, we must be sure that the procedures to measure the parameters are correct.

The analysis-by-synthesis method is an effective means to address these difficulties. This method consists of creating synthetic performances where the values of the main acoustic parameters are equal to the values previously measured. If such synthetic performances maintain the expressive characteristics of the original performances, it can be assumed that the simplifications and approximations made are acceptable, at least at the perceptual level. When expressive characteristics are lost, a comparison between the synthesis and the original allows the analysis to be refined. The application of this method requires the choice of an appropriate synthesis algorithm. Among the available ones, physical model synthesis seems to provide some considerable advantages. In fact, it allows performance control, not only event by event as in MIDI synthesis, but also at the phrase level. Moreover, the parameters controlled (e.g. speed of the bow, pressure on strings, etc.), allow an immediate reference to the expressive gestures of the musician. On the other hand, the measured parameters refer to the audio signal characteristics. In

order to provide the envelope of note amplitude with the desired shape (in terms of DRA, MA, MD etc.), it is necessary to have a deep knowledge of the IN/OUT map of the synthesizer used. In this research, the Yamaha VL1 physical model synthesizer (Yamaha Corporation, 1993) was used. Note-on and note-off information was directly inferred from IOI(n) and DR(n) measures, so as to reproduce timing and legato characteristics. Brightness (BR(n)) values, conveniently scaled so as to take into account MIDI coding, were used to control the “pressure on string” parameter in the VL1. The amplitude and velocity of the frequency vibrato were controlled, following the measured values, via corresponding vibrato parameters in the VL1. The amplitude envelopes of each note were controlled through the “bow speed” parameter. The “bow speed” is controlled by three straight lines (attack, sustain, and decay) for each note. The VL1 response to test signals (various width and slope steps and ramps) was measured so as to obtain an IN/OUT map. The data thus obtained were used to calculate, from the values of the main parameters measured, the inputs to control the bow speed of VL1. The function controlling the bow speed was set to zero only at the times of bow direction changes. In this way, it is possible to play multiple notes on the same bowing, reproducing the phrasing of a real violin. Some corrections, strictly connected to the synthesis algorithm, were needed for the generation of performances. To avoid a lowering of oscillation when the bow speed parameter value drops under a certain value, it was necessary to reduce the instrument losses at these critical points. This happens especially in the final notes of the two half phrases (8 and 16), which have a very slow decay. Decreasing the losses also causes a variation in pitch, which needs to be compensated by controlling the VL1 instrument tuning parameter.

Seven different performances were effected which, notwithstanding the timbral poverty of the VL1 violin and some defects in intonation due to imperfect tuning compensation, convey the expressive intention (hard, soft, etc.) of the originals. It is possible to listen to these syntheses in the sound examples no. 8-14 of the electronic appendix.

DISCUSSION

We saw how the sensorial adjectives allowed the accomplishment of seven performances with different characteristics. By means of a small number of parameters it was possible to measure and study such differences. Synthesis showed that the chosen parameters are enough to distinctively characterize the seven performances. We shall discuss first the average behavior of the values measured, comparing them to other studies, and then the note-by-note behavior.

Mean-value analysis

	N	Ha	S	He	L	B	D
MM		high		low	high	high	
L				high	low	low	
DRA		low			high	low	low
BR		high	low	high	low		low
UDR			low		high		
MD-MA			high		low		high
VR						high	low

Tab. 2: Behavior of statistically significant parameters on varying expressive intentions in violin performances

	N	Ha	S	He	L	B	D
MM			low	low	high	high	
L	high			low	low	low	
DRA		low	high	low		low	
BR		high	low	high	low		

Tab. 3: Behavior of parameters on varying expressive intentions in clarinet performances (Canazza et al., 1997a)

	N	Ha	S	He	L	B	D
MM				low	high	high	low
L	low		high	high	low	low	high
Key-velocity		high	low	high	low	high	

Tab. 4: Behavior of parameters on varying expressive intentions in piano performances (Battel & Fimbianti, 1997)

The data presented in the previous paragraphs suggests some considerations. The evolution of expressive deviations within each piece creates some profiles that have been shown in Fig. 5, 6, 8, 10, 12, 14, 16 and 17. Considering the various parameters one by one, it is clear that they show similar features in each of the seven performances (see double curves in Fig. 6 and 8). As all performances were obtained from the same score, it is possible to associate this behaviour to characteristics of musical structure. Independent of the expressive intention, therefore, the musical structure works as a source of deviations, inducing the musician to accomplish his performance so that the listener can easily understand it.

It was also noted that, upon changing the expressive intentions, the musician introduced systematic deviations. These deviations can be seen clearly in Fig. 6 and 8, where this effect corresponds to a curve translation along the axis of ordinates. Tab. 2 summarizes the tendencies of the statistically significant parameters on varying the expressive intentions. Each performance

shows peculiarities that makes it different from the others. But as we have already said, the musician has at his disposal more than a possibility to communicate the same content. Strategies of performance can therefore change in a significant manner. Besides, different musical instruments have different expressive means, which inevitably influence the musician's choices. In spite of this, comparison with previous studies shows some similarities in the way expressive intentions are transmitted.

Canazza et al. (1997a) and (Battel & Fimbianti, 1997) carried out a study on some performances clarinet and piano, respectively, inspired by the same sensorial adjectives used in this study, but with a different score. Not all the results can be compared as some of the parameters measured were defined differently in the two studies. Besides, quantitative comparisons are sometimes not very significant as the absolute values of parameters depend on the technical characteristics of the instrument used. Nevertheless, it is possible to make a qualitative comparison at least as far as the mean tempo (MM), legato (L), note attack time (DRA) and brightness (BR) of spectrum are concerned. Our main aim in this study is to determine, at least at a general level, those common strategies used by musicians to communicate their expressive intentions. Table 3 show the tendencies of parameters measured in the clarinet performances and Table 4 shows the qualitative results of parameters measured in piano performances. In the second case, it was only possible to compare MM and Legato parameters, considering the different technical characteristics of piano. It can be seen how, notwithstanding some differences, the pieces referring to the various expressive intentions have a similar behaviour in the different experiments. For instance, the adjective bright induced the musicians to perform their piece with a quicker metronome, a lesser legato, and a shorter attack time. In the piano, in fact, a high key velocity means a quicker attack. The main differences among the experiments have to do with a different choice in the expressive resources used, but not with a different use of these resources. In the soft version, for instance, the clarinet performer played with the values of the MM (low), DRA (high) and BR (low) parameters significantly different from the other versions. The violinist played in the same way as far as the BR (low) parameter is concerned, but unlike the clarinetist, he modified the MD-MA (high) and UDR (low) parameters. It is worth noting that a high MD-MA value means an amplitude profile slowly raising, while a high legato value in piano, together with a low key velocity, leads to an equivalent qualitative result. The only conflicting result regards the heavy piece performed by the clarinet player with a different use of the parameter L. In this case, it seems that the clarinetist used a quick note attack time and a slow metronome, causing in the listeners a sense of heavy locomotion; but the violinist and the pianist tried to communicate a sense of effort in moving things.

Perceptual studies proved how, generally speaking, it can be possible to correlate the listeners' main appraisal categories and the acoustic parameters which better characterize expressive intentions (Canazza et al., 1997b, 1997c, 1997d). Analyses carried out on the clarinet performances previously set forth (Canazza et al., 1997b) resulted in a two-dimensional semantic space containing the same sensorial adjectives used in this study. The two axes of this space were correlated to some acoustic parameters, determinants of a small number of expressive styles particularly evident in listening. The first axis was associated to the MM parameter, the second to the DRA parameter. The first axis also seemed to be correlated to a parameter that we will term climax emphasis (principal/average accent), used in the clarinet experiments to highlight the principal note of the phrase. This finding was not confirmed in the present analysis, but it is interesting to observe that the Metronome and the DRA arrange the violin performances in the same order observed in the perceptual studies, showing a possible generalization of these stylistic choices. More comparisons are possible with works analyzing the performers' expressive intentions but using a different set of adjectives. Gabrielsson and co-workers (Gabrielsson &

Juslin, 1996; Lindstrom, 1992; Juslin, 1993) carried out analyses on some flute, violin, electric guitar and singing voice recordings inspired by the adjectives happy, sad, angry, fearful, tender, solemn, no expression and natural. The happy version was performed with a quick time, a low legato value, a short attack time, a bright tone-color, a greater contrast in dotted notes and a high vibrato rate. Comparing it with our experiments, it is possible to say that the violinist used the same strategies to perform the bright piece. This result can be read as a consequence of the semantic affinity of the two adjectives.

Analysis note-by-note

Pay particular attention to the discussion of the analysis of note-by-note deviations in the violin performances. We have already seen that expressive intentions give rise to variations in performance parameters. Note-by-note analysis highlights how not every note is subjected to these variations. We shall particularly examine the legato (L) parameter, illustrated in Fig. 6, which clearly shows these peculiarities; note-by-note analysis illustrates that some notes substantially differ from the average behavior. This fact is especially evident in notes 5, 6, 8, 10, 11, 12, 13 and 14 of Fig. 6.

The data suggest two causes. The first is technical in nature and is connected to the qualities of the instrument used. The measured acoustic parameters cannot take on completely arbitrary values. In monophonic instruments, for instance, the L parameter cannot have values above 100 because a note starts only when the previous note has ended. But in the piano, the temporal superimposition of two notes is technically possible, and therefore such a limit does not exist. A further limitation regards note length (DR) that cannot be less than a certain value if the sound is to be perceived clearly. The second cause is based on the fact that not every note has the same importance within the musical structure. Some notes draw more attention and help the segmentation and the hierarchical organization of the piece itself (Drake & Palmer, 1993). The performer tends to emphasize such notes to make the piece easier to understand.

It would, then, be rather interesting to try to determine whether the peculiarities observed in the previous paragraphs on notes 5, 6, 8, 10, 11, 12, 13 and 14 are due to technical reasons or to expressive ones. Looking at the score illustrated in Fig. 1, it can be seen that notes 5, 6, 11, 12 and 14 are all sixteenth with an $IOI(n)$ duration of about 200 ms. It might be hypothesized that there is a technical limit to these notes: a low legato value on these notes, in fact, would imply a very short DR duration: this is the reason why they were played with a high Legato value in every performance. Furthermore, the score itself indicates that notes 5 and 11 should be tied to the successive ones.

A different question arises, however, with reference to notes 10, 13 and 8. It might be said that these notes have an important role within the musical structure of the phrase. Literature provides for a great number of theories allowing a structural analysis of a score (Lerdahl & Jackendoff, 1983; Narmour, 1990 etc.). The different types of analysis often privilege a particular aspect or component of the musical discourse (harmonic, melodic, rhythmic etc.). For this reason it is often necessary to support and compare different kinds of analysis in order to get more complete information. This paper illustrates the results derived from analysis of the accent structures, of harmonic charge and tonal tension. This choice is not, however, binding. The definition of accents after Drake & Palmer (1993), was applied to Corelli's score accents. In this sense, "accents" means those rhythmic-melodic events which most attract the listener's attention. These are very closely linked to the score and play an important role in the segmentation of the melody and in its structural hierarchy. Fig. 18 shows the structure of the accents that have been identified. The greatest number of accents fall on notes 8, 10, 13 and 16, and they are also those

which change least due to variation in expressive intentions. It is therefore seen that notes 8, 10 and 13 play an important role within the phase, a role that is indispensable for the transmission of the musical structure. The performer who wants to communicate this structure to the listener tends to call attention to these notes using the resources at his disposal. If these were played staccato as, for example, the bright piece would call for, then the understanding of the musical phrase would be compromised.

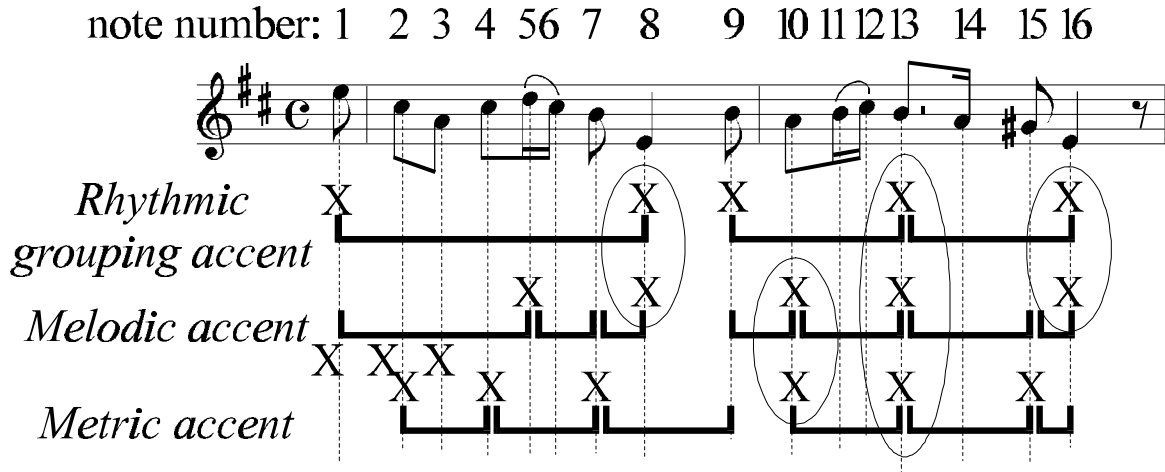


Fig. 18: Accent structures (Drake & Palmer, 1993) of the score.

A comparison with harmonic charge, as defined by Sundberg and co-worker (Friberg et al., 1991), in Fig. 19, and with Lerdahl's analysis of tonal tension (Lerdahl & Jackendoff, 1983) in Fig. 20, would seem to confirm the hypothesis that structural concerns constrain the deviations implied by expressive intention. Fig. 19 and Fig. 20 show that note 13 has a higher value (note 14 is only a passing tone). It might be said, then, that according to the tonal tension and the harmonic charge, the violinist stressed note 13, playing it legato and with a low DRA value in every performance.

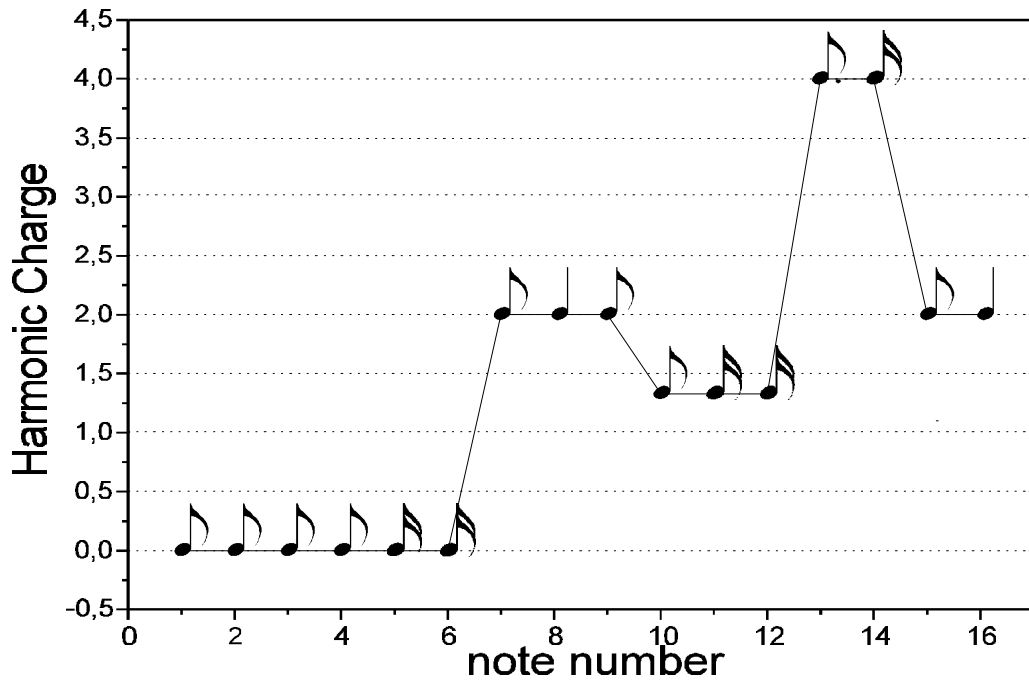


Fig. 19: Harmonic charge (Friberg et al., 1990) of the score of Fig. 1.

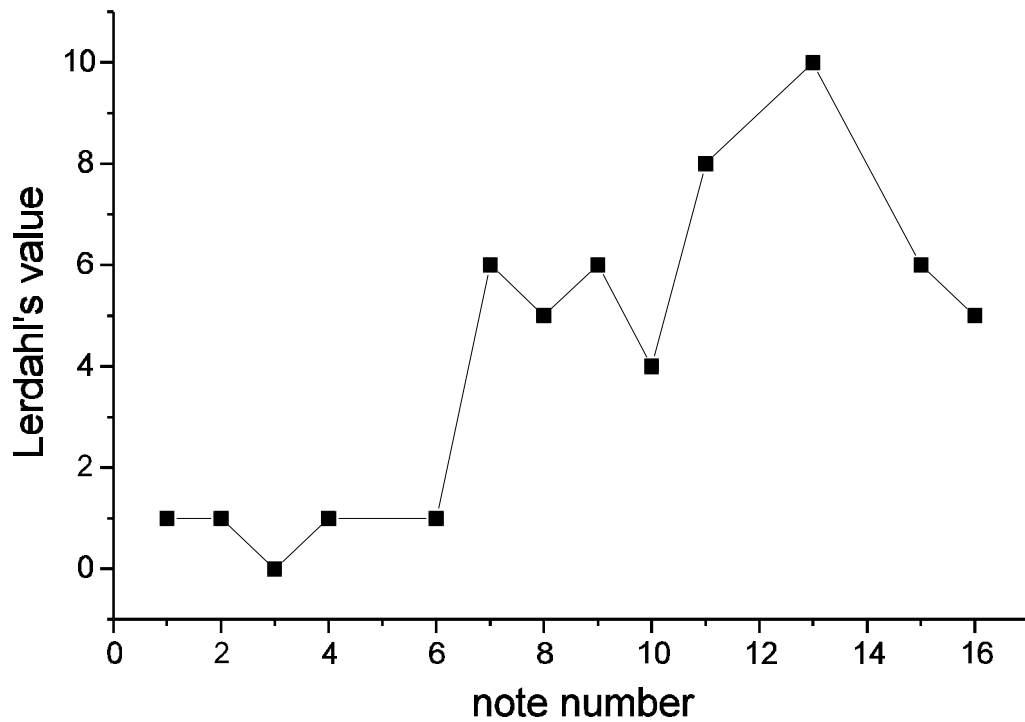


Fig. 20: Lerdahl's analysis of the slightly reduced version of the score in Fig. 1.

Notes 8 and 16 are the last ones of the two half-phrases; therefore, they have a relaxation function. To point out their peculiar function, in all the performances the musician played them

legato and preceded by a short pause. As a consequence, notes 7 and 15 will be slightly staccato. These results agree with those of Drake & Palmer (1993) who, analyzing the performances of ten pianists, found that in a rhythmic grouping, the musicians played the next-to-last event staccato. Note 10 has a support and tension function: it is in fact the resolution of an avoided cadence V-VI, it is placed on the first measure beat, and from a melodic point of view, the pitch A is the tonic of the basic tonality. In order to emphasize it, this note was played rather tied and with a reduced attack in all the performances.

Note-by-note analysis points out that the performer introduces some deviations that depend on the score. Different expressive intentions produce further deviations of certain parameters. Such deviations are not made to the same extent in each note because the score sets technical and structural limits that prevent the musical phrase from being distorted.

CONCLUSIONS

Sound analysis allowed us to identify variations in acoustic parameters as a result of varying expressive intentions. It was possible to identify two distinct expressive sources. The first refers to the musical structure of the melody, its division into phrases, and the continuous alternation of tension and relaxation points. The second depends on those expressive intentions the musician wants to convey to the listeners. Note-by-note analysis allowed us to identify technical and structural constraints in the deviations depending on the performer's expressive intentions. The parameters measured were used to control a physical model synthesizer and render different expressive intentions. The comparison of synthesized performances with originals allowed us to refine the analysis and to validate the relevant parameters. The results seem to confirm the hypothesis that various performances of the same musical piece have their own acoustic characteristics that can, in reality, be measured.

REFERENCES

- Askenfelt, A. (1986). Measurement of bow motion and bow force in violin playing. *Journal of Acoustical Society of America*, 80, 1007-1015.
- Battel G.U., & Fimbianti R. (1997). Analysis of expressive intentions in pianistic performances. In *Proceedings of the Int. Kansei Workshop 1997* (pp. 128-133). Genova: Associazione di Informatica Musicale Italiana.
- Bresin, R., De Poli, G., & Ghetta, R. (1995). A Fuzzy Approach to Performance Rules. In *Proceedings of XI CIM Colloquium on Musical Informatics* (pp. 163-167). Bologna: Associazione di Informatica Musicale Italiana.
- Bresin, R. & Vecchio, C. (1995). Neural Network plays Schumann. Analysis and synthesis of the performing action of a real pianist. In *Proceedings of the KTH Symposium on Grammars for Music Performance* (pp.5-14). Stockholm: KTH.
- Canazza, S., De Poli, G., Rinaldin, S., & Vidolin, A. (1997a). Sonological analysis of clarinet expressivity. In M. Leman (ed), *Music, gestalt, and computing. Studies in cognitive and systematic musicology* (pp. 431-440) Berlin, Heidelberg: Springer-Verlag.
- Canazza, S., De Poli, G., & Vidolin, A. (1997b). Perceptual analysis of the musical expressive intention in a clarinet performance. In M. Leman (ed), *Music, gestalt, and computing. Studies in cognitive and systematic musicology* (pp. 441-450) Berlin, Heidelberg: Springer-Verlag.
- Canazza, S., De Poli, G., Roda', A., & Vidolin, A. (1997c). Analysis and synthesis of expressive intentions in musical performance. In *Proceedings of the International Computer Music Conference 1997* (pp. 113-120). Tesseloniki: International Computer Music Association.
- Canazza, S., & Orio, N. (1997d). How are the players ideas perceived by listeners: analysis of "How high the moon" theme. In *Proceedings of the Int. Kansei Workshop 1997* (pp. 128-133). Genova: Associazione di Informatica Musicale Italiana.
- Clarke, E. F. (1987). Level of structure in the organization of musical time. *Contemporary Music Review*, 2, 211-38.
- Clynes, M. (1986). When time is music. In JR Evans, M Clynes (eds.), *Rhythm in Psychological, linguistic and musical processes* (pp. 169-224). Springfield, IL: Thomas
- Clynes, M. (1995). Microstructural musical linguistics: composers' pulses are liked most by the best musicians. *Cognition*, 55, 269-310.
- Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, 10, 343-378.
- Ehresman, D., & Wessel, D.L. (1978). Perception of timbral analogies. *IRCAM Report 13/78*. Paris, France: Centre Georges Pompidou.
- Friberg, A. (1991). Generative Rules for music performance: a formal description of a rule system. *Computer Music Journal*, 15(2), 56-71.
- Friberg, A., Frydén, L., Bodin, L. G., & Sundberg, J. (1991). Performance rules for computer controlled contemporary keyboard music. *Computer Music Journal*, 15(2), 49-55.
- Gabrielsson, A. (1995). Expressive intention and performance. In R. Steinberg (ed.) *Music and the mind machine*, (pp. 37-47), Berlin, Heidelberg, New York: Springer-Verlag.
- Gabrielsson, A. & Juslin, P. N. (1996). Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychology of music*, 24, 68-91.
- Gabrielsson, A. (1997). Music Performance. *The psychology of music*, In D. Deutsch (Ed.) *The psychology of Music*, 2nd. ed. New York: Academic.
- Juslin, P. N. (1993). The influence of expressive intention on electric guitar performance. *Unpublished thesis*. Department of Psychology. Uppsala University.
- Kendall, R. A., & Carterette, E. C. (1990). The communication of musical expression. *Music Perception*, 8, 129-164.
- Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge: The MIT Press.
- Lerdahl, F. (1996). Calculating tonal tension. *Music Perception*, 13(3), 319-363.
- Lindstrom, E. (1992). 5 x Oh my darling Clementine. The influence of expressive intention on music performance. *Unpublished thesis*. Department of Psychology. Uppsala University.

- Narmour, E. 1990. *The analysis and cognition of basic melodic structures: the implication-realization model*. Chicago: Univ. Chicago Press.
- Palmer, C. (1996). Anatomy of a performance: sources of musical expression. *Music Perception*, 13(3), 433-453.
- Palmer, C. (1997). Music Performance. *Annual Review Psychol.*, 48, 115-38
- Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *Journal of Acoustical Society of America*, 88, 622-641.
- Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's Träumerei. *Journal of Acoustic Society of America*, 92(11), 2546-2568.
- Scherer, K. R. (1991). Emotion expression in speech and music. In J. Sundberg (Ed.) *Music language speech and brain*. (pp. 146-156). London: MacMillan Press.
- Senju, M., & Ohgushi, K. (1987). How are the player's ideas conveyed to the audience?. *Music Perception*, 4, 311-324.
- Sundberg, J., Askenfelt, A., & Fryden, L. (1983). Musical Performance: a Synthesis-by-Rule Approach. *Computer Music Journal*, 7 (1), 37-43.
- Sundberg, J. & Verrillo V. (1980). On the anatomy of the retard: a study of timing in music. *Journal of Acoustic Society of America*. 68(3), 772-779.
- Todd, N. (1985). A model of expressive timing in tonal music. *Music Perception*, 3, 33-58.
- Todd, N. (1992). The dynamics of dynamics: a model of musical expression. *Journal of Acoustical Society of America*, 91, 3540-3550.
- Todd, N. (1995). The kinematics of musical expression. *Journal of Acoustical Society of America*, 97, 1940-1949.
- Yamaha Corporation (1993). *Virtual acoustic synthesizer VLI, Owner's Manual*. Hamamatsu: Yamaha Corporation.