

Investigating the effects of a postural constraint on the cellists' bowing movement and timbral quality

Jocelyn Rozé¹, Richard Kronland-Martinet¹, Mitsuko Aramaki¹, Christophe Bourdin² and Sølvi Ystad¹

¹ LMA (Laboratoire de Mécanique et d'Acoustique), CNRS, UPR 7051, Aix-Marseille Université, Ecole Centrale, 13009, Marseille, France

² ISM (Institut des Sciences du Mouvement), CNRS, UMR 7287, Aix-Marseille Université, 13288, Marseille, France
`roze@lma.cnrs-mrs.fr`

Abstract. While playing expressively, cellists tend to produce postural movements, which seem to be part of their musical discourse. This article describes how their instrumental bowing gestures and timbral features of the produced sounds may be affected when constraining these postural (or ancillary) movements. We focus here on a specific acoustic timbre alteration qualified as *harshness* in the constrained condition. A method based on Canonical Correlation Analysis (CCA) is used to extract the correlations between the bowing displacement and the sound rendition with and without postural constraint among several cellists. Then a detailed investigation of the covariation between gestural and sound data for the duration of the note is carried out, using Functional Data Analysis (FDA) techniques. Results reveal interesting effects of the postural constraint on the coupling patterns between the bowing movement and the spectro-temporal acoustical features.

Keywords: Cellist, ancillary/postural gestures, gesture-sound relationship, acoustical features, musical expressivity, functional data analysis

1 Introduction

Musical expressiveness of instrumentalists is the result of interactions within a multimodal context, in which the perceived acoustic features turn out to be embedded into continuous gesture processes [10]. By gestures, we refer to those directly responsible of the sound production, but also to so-called ancillary gestures, in particular the performers' postural movements, which may form an integral part of their subtle expressive variations [18]. In the present study, we investigate expressivity related to cello playing and focus on the influence of the musicians' bowing gesture on the timbre quality in normal playing, and in posturally constrained situations.

This study falls within a more global experimental context of sound-gesture relationship for the cello players in musical performance situations [14]. Related

works explored the instrumentalists' postural movements for the clarinet [6], the piano [16], the harp [3], or the violin [17]. It was shown that these ancillary displacements turn out to be part of the musician's motor program, and in the case of the cellist, their limitation seemed to induce an impoverishment of the expressiveness in terms of rhythmic deviations and timbre color variations. In this paper, we assess the timbral degradations which may occur on certain notes while constraining the cellist's posture.

The postural constraint should also give rise to some alterations in the bowing gesture execution, and our aim here consists in highlighting them by assessing the covarying effects with the sound characteristics. Acoustical studies carried out on the physics of the violin [15] and the cello [7] revealed the bowing pressure and velocity as the main parameters of timbral control. Furthermore, the correlations with the spectral energy distribution (referred as the spectral centroid) and its variations over time, allowed to better understand the role played by these physical parameters with respect to the perceived *brightness* [5] and musical tense [4].

After presenting the experimental methodology, we describe the sound-gesture descriptors used in the study. Analysis of type Canonical Correlation (CCA) [2] and Functional Principal Component (FPCA) [13, 1] are then carried out to investigate how these sound and gesture descriptors mutually covary, while applying a postural constraint.

2 Methodology

2.1 Experimental conditions

Seven cellists participated in the study and were asked to play a specifically designed score in the most expressive way, while being subjected to two kinds of postural conditions. Fig. 1 illustrates these two conditions. The first one was a natural condition, in which cellists were asked to play naturally as in a performance context. The second one was a physically fully constrained situation, in which the torso was attached to the back of the chair by a 5-point safety race harness and a neck collar adjusted to limit the head movements. We are aware that these kinds of physical constraints raise an epistemological issue, since bowing or acoustic alterations may result from other factors than the only limitation of postural movements, such as physical, psychological discomfort, estrangement from the concert situation... All the selected cellists were professional or very experimented, to ensure that no kind of technical weaknesses would potentially result in a lack of expressivity. These two experimental conditions are part of a larger experiment thoroughly described in [14].

2.2 Data acquisition and pre-analysis

Cellists' corporeal movements were recorded by a VICON motion capture system, composed of 8 infrared cameras acquiring data at a frame rate of 125 Hz.

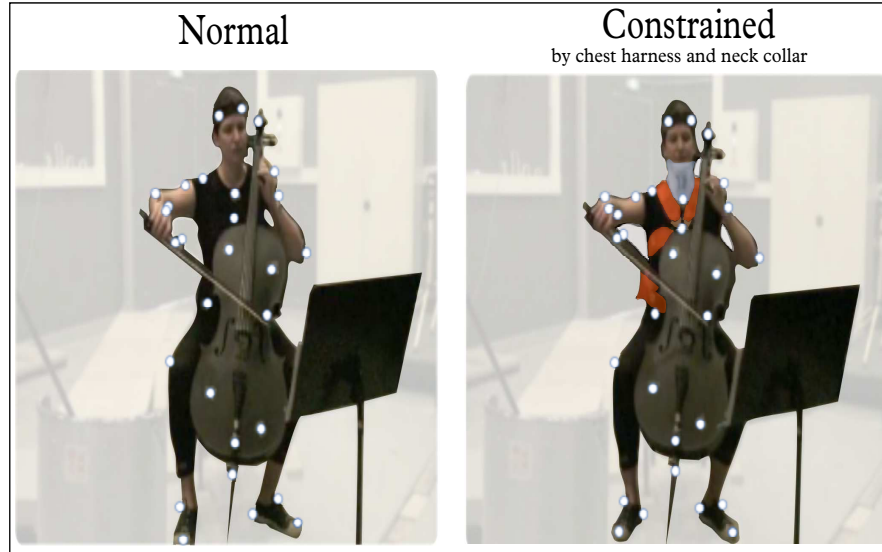


Fig. 1. The 2 types of experimental postural condition *Normal* and *Constrained*

The system tracked the 3D kinematical displacements of a set of sensors placed on the performer body, the cello and the bow. In this paper, we are interested in the bow displacements, and therefore focus on a bow marker located at the bow frog (close to the musician's right hand). Audio data were recorded at a 44.1 kHz sampling rate by a DPA 4096 microphone placed under the cello bridge and connected to a MOTU interface. The gestural and audio streams were manually synchronized by a clap at the beginning of each recording.

When analyzing acoustic data collected from the constrained condition, a specific note of the score which sounded poor, shrill and quaver as a beginners sound frequently emerged for all the cellists. Following the terminology used by one of the performers, we qualified this as a *harshness* phenomenon produced by acoustic timbre alterations. When this *harsh* feature was perceived, we extracted the note, as well as its nice (or *round*) counterpart produced in the normal condition for the same cellist. This extraction process was carefully carried out by a pitch-tracking algorithm adapted from the MIR toolbox[9]. A corpus of 8 pairs of *round/harsh* notes were hereby extracted among all the participating cellists. The corresponding sequences of bowing displacements were segmented from the motion capture stream, by using the temporal landmarks of the note in the audio stream.

To further investigate potential functional correlations between sound and bowing gesture (in particular for analysis presented in section 5), the computation of acoustic descriptors were adapted to the bowing gesture data. In practice, the frame rate of acquisition within the audio device (44.1 KHz) was much higher than that of the motion capture system (125 Hz). To synchronize the com-

putation of audio descriptors on the motion capture stream, an efficient mean consisted in splitting the audio stream in frames overlapped by a motion capture time step, i.e 8 ms (1/125Hz). The frame duration was chosen ten times higher than the hop size, i.e 80 ms, to allow a sufficient frequency resolution (12 Hz). We applied this technique for all the acoustic descriptors.

3 Observing the effects of the postural constraint

The bowing gesture and the perceived sound were explored through suitable signal descriptors. Simple statistic tests were then performed to assess the influence of the cellists' posture on these descriptors.

3.1 Bowing gesture descriptor

We focus here on a compact gestural feature which could be related to the bow displacements. In the reference frame of the motion capture system, each marker is associated to a triplet of coordinates (x, y, z) providing its position at each frame. By derivation, we can get the spatial coordinates of the velocity vector (v_x, v_y, v_z) . More generally, we worked with the absolute velocity of the bow inferred from the coordinates of the bow frog marker:

$$VEL_{bow} = \sqrt{(v_x^2 + v_y^2 + v_z^2)} \quad (1)$$

This information could have been captured by simple accelerometers, but given that the experimental context primarily focused on the musician's postural aspects, we used the data collected by the motion capture system.

3.2 Acoustic descriptors

In this paper, we tried to find acoustic descriptors, which at best could reveal the relation between the signal and the quality alteration perceived between notes played in normal and constrained conditions. This acoustic *harshness* phenomenon might correspond to a degradation of the perceived timbre, i.e a change in the spectro-temporal features of the sound signal, for equal pitches and durations. Several timbre descriptors are potential candidates to suitably characterize such time-frequency transformations.

Temporal domain From a temporal viewpoint, a *harsh* note may differ from its *round* counterpart, by the way the energy rises during the onset of the sound. The kind of features that are likely to reflect this observation imply a prior extraction of the sounds temporal envelope, that for example can be obtained from the Root Mean Square (RMS) value [8] of each audio frame l composing the signal s :

$$Rms(l) = \sqrt{\frac{1}{N_w} \sum_{n=0}^{N_w-1} s^2(lN_{hop} + n)} \quad (0 \leq l \leq L-1) \quad (2)$$

where N_w is the frame length, and N_{hop} the hop size in samples.

Attack slope The classical descriptor Attack Time could have been adopted as a temporal descriptor, but the Attack Slope (ATS) was preferred in the present case to overcome energy differences between signals. ATS represents the temporal increase or average slope of the energy during the attack phase [11]:

$$ATS = \frac{PeakValue}{AT} \quad (3)$$

where AT is the Attack Time, i.e. the time that the RMS envelope takes to deploy from 10% to 90% of its maximal value $PeakValue$.

Spectral domain From a spectral viewpoint, the shrill nature of the sound produced in the constrained situation would suggest energy reinforcement in high frequencies. Given the harmonic nature of cello sounds, harmonic spectral descriptors are believed to characterize this spectral transformation. We chose a total number of 25 harmonics to compute them.

Harmonic spectral centroid We decided to focus on the Harmonic spectral centroid instead of the standard spectral centroid (SC), since the stochastic part of the signal seemed to be negligible with regards to the deterministic part. Hence, from the harmonic instantaneous features provided by subband decomposition, we computed the Harmonic spectral centroid (HSC(l)) to characterize the barycenter of the spectral energy distribution at each frame l . This descriptor is related to the perception of *brightness* in various acoustic studies on the violin [5]. HSC represents the mean value of HSC(l) [8] :

$$HSC = \frac{1}{L} \sum_{l=1}^{L-1} HSC(l) = \frac{1}{L} \sum_{l=1}^{L-1} \frac{\sum_{h=1}^H f_h(l) A_h(l)}{\sum_{h=1}^H A_h(l)} \quad (0 \leq l \leq L-1) \quad (4)$$

where $f_h(l)$ and $A_h(l)$ are respectively the frequency and the amplitude of the h^{th} harmonic in frame l .

Harmonic tristimulus ratio To characterize more finely the spectral energy transfer, which may occur from a *round* sound to its *harsh* equivalent, we computed the harmonic tristimulus [12] at each frame. This descriptor considers the energy distribution of harmonics in three frequency bands and measures the amount of spectral energy inside each band relatively to the total energy of harmonics. The first band contains the fundamental frequency, the second one the medium partials (2, 3, 4) and the last one higher order partials (5 and more). Three spectral coordinates are hereby obtained for each frame l , corresponding to spectral barycenter distribution within each band:

$$TR_1 = \frac{1}{L} \sum_{l=1}^{L-1} TR_1(l) = \frac{1}{L} \sum_{l=1}^{L-1} \frac{A_1(l)}{\sum_{h=1}^H A_h(l)} \quad (0 \leq l \leq L-1) \quad (5)$$

$$TR_2 = \frac{1}{L} \sum_{l=1}^{L-1} TR_2(l) = \frac{1}{L} \sum_{l=1}^{L-1} \frac{\sum_{h=2}^4 A_h(l)}{\sum_{h=1}^H A_h(l)} \quad (0 \leq l \leq L-1) \quad (6)$$

$$TR_3 = \frac{1}{L} \sum_{l=1}^{L-1} TR_3(l) = \frac{1}{L} \sum_{l=1}^{L-1} \frac{\sum_{h=5}^H A_h(l)}{\sum_{h=1}^H A_h(l)} \quad (0 \leq l \leq L-1) \quad (7)$$

where $A_h(l)$ is the amplitude of the h^{th} harmonic in frame l . From here, we designed a more compact ratio focusing on the spectral transfer feature, which should increase for energy transfers towards higher partials:

$$TRIratio = \frac{1}{L} \sum_{l=1}^{L-1} TRIratio(l) = \sum_{l=1}^{L-1} \frac{TR_3(l)}{TR_1(l) + TR_2(l)} \quad (8)$$

3.3 Validation of the descriptors

To assess if these four descriptors, i.e. the bowing gesture descriptor and the three acoustic descriptors, are affected by the postural constraint, we performed statistical tests for each one, on the basis of the 8 *round/harsh* data pairs. Fig. 2 presents the quartiles of the four descriptors between the two postural conditions. It was observed that in average, the postural constraint tended to reduce the bow velocity, while giving rise to a dual effect in the spectro-temporal acoustic features resulting in a decrease of the temporal attack slope coupled with an energy increase for high-frequency partials.

The relevance of each signal descriptor was evaluated by performing a simple paired two-tailed t-test, based on the null hypothesis that the means are the same between the normal and constrained conditions. Table 1 reports the results of these t-tests, which actually reveal that the null hypothesis can be significantly rejected and thus that the postural conditions can be discriminated for all signal descriptors. This signifies that the sound-bowing gesture relationship is significantly affected when the cellists are limited in their postural movements.

Table 1. Results of paired t-tests on the defined gestural descriptor and the three acoustic descriptors. The discrimination capacity between the normal and constrained groups of 8 data for each descriptor is given by the p-value : * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Descs	BOW VELOCITY	ATS	HSC	TRIRATIO
t(7)	2.77*	4.15**	-4.21**	-3.48*

4 Correlating bow gesture to sound features

In this part, we focus on the global relation that exists between cellists bowing gesture and the resulting sound features. This connection is explored by means of raw linear and canonical correlations techniques.

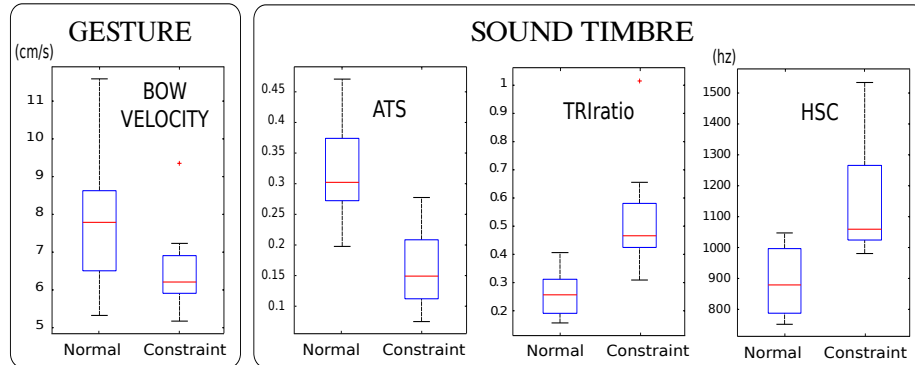


Fig. 2. Comparison of the mean gestural and acoustic features between the 2 types of postural conditions *Normal* and *Constraint*. The central marks are the medians, the edges of the boxes are the 25th and 75th percentiles

4.1 Raw linear correlations

Analysis. Assuming a linear relationship between the gestural and acoustic parameters, we computed the Pearson’s linear correlation coefficient of the bowing velocity vector with each acoustic vector. Each feature vector was composed of 16 mean data, corresponding to the 8 pairs of {normal/constrained} mean descriptor data.

Results. Raw linear correlations revealed that the bowing velocity was strongly correlated to the attack slope (ATS) of the temporal envelope ($r^2 = 0.6^*$). By contrast, spectral descriptors such as harmonic centroid (HSC) and tritrimulus ratio (TRlratio) surprisingly turned out to be weakly correlated to the bowing velocity ($r^2 = -0.11$ and $r^2 = -0.21$ respectively).

Discussion. Fig. 3 depicts the three graphs of linear correlation between the gestural variable and each acoustic variable. The interpretation of these graphs becomes interesting if we consider pairs of variables {normal/constraint}. The first graph (a) highlights the influence of the postural constraint on the bowing velocity and temporal attack slope. It reveals that once the cellists were deprived of their postural adjustments, they showed a global tendency to combine a decrease in bow velocity with a less sharp way of attacking the string, reflected by a decrease of attack slope. The second and third graph, (b) and (c) highlight the influence of the postural constraint on the bowing velocity and spectral descriptors. They present similar interesting tendencies, in spite of the weak raw correlations obtained in the results: The reduced velocity induced by the postural constraint causes an energy shift towards high frequency components. A closer examination of the canonical correlations might allow confirming this effect on the spectral variables.

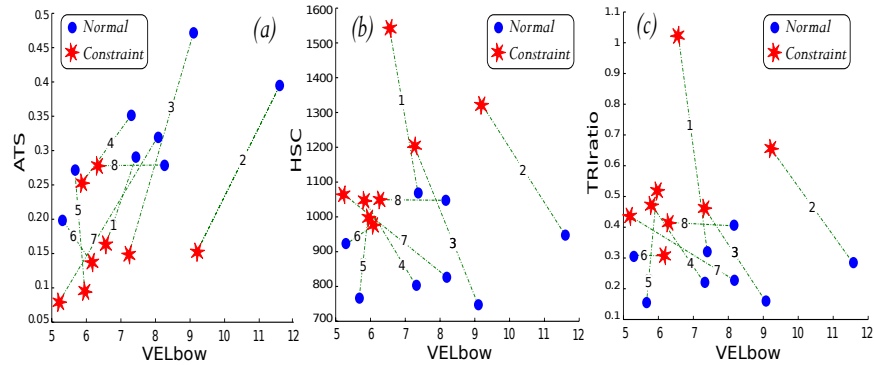


Fig. 3. Polar diagrams of the raw linear correlations obtained between the mean gestural bowing feature and the three mean acoustic ones over the 16 observations. Each {normal(N)/constraint(C)} pair is connected by a dotted line. The raw linear correlations are presented: (a) Between VELbow and ATS, (b) Between VELbow and HSC, (c) Between VELbow and TRIratio

4.2 Canonical correlations

Analysis. We performed a Canonical Correlation Analysis (CCA) to assess and quantify the nature of the interaction between the bowing gesture and all the identified acoustic features. It consisted in finding two sets of basis vectors, one for the gesture and the other for the acoustic descriptors, such that the correlations between the projections of the initial variables onto these basis vectors are mutually maximized. Compared with the previous ordinary correlations, this technique offers the advantage of being independent of the coordinate system in which the variables are described. It actually finds the coordinate system optimizing their representation. We provided the CCA with the previously used feature vectors of 16 mean data, but organized differently : the 16 mean bowing velocities were contained in a vector \mathbf{X} , and the 3 variables (ATS, HSC, TRIratio) of 16 mean acoustic data in a matrix \mathbf{Y} .

Results. The Canonical Correlation Analysis between mean gestural and mean acoustic data (\mathbf{X} and \mathbf{Y}) appeared to be highly significant ($r^2 = 0.74^*$). The analysis computed the canonical scores by projecting the initial variables \mathbf{X} and \mathbf{Y} on two matrices \mathbf{A} and \mathbf{B} , which maximize the canonical correlation $\text{corr}(\mathbf{XA}, \mathbf{YB})$. We used the canonical loadings \mathbf{A} and \mathbf{B} to express each variate $\mathbf{U} = \mathbf{XA}$ and $\mathbf{V} = \mathbf{YB}$ as a linear combination of the initial gestural and acoustic variables respectively. It thus led to the 2 following equations:

$$\begin{cases} \mathbf{U} = -0.5 \times \mathbf{VEL}_{\text{bow}} \\ \mathbf{V} = -10 \times \mathbf{ATS} - 0.01 \times \mathbf{HSC} + 5 \times \mathbf{TRIratio} \end{cases}$$

Fig. 4 presents the canonical scores (or variates \mathbf{U} and \mathbf{V}) resulting from the method.

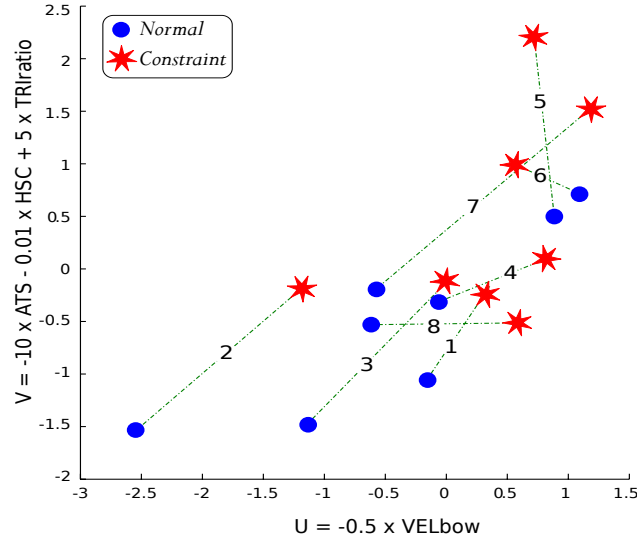


Fig. 4. Canonical variates of the CCA applied on the mean gestural and acoustic features. The variates \mathbf{U} and \mathbf{V} correspond to linear combinations of bowing gestures and acoustic features respectively. The canonical correlation between these two variates is $r^2 = 0.74^*$. Each {normal/constrained} pair has been represented by a dotted line.

Discussion. The canonical weights (or projection eigenvectors) \mathbf{A} and \mathbf{B} stand for the relative importance of each of the initial variables in the canonical relationships. We can hereby deduce from the first correlation equation that the gestural variable $\mathbf{VEL}_{\text{bow}}$ is negatively correlated to its variate \mathbf{U} ($\mathbf{A} = -0.5$). In the same manner, the second correlation equation indicates that the variate \mathbf{V} is negatively correlated with \mathbf{ATS} ($\mathbf{B}(1) = -10$), positively correlated with $\mathbf{TRIratio}$ ($\mathbf{B}(3) = 5$), and not correlated with \mathbf{HSC} ($\mathbf{B}(2) = -0.01$). By contrast with the raw linear correlation method, the CCA reveals a correlation with at least one spectral variable ($\mathbf{TRIratio}$).

If we now consider data pairs of postural conditions, as it's represented on Fig. 4, we can get an interesting interpretation of the role played by each variate in the modification of the sound-gesture relationship. Along the gestural variate \mathbf{U} , the constrained condition is rated higher as a whole compared to its normal counterpart, which indicates a global decrease of the gestural variable $\mathbf{VEL}_{\text{bow}}$. Along the acoustic variate \mathbf{V} , the constrained condition is also rated higher as a whole than its normal counterpart, which suggests a global dual effect of the two main acoustic correlated variables : A decrease of \mathbf{ATS} coupled to an increase of $\mathbf{TRIratio}$. This interpretation of the space built from the CCA variates,

reinforces the results already obtained in section 3. Furthermore, it is coherent with the results of Guettler [7], who demonstrated that the lowest bow speeds give the highest relative amplitudes for the upper partials.

5 Extracting functional covariates of the sound-gesture relationship

The previous statistic tools revealed interesting results for descriptors averaged over the whole duration of each note. However, the examined data are functional by nature and in this part, we assess if the previous findings might be confirmed in the functional domain. The modifications induced by the postural constraint should allow extracting functional covariates of the sound-gesture relationship.

5.1 Preliminary processing

By means of a canonical correlation analysis, we showed in section 4 that a gestural descriptor of bowing velocity (VELbow) and two dual acoustic descriptors - attack slope (ATS) and spectral energy transfer (TRIratio) - were suitable to model the effect of the postural constraint on the sound-bowing gesture relationship within the data corpus. Instead of focusing on their mean values as previously, we here consider the functional nature of these descriptors, in order to compare their evolution in time. Hereby, since ATS descriptor corresponds to a discrete value, we rather consider the RMS envelope in the functional temporal domain.

Even though the musicians were asked to play the score at a fixed tempo, all the notes composing the corpus presented slightly different durations, because of deviations induced from the constraint or expressive intentions of the players. These temporal variations prevent the sequences of descriptors from being directly compared. A preliminary step thus consisted in synchronizing the temporal sequences of the corpus by time-warping process. The 8 {normal/constraint} pairs of descriptor curves were fitted to the duration of the longest one, which measured 56 data points (i.e. 45 ms in the mocap frame rate). This led to 16 time-warped temporal sequences for each one of the three functional descriptors: VELbow(t), RMS(t), TRIratio(t).

5.2 Analysis

The 3 groups of 16 time-warped curves were processed by Functional Data Analysis (FDA) techniques [13], which consisted in modeling each time-serie as a linear combination of equally spaced 6-order B-spline basis functions. We chose a semi-sampled basis with respect to the total number of data points in each curve, i.e. a basis of 28 (56/2) B-spline functions, to keep a fine-grained definition of each curve and limit the inner smoothing FDA mechanism.

This functional spline-based representation of time-point data turned out to be particularly suited to analyze the sources of variability encompassed within

the sound-bowing gesture interaction. It was achieved by combining the FDA modelling to classical multivariate Principal Component Analysis (PCA), a technique known as Functional PCA (FPCA). We thus extracted the major modes of variability of this interaction by carrying out two *bivariate* FPCAs : The first one between the functional bow velocities $VELbow(t)$ and the temporal sound envelopes $RMS(t)$; The second one between the same bow velocities $VELbow(t)$ and the functional descriptor of high-frequency spectral distribution called the $TRIratio(t)$.

5.3 Results

The results of the two bivariate FPCAs are respectively presented Fig. 5 and Fig. 6. In each case, we focused on the first two principal components returned by the method, since they accounted for more than 90% of the overall variability among the bivariate set of 16 curves. A bivariate functional principal component was defined by a double vector of weight functions:

$$\begin{cases} \xi_m = (\xi_m^{VELbow}, \xi_m^{RMS}) & \text{for the first FPCA} \\ \eta_m = (\eta_m^{VELbow}, \eta_m^{TRIratio}) & \text{for the second FPCA} \end{cases}$$

ξ_m^{VELbow} denotes the principal m -variations of the bow velocity curves, relatively to ξ_m^{RMS} , the m -variations of temporal sound envelopes. η_m^{VELbow} denotes the principal m -variations of the bow velocity curves, relatively to $\eta_m^{TRIratio}$, the m -variations of high-frequency spectral distributions. The index m equals 1 or 2, since only two components are necessary to account for the principal variations.

For the first FPCA, we obtained two orthonormal bivariate eigenfunctions ξ_1 and ξ_2 , respectively accounting for 72% and 21% of the variations. The effects of these bivariate eigenfunctions ξ_1 and ξ_2 are represented in Fig. 5, by specific perturbations of the mean functional variables $VELbow(t)$ and $RMS(t)$. These perturbations reflect the fact of adding or subtracting each eigenfunction to the two mean curves, i.e. $VELbow(t) \pm \xi_m^{VELbow}(t)$ and $RMS(t) \pm \xi_m^{RMS}(t)$. Interestingly, we notice that the first eigenfunction reflects a *positive* correlation, visible as an overall amplitude shift of the two functional means. The second eigenfunction correlates a distortion in the timing of the mean velocity profile with an amplitude shift of the mean sound envelope. The bivariate effect of these eigenfunctions is synthesized in polar diagrams (Fig. 5.c1 and Fig. 5.c2), which report the position of the mean function values $(VELbow(t), RMS(t))$ by a dot in the (x, y) plane. Each point of the polar mean curve is linked to a line indicating the direction of the perturbation $(VELbow(t) + \xi_m^{VELbow}(t), RMS(t) + \xi_m^{RMS}(t))$.

For the second FPCA, we obtained two orthonormal bivariate eigenfunctions η_1 and η_2 , respectively accounting for 71% and 20% of the variations. The effects of these bivariate eigenfunctions η_1 and η_2 have been represented Fig. 6 by specific perturbations of the mean functional variables $VELbow(t)$ and $TRIratio(t)$, as previously. Interestingly, we notice that the first eigenfunction reflects a *negative* correlation, visible as an opposed amplitude shift of the

two functional means. The second eigenfunction correlates a distortion in the timing of the mean velocity profile with an amplitude distortion of the mean high-frequency spectral distribution. Polar diagrams of these two mean variables have also been reported with their eigenfunction perturbations (Fig. 6.c1 and Fig. 6.c2).

5.4 Discussion

Results revealed that across the two FPCAs, the behavior of the bow velocity eigenfunctions ξ_m^{VELbow} and η_m^{VELbow} remained consistent. Indeed, their major contribution, denoted ξ_1^{VELbow} and η_1^{VELbow} , reflected an overall amplitude increase of the mean bow velocity curve, according to an explained variance around 70% in both cases. Similarly, their minor contribution, denoted ξ_2^{VELbow} and η_2^{VELbow} , reflected an amplitude distortion with respect to the middle of the mean velocity profile, according to an explained variance around 20% in both cases. This minor contribution might be interpreted in terms of accelerations/decelerations of the mean bowing gesture. The common perturbations induced by these bow velocity eigenfunctions, may be viewed as a leveraging tool to interpret the joint effects on acoustic variables $RMS(t)$ and $TRIratio(t)$.

First, the major mode of variations (ξ_1, η_1) transversal to the variables, reveals that increasing the overall mean bow velocity, results in a slope rise of the mean sound temporal envelope (Fig. 5.b1), combined with a drop of the mean high-frequency spectral distribution (Fig. 6.b1). More specifically, the polar diagram $(VELbow(t), RMS(t))$ shows a strong linear increase of the interaction between these variables until the velocity peak. The direction being taken by the major eigenfunction perturbations $(\xi_1^{VELbow}, \xi_1^{RMS})$ reflects a positive covariation of these parameters over the whole duration of the note (Fig. 5.c1). Similarly, the polar diagram $(VELbow(t), TRIratio(t))$ shows a weakest linear decrease of the interaction between these variables until the velocity peak. The direction being taken by the major eigenfunction perturbations $(\eta_1^{VELbow}, \eta_1^{TRIratio})$ reflects a negative covariation of these parameters over the whole duration of the note (Fig. 6.c1). This coupled effect is coherent with the findings of section 4 and hence turns out to be a major trend of the sound-bowing gesture relationship.

Then, the minor mode of variations (ξ_2, η_2) transversal to the variables, reveals that accelerating the mean bowing gesture results in an overall increase of the mean sound temporal envelope (Fig. 5.b2), combined with a progressive rise of the mean high-frequency spectral distribution (Fig. 6.b2). The polar diagrams of minor eigenfunctions $(\xi_2^{VELbow}, \xi_2^{RMS})$ in Fig. 5.c2, and $(\eta_2^{VELbow}, \eta_2^{TRIratio})$ in Fig. 6.c2, reflect this tendency by an inversion of the direction being taken by the perturbations towards the middle of the mean sequences. The first one highlights the combined effect of bow acceleration with a gain of global sound energy, while the second one suggests that decelerating the bow speed might induce a quicker extinction of the high-frequency sound partials.

This last result seems to be coherent with the physics of the instrument, since within a pulling bow movement, the cellists can choose to independently

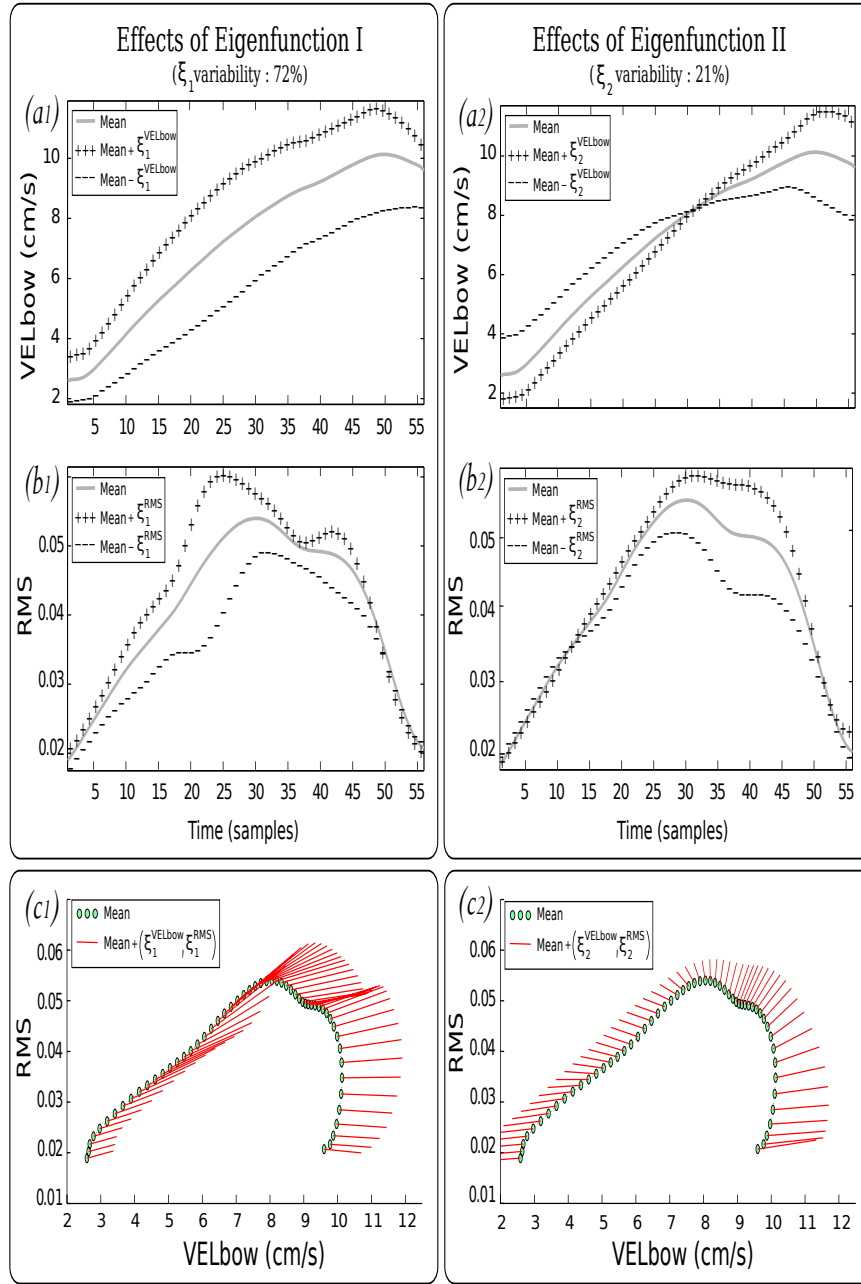


Fig. 5. *Left* : Effects of adding or subtracting the 1st bivariate eigenfunction ξ_1 to or from the mean curve of Bow velocity (a_1) and RMS (b_1). *Right* : Effects of adding or subtracting the 2nd bivariate eigenfunction ξ_2 to or from the mean curve of Bow velocity (a_2) and RMS (b_2). The covariate effect of each eigenfunction on both variables is presented in polar diagrams: (c_1) for eigenfunction I and (c_2) for eigenfunction II

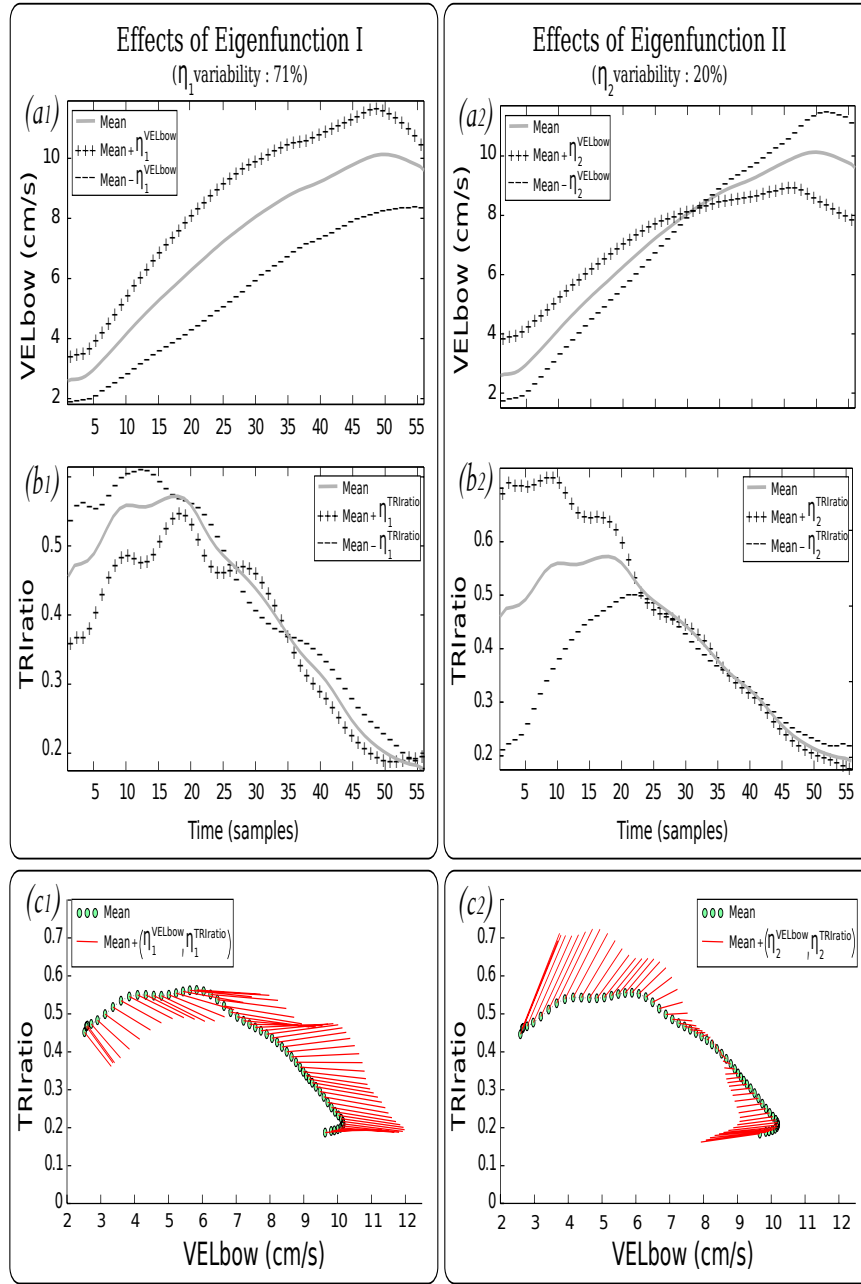


Fig. 6. *Left* : Effects of adding or subtracting the 1st bivariate eigenfunction η_1 to or from the mean curve of Bow velocity (a_1) and TRIratio (b_1). *Right* : Effects of adding or subtracting the 2nd bivariate eigenfunction η_2 to or from the mean curve of Bow velocity (a_2) and TRIratio (b_2). The covariate effect of each eigenfunction on both variables is presented in polar diagrams: (c_1) for eigenfunction I and (c_2) for eigenfunction II

accelerate the speed or reinforce the pressure, to ensure an optimal timbre quality and high-frequency content, along the bow movement. An additional FPCA carried out on bow velocities $VEL_{bow}(t)$ and spectral barycenter evolutions $HSC(t)$, strengthened this link between a bow acceleration and a better stabilization of the spectral barycenter. Nevertheless, further investigations should be conducted with other essential physical parameters like the bow pressure, to completely validate these assumptions.

6 Conclusion

This paper presented a contribution to a better understanding of the musician/instrument interaction in the case of the cello playing. It aimed to better identify the connections between the cellists' motor control principles and the related expressive timbral features. More specifically, the role played by ancillary movements was investigated through a full postural constraint, whose effects were assessed both on the cellist's bowing gesture velocity and spectro-temporal acoustic features of the produced sounds.

The results turned out to be coherent with the bowed-string physics [7] and synthesis models [5]. Indeed, a canonical correlation analysis (CCA) performed on the mean values of sound-gesture parameters, revealed that the discomfort caused by the postural constraint, was mainly linked to an overall decrease of bow velocity. This gestural variation was coupled to an *harsh* spectro-temporal acoustic transformation, combining a slope attack fall of the sound temporal envelope, with an energy rise in the high-frequency partials. Furthermore, a functional analysis of the descriptor data allowed to extract principal components (FPCs) characterizing the sound-bowing gesture covariations in time. These eigenfunctions first confirmed the major trend already identified by CCA. On the other hand, they also highlighted a minor influence of the postural constraint on a bow deceleration, coupled with a reinforcement in *brightness* (i.e. the upper partials in the spectrum) at the beginning of the sound.

Further investigations might consist in identifying more finely the type of cellists' postural movements, ensuring an optimal coupling between their bow velocity and the preservation of spectro-temporal acoustic features.

Acknowledgments. This work was partly supported by the French National Research Agency and is part of the "Sonimove" project (ANR-14-CE24-0018).

References

1. Bianco, T., Freour, V., Rasamimanana, N., Bevilacqua, F., Caussé, R.: On gestural variation and coarticulation effects in sound control. In: *Gesture in Embodied Communication and Human-Computer Interaction*, pp. 134–145. Springer (2009)
2. Caramiaux, B., Bevilacqua, F., Schnell, N.: Towards a gesture-sound cross-modal analysis. In: *Gesture in embodied communication and human-computer interaction*, pp. 158–170. Springer (2009)

3. Chadeaux, D., Le Carrou, J.L., Wanderley, M.M., Fabre, B., Daudet, L.: Gestural strategies in the harp performance. *Acta Acustica united with Acustica* 99(6), 986–996 (2013)
4. Chudy, M., Carrillo, A.P., Dixon, S.: On the relation between gesture, tone production and perception in classical cello performance. In: *Proceedings of Meetings on Acoustics*. vol. 19, p. 035017. Acoustical Society of America (2013)
5. Demoucron, M.: On the control of virtual violins-Physical modelling and control of bowed string instruments. Ph.D. thesis, Université Pierre et Marie Curie-Paris VI; Royal Institute of Technology, Stockholm (2008)
6. Desmet, F., Nijs, L., Demey, M., Lesaffre, M., Martens, J.P., Leman, M.: Assessing a clarinet player’s performer gestures in relation to locally intended musical targets. *Journal of New Music Research* 41(1), 31–48 (2012)
7. Guettler, K., Schoonderwaldt, E., Askenfelt, A.: Bow speed or bowing position - which one influence spectrum the most ? *Proceedings of the Stockholm Music Acoustic Conference (SMAC 03)* (August 6-9 2003)
8. Kim, H.G., Moreau, N., Sikora, T.: *MPEG-7 audio and beyond: Audio content indexing and retrieval*. John Wiley & Sons (2006)
9. Lartillot, O., Toivainen, P.: A matlab toolbox for musical feature extraction from audio. In: *Proceedings of the International Conference on Digital Audio Effects*. pp. 237–244 (2007)
10. Leman, M.: *Embodied music cognition and mediation technology*. Mit Press (2008)
11. Peeters, G.: A large set of audio features for sound description (similarity and classification) in the *cuidado* project. Tech. rep., IRCAM (2004)
12. Pollard, H.F., Jansson, E.V.: A tritestimulus method for the specification of musical timbre. *Acta Acustica united with Acustica* 51(3), 162–171 (1982)
13. Ramsay, J.O.: *Functional data analysis*. Wiley Online Library (2006)
14. Rozé, J., Ystad, S., Aramaki, M., Kronland-Martinet, R., Voinier, T., Bourdin, C., Chadeaux, D., Dufrenne, M.: Exploring the effects of constraints on the cellist’s postural displacements and their musical expressivity. To appear in *post-proceedings of CMMR 2015 - Music, Mind and Embodiment*, Plymouth (2015)
15. Schelleng, J.C.: The bowed string and the player. *The Journal of the Acoustical Society of America* 53(1), 26–41 (1973)
16. Thompson, M.R., Luck, G.: Exploring relationships between pianists’ body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae* 16(1), 19–40 (2012)
17. Visi, F., Coorevits, E., Miranda, E., Leman, M.: Effects of different bow stroke styles on body movements of a viola player: an exploratory study. Ann Arbor, MI: Michigan Publishing, University of Michigan Library (2014)
18. Wanderley, M.M., Vines, B.W., Middleton, N., McKay, C., Hatch, W.: The musical significance of clarinetists’ ancillary gestures: An exploration of the field. *Journal of New Music Research* 34(1), 97–113 (2005)