



**HAL**  
open science

## A Custom Feature Set For Analyzing Historically Informed Baroque Performances

Nolan Eley, Théodora Psychoyou, Catherine Lavandier, Brian F. G. Katz

► **To cite this version:**

Nolan Eley, Théodora Psychoyou, Catherine Lavandier, Brian F. G. Katz. A Custom Feature Set For Analyzing Historically Informed Baroque Performances. ICA 2022 Proceedings, Oct 2022, Gyeongju, South Korea. hal-03836558

**HAL Id: hal-03836558**

**<https://hal.science/hal-03836558>**

Submitted on 2 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ABS-0872

## A Custom Feature Set For Analyzing Historically Informed Baroque Performances

Nolan ELEY<sup>(1,2)</sup>, Théodora PSYCHOYOU<sup>(3)</sup>, Catherine LAVANDIER<sup>(1)</sup>, Brian F.G. KATZ<sup>(2)</sup>

<sup>(1)</sup>ETIS Laboratory, CY Cergy Paris University, ENSEA, CNRS, France, nolan.eley@dalembert.upmc.fr

<sup>(2)</sup>Sorbonne Université, CNRS, Institut Jean Le Rond d'Alembert UMR7190, Paris, France, brian.katz@sorbonne-universite.fr

<sup>(3)</sup>Sorbonne Université, Institut de Recherche en Musicologie UMR 8223, France

### ABSTRACT

A common strategy for analyzing music performances is to extract a high number of low-level features followed by dimensionality reduction. While this can be an effective way of identifying differences in performances, interpreting these differences in a musically meaningful way is difficult. An alternative strategy is to select or design features which are appropriate for the task at hand, however, this must be done carefully and in accordance with musicological principles. In this study, a custom feature set was designed to capture baroque expressiveness in music performances, motivated by principles of historically informed performance practice. This feature set was verified on a small set of annotated performances then applied to a larger group of recordings of historically informed solo baroque performances in two halls, one historically appropriate and one modern. These features were examined to study the effect of room acoustics on the baroque expressiveness of the performances in order to gain insight into the role of acoustics in historically informed performances.

Keywords: Music performance analysis, chamber music, room acoustics, historically informed performance

### 1 INTRODUCTION

There is often a large variety among expressive musical performances of a single composition. The field of music performance analysis is concerned with systematically identifying and assessing these variations. There have been many strategies for analyzing different aspects of musical performances, which are well described in [1]. Because each genre may have a different set of intrinsic mechanisms which contribute to an idealized performance within that genre, it is important that the analysis framework take into consideration the musicological context of the task at hand.

One way to address this is through careful selection of low-level features. However, the available low-level features may be limiting and often lack a direct musical meaning. Furthermore, it is not always known which features will be most discriminative and one runs the risk of leaving out meaningful features whose capabilities may be currently unknown. Another strategy, therefore, is to extract a large number of low-level features then apply dimensionality reduction, followed by statistical analysis [2, 3]. While this can be effective at identifying similarities and differences among performances, one disadvantage is that the resulting dimensions can often be difficult to interpret, especially in a way that is musically meaningful. One way to overcome these problems is to design custom features for identifying specific musical parameters which are known to be of interest to the current task, based on musicological principles.

### 2 CONTEXT

There are many factors which can affect musical performance, one of which is the acoustic environment [4]. This study takes place as part of a larger effort to study the interactions between acoustics and musical performance from a historical perspective, merging the study of acoustics with the fields of historically informed

performance (HIP) and historical musicology [5, 6]. More specifically, this study focuses on baroque HIP and whether the acoustics of spaces of the era facilitate HIP musicians' play. To better understand this problem, it is helpful to know which musical parameters are most salient in the genre. Through the study of period musical treatises and manuals, musicologists have identified a number of characteristics which are responsible for defining baroque performance style and serve as the basis for HIP [7, 8, 9].

Musicologists have adopted the term "stylishness" as a way to describe musical expressiveness in terms of its appropriateness for a specific musical and historical context [10]. For example, the parameters which define an expressive performance of 19<sup>th</sup> century romantic music are very different from those which define an equally expressive performance of 18<sup>th</sup> century baroque music. While both performances may be perceived as equally stylish, their perceived expressiveness may vary widely due to this term's strong association with mainstream romantic musical gestures. For this reason, researchers have started utilizing the term "baroque expressive" in order to differentiate a type of expressive performance which is deemed as being baroque stylish [11].

The goal of this paper is to design an analysis framework for identifying the parameters which make up baroque expressiveness among other performance styles of baroque music. This framework will be verified against a small set of recordings which have already been authenticated as representing distinct performance styles (see Section 3). Following this verification, the designed features will be applied to a set of recordings from a previous study [6] to shed light on the role of acoustics in facilitating baroque HIP practice (see Section 4).

### 3 FEATURE DESIGN AND VERIFICATION

#### 3.1 Dataset

Three distinct schools of baroque performance practice are identified in [11], and we refer the reader to that paper for more information. The three schools are referred to as expressive-emotional, modern-literalistic, and baroque-expressive. In that study, listeners were able to identify these three schools of performance as evidenced by ratings in several musical categories. Relying only on the musical example used in that study (the first 8 bars of J.S. Bach's *Sarabanda* from the D minor Partita for solo violin) to verify our analysis framework would not suffice as that would not provide enough data to yield meaningful results. So, the number of recordings was expanded to three full solo violin pieces within each performance school. These pieces, all composed for solo violin by J.S. Bach, were the aforementioned *Sarabanda*, the *Gavotte en Rondeau* from the E major Partita, and the *Largo* from the C major Sonata. While these additional recordings have not been verified by listeners as embodying specific performance parameters, it is assumed that additional recordings by the same musicians would carry similar performance characteristics to the one previously examined.

Below follows the development of features to identify baroque expressiveness in three major areas which listeners have already identified in [11], namely, phrasing, tone quality, and vibrato. The verification for each feature set is done using support vector machines (SVM) using a one versus one approach with cross validation (10 random folds) to predict the musician which serves as a proxy for performance style.

#### 3.2 Phrasing

The phrasing style in a baroque expressive performance is often described as *rhetorical*, since it tends to mimic the patterns found in the art of oratory, in that it is varied, locally nuanced, and rhythmically flexible [12]. By contrast, other performance styles tend to use *legato*, or continuous phrasing, in order to highlight cadences and large phrase boundaries.

Researchers have found that phrasing is closely related to changes in tempo and loudness, and correlations are typically found between these two musical parameters [13, 14]. Our methodology for capturing phrasing was to separate the recordings into segments of different lengths (1, 2, 4, 8, and 16 bars) and extract features from the tempo and loudness curves of these segments as well as the entire recordings. In total, the dataset yielded 1029 segments. The tempo curves were made up of the note-level tempo values smoothed with a moving average filter since smoothed tempo curves have been found to be helpful in modeling expressive timing [15, 16]. The loudness curves were calculated using the frame-wise root-mean square (RMS) with a window size of 0.05 s and a hop size of 0.02 s. The features extracted from these curves are range, standard deviation

(SD), and coefficients of a parabolic function (2nd-order polynomial) which have been shown to be effective in modeling expressive performance [17, 18].

Using these features as input to an SVM tasked with predicting the musician, we were able to achieve an average F-score (the geometric mean of precision and recall) of 72%. As further exploration, the same classifier was used to predict the composition, and this achieved an average F-score of 87%. As the different compositions are stylistically different, they should exhibit very different phrasing characteristics, meaning this F-score may reflect somewhat the upper bound of the discriminability of this feature set.

### 3.3 Tone Quality

Tone quality or timbre is multidimensional and complex. One dimension of tone quality which was found to be representative of baroque expressive play is “lightness of tone.” Our methodology for identifying this tone quality was to use mel-frequency cepstral coefficients (MFCC) on the harmonic and percussive components of the signal which were extracted using the median filter approach [19]. This approach applies a median filter to the spectrogram either across successive bins or frames which enhances and suppresses either the harmonic or percussive components of the signal. A typical use case for such an algorithm might be to separate the drums in a pop song to facilitate remixing. However, when applied to solo instrumental music, the percussive components end up highlighting the nontonal, transient, and stochastic aspects of the signal which translate to elements such as bow and articulation technique which are likely correlated with one’s perception of tone intensity [20]. Additionally, the nontonal spectrum has previously been identified as being particularly discriminative in instrument classification [21]. MFCCs were extracted from these harmonic and percussive signals in addition to the original, unprocessed audio. The MFCCs were calculated frame-wise using a window size of 0.04 s and a hop size of 0.01 s, and 21 coefficients were used.

To quantify the effectiveness of these various MFCC groups in identifying the performer, four classification tasks were considered. One using each the percussive, harmonic, and unprocessed audio, and one which used the MFCCs from both the harmonic and percussive audio. Using the unprocessed audio, the classifier was able to achieve a class-weighted average F-score of 68%. Using the harmonic and percussive signals improved the classification to 72% and 77%, respectively. And finally, using MFCCs from both the harmonic and percussive signals together achieved the highest class-weighted average F-score of 83%. These results suggest that the harmonic, and especially the percussive components of the signal can be more discriminative than the unprocessed signal, although more testing is needed to see how robust these components are for different types of signals.

### 3.4 Vibrato

Vibrato is a musical ornamentation made up of semi-periodic fluctuations of pitch and is often described by two features: rate and extent [22]. The rate refers to the frequency of modulation which is typically in the range of 4 Hz to 8 Hz [23], while the extent refers to the depth or intensity of the vibrato, measured in cents. Baroque expressive performances tend to use less intense vibrato, less frequently than mainstream performance styles.

Most methods calculate extent in reference to a pitch center, usually the mean [24, 23, 25, 26]. However, this assumes that the fundamental frequency remains unchanged throughout the duration of the note. While this may be the case in controlled laboratory experiments and therefore would not pose a problem to those examining vibrato parameters under such conditions [27], in real performance situations, performers do not always start and end each note at the precise pitch center and may adjust their intonation while maintaining vibrato. One previous attempt to control for this used linear regression to model the pitch center [28].

The proposed methodology aims to improve upon previous methods in two ways. First, the mean removal step is modified by modeling the change in fundamental frequency over the duration of the note with a polynomial curve, aiming to improve the accuracy of the measured vibrato extent. Second, it models the evolution of the vibrato rate as it changes throughout the duration of the note using polynomial fit coefficients, rather than relying on a single measure (such as the median) for the vibrato rate.

The first step in the process is to extract pitch contours [29]; this was done using the Melodia plug-in<sup>1</sup> for Sonic Visualiser<sup>2</sup>. Pitch contours are continuous curves of varying length and can be output as multi-pitch

---

<sup>1</sup><https://www.upf.edu/web/mtg/melodia>

<sup>2</sup><https://www.sonicvisualiser.org/>

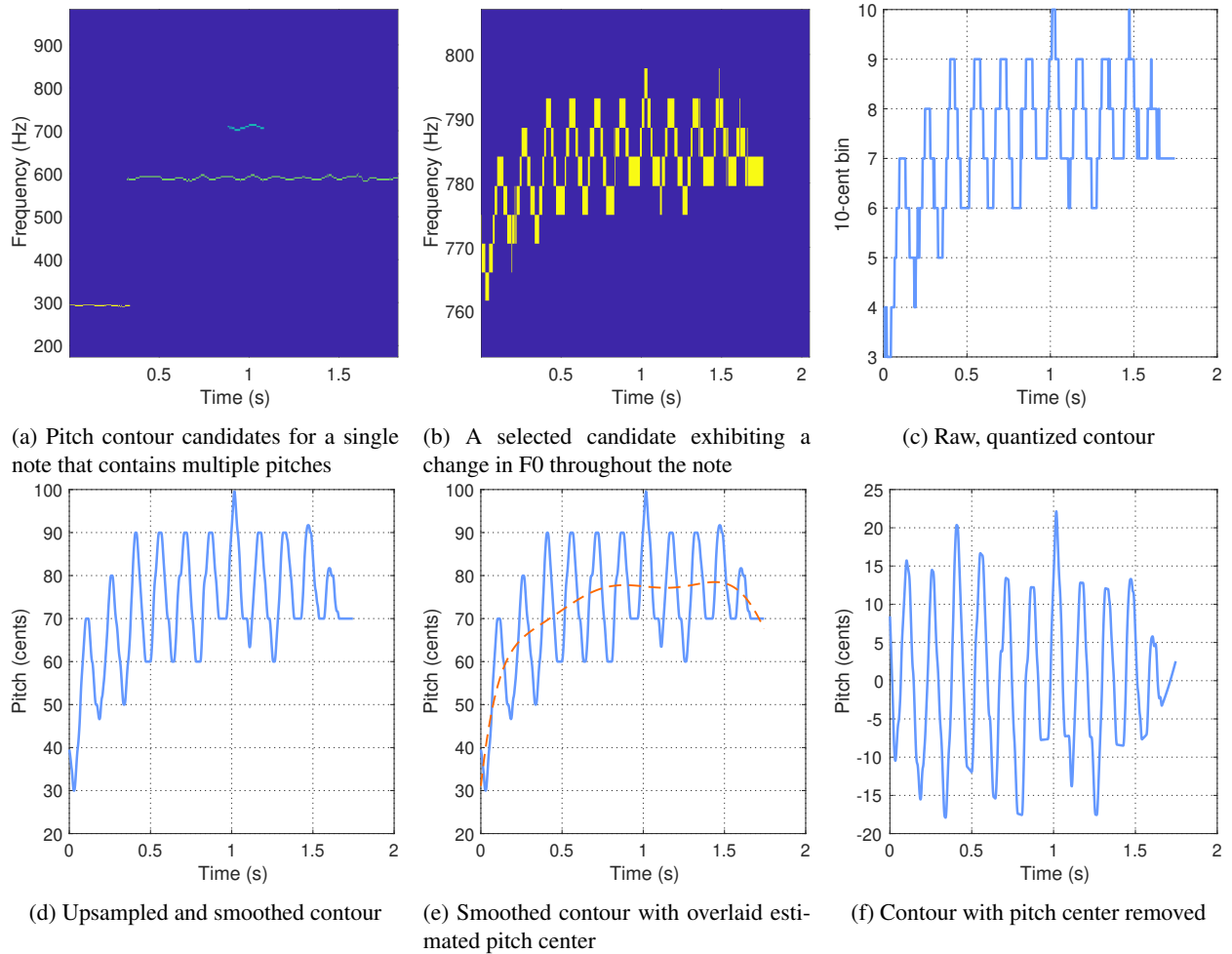


Figure 1. Vibrato feature extraction.

representations giving them an advantage to other fundamental frequency algorithms such as pYin [30] when applied to polyphonic contexts. The pitch contours were separated by note using note onset information. In cases where multiple contours were detected during the span of one note, the best pitch contour was chosen by summing the energy in each bin (representing 10 cents) and then selecting the bin with the most energy. This bin, along with the six bins above and below (equivalent to a range of 130 cents) were selected for further processing (see Fig. 1b).

Because the raw contour was quantized to the nearest 10-cent bin (see Fig. 1c), the contour was upsampled, smoothed using a moving average filter, and factored by a gain of 10 so that each bin would represent one cent (see Fig. 1d). The presence or lack of vibrato was detected if the signal met three conditions: a minimum length, zero crossing rate (on the pitch-centered signal), and extent. These thresholds were arrived at empirically through preliminary tests.

If the presence of vibrato was detected the pitch center was approximated using a polynomial curve of degree 5 (see Fig. 1e). This pitch center was then subtracted from each value, resulting in the final signal seen in Fig. 1f from which the vibrato parameters were extracted. The extent was calculated by taking the mean of the doubled absolute value of each peak and trough. The rate was calculated frame-wise by taking an auto-correlation of the signal followed by peak picking with a peak distance threshold set to stay within the typical rate of vibrato. The median value across all frames was selected as the rate. Polynomial fit coefficients

Table 1. Vibrato characteristics for each performance style.

| Style                | Prevalence | Rate (mean & SD)  | Extent (mean & SD)   |
|----------------------|------------|-------------------|----------------------|
| Expressive-emotional | 32%        | 7.1 Hz $\pm$ 0.75 | 16.8 cents $\pm$ 5.8 |
| Modern-literalistic  | 22%        | 7.3 Hz $\pm$ 0.73 | 19.1 cents $\pm$ 7.3 |
| Baroque-expressive   | 13%        | 6.8 Hz $\pm$ 0.87 | 12.7 cents $\pm$ 3.9 |

of order 3 were taken of the frame-wise rate to model the evolution of the rate across the duration of the note.

In order to quantify the discriminability of these proposed vibrato features, four classification tasks were considered. Two tasks used mean removal while the other two used the polynomial curve modeled pitch center removal. These two groups either used only the baseline features (rate and extent) or the baseline features in addition to polynomial fit coefficients. Overall, 972 notes with vibrato were identified out of 4374 played notes.

In general, the different classification tasks did not yield very different results. Using the baseline features (vibrato rate and extent) yielded class-weighted average F-scores of 47% and 49% using the mean and polynomial curve as pitch center estimation, respectively. Adding the additional features meant to model the change of rate yielded class-weighted average F-scores of 46% and 52% using the mean and polynomial curve as pitch center estimation, respectively. These F-scores are all significantly above random chance (33%) and, overall, the proposed changes in methodology and additional features yielded an approximately 5% increase in classification accuracy.

Various vibrato characteristics are shown in Table 1 and reveal that, as expected, the baroque-expressive style exhibits the least amount of vibrato, in terms of prevalence (the ratio of notes with vibrato to those without), rate, and extent. The expressive-emotional style predictably shows the most prevalent vibrato, and both it and the modern-literalistic style show somewhat comparable rates and extents.

## 4 APPLIED ANALYSIS

The dataset that this analysis framework was developed for was part of a larger study intended to help understand the role of acoustics in historically informed performance. A broad overview of the study can be found in [6], but a brief description follows. Musicians specializing in historically informed baroque performance practice played in two rooms, one historically appropriate to the musical era, and one modern. The historical room is the *Salon des Nobles* in Versailles, and the modern hall is the amphitheater from the *Cité de la Musique* in Paris. 10 musicians participated, 4 violists da gamba, 3 flutists, and 3 theorbists. The musicians each played 3 pieces of music which were chosen for each instrument by musicologists, and they repeated each piece around 3 times in each room, rendering a total of more than 180 recordings. The same classification exercises used to verify the features were used when applying the analysis framework to these performances, but instead of being tasked with predicting the *musician*, they were tasked with predicting the *room* in which the performances occurred.

Using the phrasing set of features yielded average F-scores of 90%, 95%, and 83% for the viol, flute, and theorbo, respectively. While this is fairly high, it is possible that some of the features which make up the phrasing may have been influenced by the sound of the room captured by the microphone. Even though each instrument was recorded with a directional microphone situated 1 m away, resulting in a high ratio of instrument sound compared to room sound, the sound of the room can still be identified in most recordings. Some of the features which make up the phrasing set are related to the RMS level, which could be influenced by this incidental room sound. It is difficult to tell exactly which identified differences are due to this room sound and which are due to actual differences in the performance. In order to clarify this, the classification task was repeated on these three instruments using only the tempo-related features in the phrasing set. This yielded average F-scores of 64%, 73%, and 53%, respectively. This indicates that at least some of the phrasing expressiveness was changed due to the room, at least for the viol and flute. It was previously reported in [6]

that the performances by the theorbists included many mistakes, so it is not surprising that the results using only tempo-related features resulted in an accuracy that is only slightly above random chance.

The vibrato analysis is limited to the viol since it is the only instrument used in the study which is capable of vibrato. In total, 1145 notes were identified with vibrato out of 6888 played notes, meaning almost 17% of the played notes were embellished with vibrato, compared to the 22% in the test set of Bach recordings. The results of the classifier using the vibrato feature set performed no better than random chance, indicating that the vibrato style did not change as a function of the room. This is not very surprising as it was the least discriminative feature in the test set. Despite the poor performance of vibrato as a predictor of *room*, some interesting findings can still be reported. 54% of the vibrato occurred in the Salon des Nobles. Furthermore, about 30% of the vibrato can be attributed to one of the four violists who had 3 years of performing experience, compared to an average of 37.3 (SD: 5.5) years of experience the other 3 violists had. This implies that perhaps the more liberal use of vibrato may be correlated with having less experience or different training. Lastly, the average vibrato rate was 6.8Hz and the average extent was 15.2 cents.

In general, the features designed to describe tone quality were able to predict the room with fairly high accuracy, with class-weighted average F-scores among the 90% range using the normal, harmonic, or percussive components of the signal. There was not much of a difference between the unprocessed and the harmonic signals, while the percussive signal resulted in an F-score of about 5-7 points lower. However, it should be noted that these MFCCs are likely influenced by the room sound to some degree though it is difficult to say precisely how much. The percussive signal, with its emphasis on transients, is likely influenced by the room less, and therefore, its slightly lower F-scores may more accurately reflect an actual difference in tone quality.

## 5 Conclusions

The goal of this study was to simplify the analysis of music performance to a few musically meaningful dimensions which are relevant to the task at hand, in this case, historically informed baroque performance. We identified three major areas which are important to baroque expressiveness, namely, phrasing, vibrato, and tone quality, and developed an analysis framework for identifying differences in these areas. This framework was first successfully verified on recordings from a published study in which musicologists and listeners had identified significant differences among baroque performance styles.

Subsequently, this analysis framework was applied to an unrelated set of recordings of baroque performances, with the goal of identifying significant differences among the performances which may have been influenced by the room and its acoustics. Judging the success of the analysis framework on this dataset was difficult, as it was not previously known how these performances differed, nor to what degree. Overall, the phrasing and tone quality features identified significant differences between performances in the two rooms. However, it was noted that some of this may be due to the sound of the room incidentally picked up by the microphone. The vibrato feature set did not show any significant differences between the two rooms, however, it did reveal a difference between performers which may be correlated with their experience or training.

While this method of performance analysis was able to reduce the problem to a few components, these components are still described by high-dimensional feature sets which exhibit problems of interpretability. More work (and larger datasets) would be needed to reduce the dimensionality of these feature sets in a way that makes the results easier to interpret. For example, we are able to identify differences in phrasing, but cannot say definitively what this means on the continuum of “legato” to “separated”. A complementary listening test performed on a subset of these recordings may help identify what these differences are.

One issue with developing an analysis framework on a few recordings of Bach violin pieces is the major risk of over-fitting to such a small and unrepresentative dataset. Ultimately, without larger datasets of performances of baroque music annotated along specific dimensions of performance, it is difficult to create a truly robust analysis framework. Still, the approach outlined in this paper suggests that a tailored method of analyzing music performance can be effective and yield meaningful results that are more interpretable than broad methods of low-level feature extraction.



## ACKNOWLEDGEMENTS

This work was supported by the Paris Seine Graduate School Humanities, Creation, Heritage, Investissement d'Avenir ANR-17-EURE-0021 – Foundation for Cultural Heritage Science. This project takes place as part of [EVAA\\_Ver: Expérience Virtuelle en Acoustique Archéologique](#). We would also like to thank the research center of the Château de Versailles and the Musée de la Musique of the Philharmonie de Paris, as well as the musicians who participated in the study.

## REFERENCES

- [1] Lerch A, Arthur C, Pati A, Gururani S. An Interdisciplinary Review of Music Performance Analysis. *Transactions of the Intl Soc for Music Information Retrieval*. 2020 Nov;3(1):221-45.
- [2] Schärer Kalkandjiev Z. The Influence of Room Acoustics on Solo Music Performances. An Empirical Investigation [PhD Dissertation]. Berlin: Technischen Universität; 2015.
- [3] Amengual Garí S, Kob M, Lokki T. Analysis of trumpet performance adjustments due to room acoustics. In: *Proc of the Intl Symposium on Room Acoustics*. Amsterdam; 2019. p. 65-73.
- [4] Kob M, Amengual Garí SV, Schärer Kalkandjiev Z. Room Effect on Musicians' Performance. In: Blauert J, Braasch J, editors. *The Technology of Binaural Understanding*. Springer Intl Publishing; 2020. p. 223-49. Series Title: *Modern Acoustics and Signal Processing*.
- [5] Eley N, Mullins S, Stitt P, Katz BFG. Virtual Notre-Dame: Preliminary results of real-time auralization with choir members. *Immersive and 3D Audio: from Architecture to Automotive (I3DA)*. 2021.
- [6] Eley N, Lavandier C, Psychoyou T, Jossic M, Katz BFG. Performance analysis of solo baroque music played in a period and modern hall. *16th french acoustics congress of the french acoustics society*. 2022 Apr.
- [7] Lawson C, Stowell R. *The historical performance of music: an introduction*. Repr ed. Cambridge handbooks to the historical performance of music. Cambridge: Cambridge University Press; 2003.
- [8] Haynes B. *The End of Early Music: A Period Performer's History of Music*. New York, NY: Oxford University Press; 2007.
- [9] Donington R. *Baroque Music Style and Performance: A Handbook*. W.W. Norton & Company; 1982.
- [10] Schubert E, Fabian D. The dimensions of baroque music performance: a semantic differential study. *Psychology of Music*. 2006 Oct;34(4):573-87.
- [11] Fabian D, Schubert E. Baroque expressiveness and stylishness in three recordings of the D minor Sarabanda for solo violin by J.S. Bach. *Music Performance Research*. 2009 Jan;3:36-56.
- [12] Ponsford D. Instrumental performance in the seventeenth century. In: Lawson C, Stowell R, editors. *The Cambridge History of Musical Performance*. Cambridge University Press; 2012. p. 421-47.
- [13] Palmer C. Mapping Musical Thought to Musical Performance. *Journal of Experimental Psychology: Human Perception and Performance*. 1989;15(12):331-46.
- [14] Palmer C. Music Performance. *Annual Review of Psychology*. 1997;48:115-38.
- [15] Cambouropoulos E, Dixon S, Goebel W, Widmer G. Human Preferences for Tempo Smoothness. In: *Proc of the VII Intl Symposium on Systematic and Comparative Musicology*; 2001. p. 18-26.
- [16] Schreiber H, Zalkow F, Müller M. Modeling and estimating local tempo: A case study on Chopin's Mazurkas. In: *Proc of the 21st Intl Soc for Music Information Retrieval Conf*; 2020. p. 773-9.



- [17] Li S, Dixon S, Plumbley MD. Clustering Expressive Timing With Regressed Polynomial Coefficients Demonstrated By A Model Selection Test. In: Proc of the 18th Intl Soc for Music Information Retrieval Conf; 2017. p. 457-63.
- [18] McAngus Todd NP. The dynamics of dynamics: A model of musical expression. The Journal of the Acoustical Soc of America. 1992 Jun;91(6):3540-50.
- [19] Fitzgerald D. Harmonic/Percussive Separation using Median Filtering. 13th Intl Conf on Digital Audio Effects (DAFx-10). 2010 Jan.
- [20] Schoonderwaldt E, Guettler K, Askenfelt A. Effect of the Width of the Bow Hair on the Violin String Spectrum. In: Proc of the Stockholm Music Acoustics Conf. Stockholm, Sweden; 2003. p. 91-4.
- [21] Fragoulis D, Papaodysseus C, Exarhos M, Roussopoulos G, Panagopoulos T, Kamarotos D. Automated classification of piano-guitar notes. IEEE Trans Audio Speech Lang Process. 2006 May;14(3):1040-50.
- [22] Seashore CE. Psychology of Music. New York and London: McGraw-Hill Book Company; 1938.
- [23] Weninger F, Amir N, Amir O, Ronen I, Eyben F, Schuller B. Robust feature extraction for automatic recognition of vibrato singing in recorded polyphonic music. In: Intl Conf on Acoustics, Speech and Signal Processing (ICASSP). Kyoto, Japan: IEEE; 2012. p. 85-8.
- [24] Herrera P, Bonada J. Vibrato Extraction and Parameterization in the Spectral Modeling Synthesis Framework. Proc of the Digital Audio Effects Workshop (DAFX98). 1998.
- [25] Pang HS, Yoon DH. Automatic detection of vibrato in monophonic music. Pattern Recognition. 2005 Jul;38(7):1135-8.
- [26] Regnier L, Peeters G. Singing Voice Detection in Music Tracks using Direct Voice Vibrato Detection. In: Intl Conf on Acoustics, Speech and Signal Processing (ICASSP). Taipei, Taiwan; 2009. p. 1685-8.
- [27] Brown JC, Vaughn KV. Pitch center of stringed instrument vibrato tones. The Journal of the Acoustical Soc of America. 1996 Sep;100(3):1728-35.
- [28] De Muynke J, Eley N, Ferrando J, Katz BFG. Preliminary analysis of vocal ensemble performances in real-time historical auralizations of the Palais des Papes. In: Proc on the 2nd Symposium of The Acoustics of Ancient Theatres. Verona, Italy; 2022. .
- [29] Salamon J, Gomez E. Melody Extraction From Polyphonic Music Signals Using Pitch Contour Characteristics. IEEE Transactions on Audio, Speech, and Language Processing. 2012;20(6):1759-70.
- [30] Mauch M, Dixon S. PYIN: A fundamental frequency estimator using probabilistic threshold distributions. In: Intl Conf on Acoustics, Speech and Signal Processing (ICASSP). Florence, Italy: IEEE; 2014. p. 659-63.