and MuRET [16, 17]. MuRET, a framework developed by David Rizo at the University of Alicante as part of the Hispamus Project [18], is the only one with support for handwritten mensural notation. [3]

As with other OMR tools, MuRET provides information on the pitch and shape of notes as part of its workflow and allows the user to correct the recognized symbols. It also has support for encoding the imperfect / perfect / altered durational values of mensural notes, though these values are not retrieved automatically—as in the case of pitches and note shapes—and instead have to be entered manually in **mens—a Humdrum format for mensural notation [19]—using MuRET's interface.

At the end of the OMR process, the user can export the encoded music into MEI (Music Encoding Initiative)—a symbolic file format that has support for encoding mensural notation [20]. It is possible to export two kinds of MEI files in MuRET, *score-based* and *parts-based*. The score-based file has the voices stacked and aligned based on the duration of the notes; however, for the notes to be correctly lined up, the user must have provided the durational values for the notes in **mens (i.e., include imperfections and alterations). The parts-based file provides the music encoded as separate parts (just as in the original sources), encoding only pitch and note-shape information for each note. The information about the notes' imperfect, perfect, or altered values is calculated by the next tool in the MIR pipeline, the MP Editor, which will return the score-based Mensural MEI file.

2.3 The Measuring Polyphony Editor (MP Editor)

The MP Editor is an online editor for mensural notation [21], developed under the Measuring Polyphony Project directed by Karen Desmond. [4] Until recently, this was the only mensural notation editor that performed automatic voice alignment, [5] a functionality that I implemented. The MP Editor takes in a series of notes (i.e., pitches and note shape) and automatically lines up the voices into a score using a re-implementation of the scoring-up script in [9]; additionally, after scoring up the voices, the MP Editor allows the user to perform editorial corrections. In a previous paper, I presented the work done to use MuRET's parts-based output as MP Editor's input, which allows for a complete and semi-automatic pipeline of the scoring up of each piece of the GuatC 1 corpus [23].

2.4 The Dissonance Filter (DF)

The goal of this article is to evaluate the efficiency of using counterpoint error markers while making editorial corrections in the last step of the MIR pipeline. On a previous paper, I presented the work done to integrate humlib—Humdrum's data-parsing library—within the MP Editor [23], [6] which allows humlib's Renaissance dissonance filter to label the dissonant notes. [7] The complete list of dissonance labels available in the filter can be consulted on the Verovio Humdrum Viewer (VHV) page. [8] In the MP Editor, the DF marks the dissonances in the score, and makes a distinction between the legal (e.g., passing tones, lower/upper neighbours, and suspensions) and illegal dissonances, rendering the former in blue and the latter in orange. From the set of dissonance labels presented in the VHV documentation page, the ones considered illegal (and rendered in orange font) are *Z*/*z*, *Y*/*y*, *L*/*l*, and *x*. For an example of how the DF labels the dissonances in the MP Editor, see Figures 3–5.

3. METHODOLOGY

3.1 Data Preparation

We randomly selected fifteen pieces from the GuatC 1 manuscript, where each piece was either a mass movement or a short polyphonic piece. The fifteen pieces represent around 20% of the total corpus. To ease the editorial work in the MP Editor, we subdivided the long mass movements into smaller units that we called "self-contained units." Here, we are considering as a "long movement" anything that has more than two openings (i.e., four pages) in the manuscript. We defined the "self-contained unit" as the minimum number of consecutive sections (one or more) that start at a page beginning and end at a page ending. This definition was needed because we had to face the problem of handling sections whose beginning or end do not correspond to the beginning or ending of an image (i.e., a page) since MuRET can only produce an MEI output from a set of one or more selected images.

We obtained a total of 23 self-contained units and divided these into two datasets: a *DF Dataset* where the dissonance filter would be activated during the correction process, and a *NDF Dataset* where no dissonance filter would be used during correction. To guarantee a balanced dataset, we arranged for the DF and NDF datasets to have the same average number of measures, voices, and illegal dissonances. We also made sure that each dataset had close to the same number of CPDL transcriptions to use as a reference, providing another musician's interpretation of where the mistakes are found and how to fix them.

Table 1 shows the pieces in the two datasets. The first column contains the name of the self-contained unit (e.g., Missa9.2_Gloria1), which shows the mass number in the inventory table of the GuatC 1 found in [25] (e.g., Missa9 for the 9th mass in GuatC 1), the movement number and name (e.g., Missa9.2_Gloria1 indicates the second movement in the mass, which is always a *Gloria*), and the unit number within that movement (the first unit of the Gloria).

[3] `https://muret.dlsi.ua.es/muret/#/home`

[4] `https://editor.measuringpolyphony.org`

[5] Imperfections and alterations are entered manually in the editor from the Computerized Mensural Music Editing (CMME) project. The new Mensural Rhythm Interpretation Tool (MeRIT) automatically scores up the piece, but its focus is on introspection and pedagogical use and not on producing an edited score [22].

[6] `https://github.com/craigsapp/humlib`

[7] This filter was developed by Alex Morgan for the Josquin Research Project (`https://josquin.stanford.edu`). It is based on Peter Schubert's book on modal counterpoint [24].

[8] `https://doc.verovio.humdrum.org/filter/dissonant/`

3.2 Experiment Setup and Evaluation Procedure

The experiment was conducted by a Bachelor of Music student with a major in voice and a minor in early music. She has knowledge of mensural notation and counterpoint from courses in paleography and Renaissance musicianship, in addition to her participation in early-music vocal ensembles. She corrected each piece following these steps: (1) look at the ends of phrases and sections to see if the voices line up at the cadence in order to determine if there is a note value missing earlier in the piece; and (2) look at the notes preceding that cadence, this would mean looking at all these notes in the NDF Dataset, while focusing on the places with orange labels and the notes preceding them in the DF Dataset. She was also instructed to report the correction time for each piece, annotate any comments about the piece that she considered relevant, and provide the files downloaded from the MP Editor at the end of the correction process. Although most of the corrections at this point in the MIR pipeline should be editorial, we asked the experimenter to still keep an eye out for OMR errors.

We recorded the correction time per piece, the time invested in the correction of the pieces per dataset (DF and NDF), and the accuracy of those corrections. While the average correction time of the DF and NDF datasets is easy to obtain, analyzing the accuracy of the corrections is more involved. To study the accuracy of the corrections, we used three sets of files: (1) the OMR scored-up files obtained by uploading the OMR Parts-based MEI file into the MP Editor and exporting the score with no corrections; (2) the scores corrected by the experimenter; and (3) the CPDL scores. We compared the OMR scores against the experimenter's scores to identify the experimenter's corrections. And we compared the experimenter's scores against the CPDL scores to identify discrepancies between the two transcriptions. [9] While the CPDL scores cannot be considered ground truth, they record another musician's ideas about mistakes in the sources, and help in the process of checking the experimenter's corrections. Whenever we found a discrepancy, we analyzed both versions (the experimenter's and the CPDL's) and chose the best solution, always favoring the one that removed "true" illegal dissonance labels and that guaranteed imitation and motivic consistency. [10] Here, we are distinguishing between "true" and "false" illegal dissonances, where **true** illegal dissonances are notes that are correctly labelled as such (i.e., true positives) and, on the other hand, **false** illegal dissonances are actually legal dissonances that are incorrectly classified by the DF (i.e., false positives). We also re-uploaded the experimenter's files into the MP Editor and activated the DF on her files to check if there were any orange labels left to determine whether these were counterpoint errors missed by the experimenter or if they were "false" illegal dissonances. We found a few instances of these false illegal dissonances as reported in Section 4.1.

[9] The OMR and the experimenter's scores are encoded in MEI, while the CPDL scores are in MusicXML. Therefore, to compare them, we wrote transformation scripts for both formats to convert them into text files that encode the notes in each voice as a sequence of tokens and then compared these text files using a *diff* tool.

[10] Imitation refers to repeated melodic motifs in different voices, typical of sacred polyphony.

4. RESULTS AND DISCUSSION

There are two types of results in this experiment: (1) the average correction time of the NDF and DF datasets, which can be seen in Table 1; and (2) the experimenter score files, which are stored in GitHub. [11]

NDF Dataset					
Self-contained Units	CPDL	Number of			Time (min)
		M	V	IllD	
Missa8.4_Sanctus0	yes	29	4	9	8
Missa9.2_Gloria1	yes	44	4	46	13
Missa9.2_Gloria2	yes	37	4	8	17
Missa10.1_Kyrie0	**part**	71	4	4	18
Missa10.3_Credo1	yes	67	4	64	16
Missa10.5_AgnusI0	yes	23	4	2	7
Missa15.3_Credo2	yes	6	4	43	15
Missa15.3_Credo3	yes	48	3	0	15
Missa16.4_Sanctus2	yes	13	4	17	11
Missa17.5_Agnus0	yes	26	4	1	6
Piece21_Surrexit0	**no**	17	5	21	10
Average		**39.2**	**4**	**19.5**	**12.4**
Standard Deviation					**4.2**

DF Dataset					
Self-contained Units	CPDL	Number of			Time (min)
		M	V	IllD	
Missa7.4_Sanctus1	**no**	29	4	0	1
Missa7.4_Sanctus2	**no**	50	4	52	10
Missa9.6_AgnusII0	yes	35	5	14	7
Missa10.3_Credo2	yes	40	4	3	2
Missa10.3_Credo3	yes	64	4	17	30
Missa13.5_Agnus0	yes	30	4	6	6
Missa14.1_Kyrie0	yes	56	4	23	7
Missa15.1_Kyrie0	yes	44	4	0	1
Missa15.3_Credo1	yes	64	4	61	2
Missa16.4_Sanctus1	yes	22	4	2	1
Missa16.4_Sanctus3	yes	27	3	16	9
Piece25_Surrexit0	yes	15	4	13	6
Average		**39.7**	**4**	**17.3**	**6.8**
Standard Deviation					**8.0**

Table 1: Voice-alignment correction time for each piece in the No Dissonance Filter (NDF) Dataset and the Dissonance Filter (DF) Dataset. The pieces shaded in gray did not require correction. The *M* stands for measures, *V* for voices, *IllD* for illegal dissonances, and the *CPDL* column indicates if a modern transcription exists on CPDL.

[11] https://github.com/martha-thomae/GuatC1/tree/Experiment/MPeditor_files/Score_files

4.1 Discussion

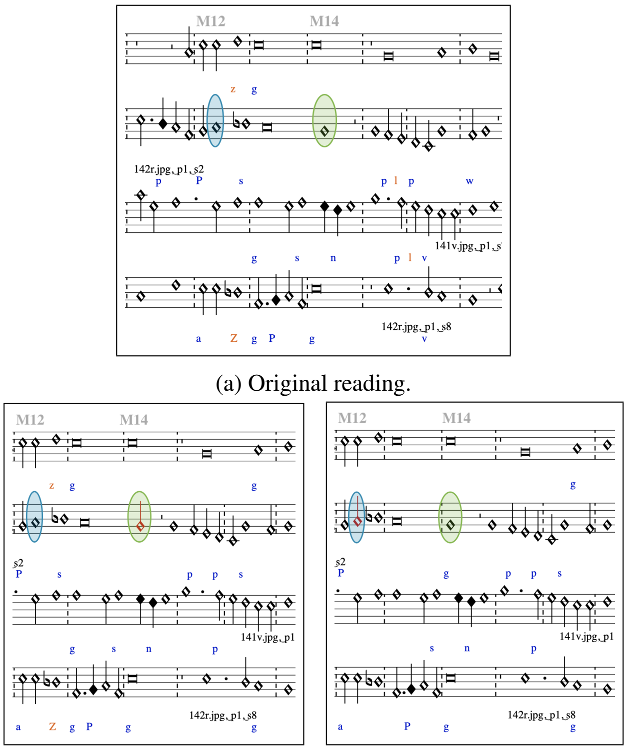
By analyzing the corrections made to the individual pieces and looking at their correction times, the following points become apparent:

1. **The use of the DF reduces the correction time.** As shown in Table 1, the use of the DF reduces the correction time from 12.4 to 6.8 minutes, almost halving it. However, the standard deviation of the DF Dataset is high compared to the one in the NDF Dataset. This is due to an outlier in the DF Dataset. While the correction time for all pieces in the DF Dataset is in the 1–10 minute range, the Missa10.3_Credo3 correction time is 30 minutes. Details about this mass will be provided later in point 4 below. Removing this outlier from the DF Dataset reduces the average correction time to 4.7 minutes (and the standard deviation is reduced from 8.0 to 3.4). Eliminating the outlier reduces the correction time with the DF to almost a third of the correction time for the NDF. Moreover, for pieces that have around the same amount of information—number of measures, voices, and illegal dissonances—and that did not require any corrections (see grey entries in Table 1), the reduction in time when using the DF is considerable. The pieces in grey with 26 and 29 measures took the experimenter 6 minutes to correct without the DF, and 1 minute with the DF. Similarly, the pieces with 48 and 44 measures took the experimenter 15 minutes without the DF and 1 minute with the DF, despite the fact that the latter (the piece in the DF data) has an extra voice.

2. **The use of the DF increases accuracy in the correction.** There were numerous cases where the experimenter made a rhythmic change too late. [12] An example is shown in Figure 3. Figure 3a shows the original manuscript reading of Missa15.3_Credo2. Looking at this scored-up version of the piece, the experimenter noticed that the cadence to G at the end of the example was not correct and knew that she had to cut a minim in the alto voice sometime before the cadence. Her correction, halving a semibreve in measure 14, is shown in Figure 3b. This change still resulted in an illegal dissonance, however, shown in Figure 3b by the orange "z" under the *Bb* in the bass (not shown in the experimenter's NDF dataset). The rhythmic change should have been made a few measures before (in measure 12), as shown in Figure 3c. While the experimenter's correction removes all following illegal dissonances (Figure 3b), it still leaves the ones preceding it. The DF would have shown her where to look—just before the first orange label—to change the note value. This change matches the CPDL correction.

3. **The DF not only aids in identifying scribal errors but also OMR errors that affect the voice alignment.** For pieces where MURET missed a note, or assigned the wrong value of a note or rest in the manuscript, the DF helped to identify the missing or incorrect symbol. [13] Although these pieces were part of the NDF Dataset (where the DF was not used during the experiment), applying the DF on the scored-up OMR file could have saved the experimenter some time as the orange labels help in noticing these errors. By finding the first orange label and checking the manuscript around that spot, the experimenter would have realized that a note/rest in the manuscript was missing (or had the wrong value) in the rendered file.

4. **Style issue – Eighteenth-century works.** The DF was not designed to handle eighteenth-century compositional style. Mass 10 is the only mass in the corpus that we know was composed in the eighteenth century. In many of its self-contained units, one can see the use of seventh chords and ties. [14] In this mass, the DF did not work well, because dissonances that are illegal in Renaissance style are legal in the eighteenth century. The experimenter pointed out a passage in Missa10.3_Credo3, which took her 30 minutes to correct. She invested a lot of time trying to figure out how



(a) Original reading.

(b) Experimenter's version. (c) Correct version.

Figure 3: Missa15.3_Credo2 beginning at measure 12. The circled semibreve in measure 14 is the one halved by the experimenter, while the circled semibreve in measure 12 is the one that should have been halved. **Clefs from top to bottom:** G, G, suboctave G, suboctave G.

[12] This happened in Missa7.4_Sanctus2, Missa15.3_Credo2, and Missa16.4_Sanctus3.

[13] This happened in Missa9.2_Gloria1, Missa9.2_Gloria2, and Missa10.3_Credo1.

[14] Seventh chords are found in Missas10.1_Kyrie0 (measure 4), Missa10.3_Credo1 (measures 11, 12, and 39), Missa10.3_Credo3 (measure 20), and Missa10.5_AgnusI0 (measure 16). While ties are found in Missa10.3_Credo1 and Missa10.3_Credo3.

to get rid of the orange labels and finally left the passage as it was. This was the correct decision as this passage combines Baroque and Renaissance styles, with seventh chords (see Figure 4), such as the third inversion seventh chord over the red "D," shown by the orange "y."

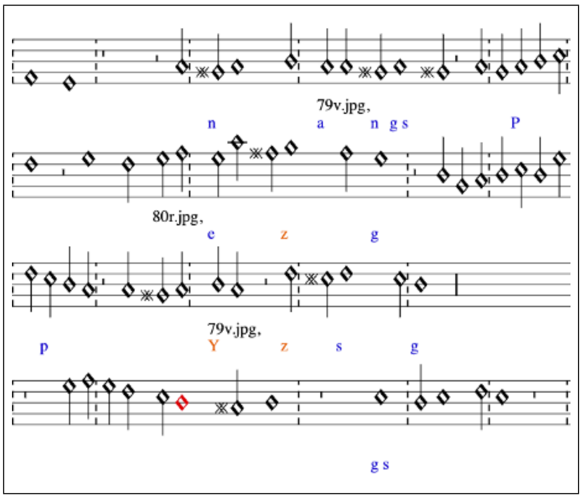


Figure 4: Passage where the experimenter invested a lot of time in Missa10.3_Credo3. It is a combination of Renaissance and Baroque style with fake suspensions and seventh chords. **Clefs from top to bottom:** G, suboctave G, suboctave G, F.

5. **False illegal dissonances.** Rearticulations, minims and semiminims doubling the pitch class of the agent, and extremely short notes like fusas are wrongly labelled as illegal dissonances. This is an issue to correct on the dissonance filter.

Finally, our methodology of re-uploading the file corrected during the experiment, activating the DF, and looking for illegal dissonances left allowed us to catch an error missed by both the experimenter and the CPDL transcriber. Moreover, our methodology of looking for imitative textures allowed us to easily correct this error (see Figure 5).

5. CONCLUSIONS

In this paper, we presented a MIR pipeline designed to obtain corrected symbolic scores for the pieces of a set of Guatemalan choirbooks, and we tested it on the first of these books (GuatC 1). We also presented a way to facilitate correcting the results of the last step of this pipeline, the automatic alignment of the mensural voices recognized by the OMR, conducted in the MP Editor. We used the dissonance filter from the humlib library within the MP Editor to highlight illegal dissonances according to the rules of Renaissance counterpoint. Our findings reveal that using the dissonance filter as a counterpoint error marker for the scoring up of Renaissance music facilitated the spotting of scribal errors and even some OMR errors that went undetected in the previous step of the MIR pipeline. Using the MP Editor and the dissonance filter, the user can see the scored-up OMR file, look for the first illegal dissonance

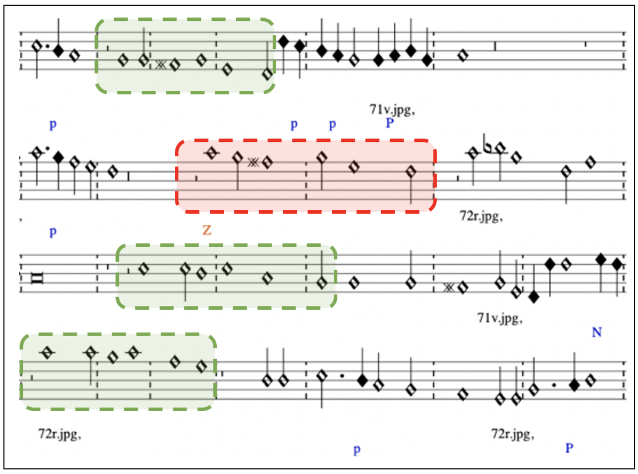


Figure 5: Experimenter and CPDL version of the Christe section of Missa10.1_Kyrie0 with the DF turned on. The alto's A marked by an orange Z label is an undetected error by the experimenter and the CPDL transcriber. Moving this note down to a G results in imitation of the other voices, marked in green boxes. **Clefs from top to bottom:** G, suboctave G, suboctave G, F.

(i.e., the first orange label in the MP Editor), and evaluate if there is an OMR or a scribal error at or before this point. The use of the DF reduces the correction time (to almost a third) and increases the accuracy of the correction process for pieces written in Renaissance style. Further improvements in the time and accuracy of the corrections could be obtained by improving the DF (i.e., resolving the issue of false illegal dissonances). The DF proved to be sensitive to changes in style, as shown in the eighteenth-century mass of the corpus that combines Baroque and Renaissance styles. Future work consists of extending the filter to work on non-Renaissance music.

In conducting this work, the first author has collaborated with the developers of MuRET, MP Editor, and humlib, improving these tools' support for mensural notation and allowing for their interoperability. Interested students and scholars can use the presented MIR pipeline to retrieve editorial music scores for other mensural sources semi-automatically. The OMR part of the proposed pipeline can be substituted by any other OMR tool for mensural notation (e.g., Aruspix) as long as its output conforms with the input expected by the MP Editor [23]. Although this experiment was small, it already showed promising results. The following steps would include replicating these findings with more musicologists and exploring other counterpoint errors (e.g., parallel perfect intervals).

6. ACKNOWLEDGEMENTS

This work draws on research supported by the Fonds de recherche du Québec – Société et culture (FRQSC, Bourse au doctorat en recherche, 13D - Musique, 2019-B2Z-261749) and the Social Sciences and Humanities Research Council of Canada (SSHRC). Special thanks to the developers of the various tools involved in the MIR pipeline, David Rizo (of MuRET), Juliette Regimbal (of the MP Editor), and Craig Sapp (of humlib), as they assisted in the process of allowing for the interoperability of these tools. Thanks to Alexander Morgan, who developed the dissonance filter for the Josquin Research Project; to Geneviève Gates-Panneton, who was the experimenter and provided valuable feedback for this paper; and to Professor Peter Schubert, who served as an expert consultant for questions regarding modal counterpoint. Finally, thanks to Timothy de Reuse and Andres Lou for their help in editing this paper.

7. REFERENCES

[1] O. Morales Abril, "Música local y música foránea en los libros de polifonía de las Catedrales de Guatemala y México," *Jahrbuch für Renaissancemusik: Troja*, vol. Vokalpolyphonie zwischen Alter und Neuer Welt: Musikalische Austauschprozesse zwischen Europa und Lateinamerika im 16. und 17. Jahrhundert, no. 14, pp. 95–124, 2015.

[2] M. E. Thomae, J. E. Cumming, and I. Fujinaga, "Digitization of choirbooks in guatemala," in *Proceedings of the 9th International Conference on Digital Libraries for Musicology (DLfM)*. Prague, Czech Republic: Association for Computing Machinery, 2022, pp. 19–26. [Online]. Available: https://doi.org/10.1145/3543882.3543885

[3] Y.-H. Huang, X. Chen, S. Beck, D. Burn, and L. Van Gool, "Automatic handwritten mensural notation interpreter: from manuscript to MIDI performance," in *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, Málaga, Spain, 2015, pp. 79–85.

[4] D. Pujol, "Polifonía española desconocida conservada en el Archivo Capitular de la Catedral de Guatemala y de la Iglesia parroquial de Santa Eulalia de Jacaltenango," *Anuario Musical*, vol. 20, pp. 3–10, 1965.

[5] R. Stevenson, *Renaissance and Baroque Musical Sources in the Americas*. General Secretariat, Organization of American States, 1970.

[6] R. J. Snow, *A New-World Collection of Polyphony for Holy Week and the Salve Service: Guatemala City, Cathedral Archive, Music MS 4*, ser. Monuments of Renaissance Music. Chicago: University of Chicago Press, 1996, no. 9.

[7] O. Morales Abril, "Catálogo de los acervos musicales del Archivo Histórico Arquidiocesano de Guatemala," N.D., unpublished.

[8] D. Lehnhoff, *Rafael Antonio Castellanos: vida y obra de un músico guatemalteco*. Universidad Rafael Landívar, Instituto de Musicología, 1994. [Online]. Available: http://www.url.edu.gt/PortalURL/Biblioteca/Contenido.aspx?o=6708&s=49

[9] M. E. Thomae, J. E. Cumming, and I. Fujinaga, "The mensural scoring-up tool," in *Proceedings of the 6th International Workshop on Digital Libraries for Musicology (DLfM)*. The Hague, Netherlands: ACM, pp. 9–19.

[10] I. Fujinaga, "Single interface for music score searching and analysis (SIMSSA) project: Optical music recognition workflow for neume notation," in *Proceedings of the Computers and the Humanities Symposium (JinMonCom)*. Osaka, Japan: Information Processing Society of Japan, 2019, pp. 281–286.

[11] I. Fujinaga and A. Hankinson, "Single interface for music score searching and analysis (SIMSSA)," in *Proceedings of the 1st International Conference on Technologies for Music Notation & Representation (TENOR)*, Paris, France, 2015, pp. 109–15.

[12] G. Burlet, A. Porter, A. Hankinson, and I. Fujinaga, "Neon.js: Neume editor online." in *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR)*, Porto, Portugal, 2012, pp. 121–126.

[13] J. Regimbal, Z. McLennan, G. Vigliensoni, A. Tran, and I. Fujinaga, "Neon2: A Verovio-based square-notation editor," in *Proceedings of the Music Encoding Conference (MEC)*, Vienna, Austria, 2019.

[14] C. Wick and F. Puppe, "OMMR4all – A semiautomatic online editor for Medieval music notations," in *Proceedings of the 2nd International Workshop on Reading Music Systems (WoRMS)*, Delft, Netherlands, 2019, pp. 31–34.

[15] L. X. Pugin, "Editing Renaissance music: The Aruspix project," *Beihefte zu editio*, pp. 64–103, 2009.

[16] J. M. Iñesta, D. Rizo, and J. Calvo-Zaragoza, "MuRET as a software for the transcription of historical archives," in *Proceedings of the 2nd International Workshop on Reading Music Systems (WoRMS)*, Delft, Netherlands, 2019, pp. 12–15.

[17] D. Rizo, J. Calvo-Zaragoza, and J. M. Iñesta, "MuRET: A music recognition, encoding, and transcription tool," in *Proceedings of the 5th International Conference on Digital Libraries for Musicology (DLfM)*. Paris, France: ACM, 2018, pp. 52–56.

[18] J. M. Iñesta, P. J. P. de León, D. Rizo, J. Oncina, L. Micó, J. R. Rico-Juan, C. Pérez-Sancho, and A. Pertusa, "Hispamus: Handwritten Spanish music heritage preservation by automatic transcription," in *Proceedings of the 1st International Workshop on Reading Music Systems (WoRMS)*, Paris, France, 2018, pp. 17–18.

[19] D. Rizo, N. P. León, and C. S. Sapp, "White mensural manual encoding: from Humdrum to MEI," *Cuadernos de Investigación Musical*, no. 6, pp. 373–393, 2018.

[20] P. Roland, A. Hankinson, and L. Pugin, “Early music and the Music Encoding Initiative,” *Early Music*, vol. 42, no. 4, pp. 605–11, November 2014.

[21] K. Desmond, A. Hankinson, L. Pugin, J. Regimbal, C. S. Sapp, and M. E. Thomae, “Next steps for Measuring Polyphony – A prototype editor for encoding mensural music [poster],” in *Proceedings of the Music Encoding Conference (MEC)*, E. De Luca and J. Flanders, Eds. Boston, MA: Humanities Commons, 2020, pp. 121–124.

[22] A. Plaksin and D. Lewis, “Merit: An interactive annotation tool for mensural rhythms,” in *Proceedings of the 9th International Conference on Digital Libraries for Musicology (DLfM)*. Prague, Czech Republic: Association for Computing Machinery, 2022, pp. 55–59. [Online]. Available: https://doi.org/10.1145/3543882.3543888

[23] K. Desmond, L. Pugin, J. Regimbal, D. Rizo, C. S. Sapp, and M. E. Thomae, “Encoding polyphony from Medieval manuscripts notated in mensural notation [panel],” in *Proceedings of the Music Encoding Conference (MEC)*, S. Munich and D. Rizo, Eds. Alicante, Spain: Humanities Commons, 2021.

[24] P. Schubert, *Modal counterpoint, Renaissance style*. Oxford University Press, USA, 2008.

[25] M. E. Thomae, “The Guatemalan choirbooks: Facilitating preservation, performance, and study of the colonial repertoire,” in *Christian Music Traditions in the Americas*, A. Shenton and J. Smolko, Eds. New York: Rowman & Littlefield, 2021.

A DATASET OF SYMBOLIC TEXTURE ANNOTATIONS IN MOZART PIANO SONATAS

Louis Couturier[1] **Louis Bigo**[2] **Florence Levé**[1,2]
[1] MIS, Université de Picardie Jules Verne, Amiens, France
[2] CRIStAL, UMR 9189 CNRS, Université de Lille, France
{louis.couturier, florence.leve}@u-picardie.fr, louis.bigo@univ-lille.fr

ABSTRACT

Musical scores are generally analyzed under different aspects, notably melody, harmony, rhythm, but also through their texture, although this last concept is arguably more delicate to formalize. Symbolic texture depicts how sounding components are organized in the score. It outlines the density of elements, their heterogeneity, role and interactions. In this paper, we release a set of manual annotations for each bar of 9 movements among early piano sonatas by W. A. Mozart, totaling 1164 labels that follow a syntax dedicated to piano score texture. A quantitative analysis of the annotations highlights some characteristic textural features in the corpus. In addition, we present and release the implementation of low-level descriptors of symbolic texture, that are preliminary experimented for textural elements prediction. The annotations and the descriptors offer promising applications in computer-assisted music analysis and composition.

1. INTRODUCTION

1.1 Texture and symbolic texture

Musical texture generally refers to two distinct levels of abstraction used to describe musical content [1]. On the one hand, *sound* related texture, that can be referred to as *orchestral texture*, results from orchestration, instrumentation and timbral characteristics of instruments and performances. On the other hand, *symbolic texture*, or *compositional texture*, results from the organization of notes, chords and voices in the musical score. Naturally, these notions are closely related, and Hérold studies both textural and instrumental factors of timbre [2], highlighting the impact of compositional texture on the final sound field. Symbolic texture, which is the focus of the dataset presented in this paper, can be described through high level musical concepts such as layer separation, diversity of sonic activities, layer roles, note density and interactions [3–7]. For instance, in piano music, an accompanying chord sequence can be performed in various ways, each one being identified by a specific texture contributing to a stylistic identity. Symbolic texture stands at the center of the compositional process and can sometimes be understood as a notion of style [8]. Musical style is however commonly associated with a whole piece [9] or section whereas the notion of symbolic texture considered in this work tends to describe much shorter time spans in the musical score.

1.2 Related work

Computational methods to analyze symbolic texture have been elaborated in various musical styles including classical string quartet music [10] and modern popular guitar music [11]. Given the wide range of playing modes offered by the instrument, piano music brings unprecedented challenges for the task of in-score texture identification. The present work builds on a recent formal syntax elaborated to classical piano music modeling [7]. As a crucial parameter of musical style, symbolic texture has also recently raised an important attention in the tasks of music generation [12] and style transfer [8,13], where musical texture is efficiently learned by deep neural networks but with limited perspectives of explicit categorization and musicological interpretation. Alternatively, the dataset proposed in the present work promotes a pedagogical and transparent expression of symbolic piano texture, aiming at facilitating its use to design computational tools intended to assist music pedagogy, analysis and composition.

The MIR – Music Information Retrieval – community dedicates an important part of its work effort in building musical expert annotations accompanying music datasets to facilitate training and evaluation of computational models. The classical repertoire, in particular its piano music, gives raise to a number and variety of annotation needs given the richness of its musical language. Key regions, cadences, phrases and harmonies were annotated in the Mozart piano sonatas from the New Mozart Edition (Neue Mozart Ausgabe) [14]. As other representative initiatives, the TAVERN dataset includes harmony and phrase annotations on Mozart and Beethoven's piano variations [15] and the Fugue dataset includes form annotations specific to this genre [16]. Beyond piano music, classical string quartets have also raised a number of key, harmony and structure annotation efforts including repertoires of Haydn [17], Mozart [18] and Beethoven [19, 20]. Although a corpus has been proposed in [21], symbolic texture has still rarely

Attribution: L. Couturier, L. Bigo, and F. Levé, "A Dataset of Symbolic Texture Annotations in Mozart Piano Sonatas", in *Proc. of the 23rd Int. Society for Music Information Retrieval Conf.,* Bengaluru, India, 2022.

Figure 1. Examples of texture annotations using the syntax defined in [7]. a) K. 283.III m.1-2: the melody is doubled at the third (`M2p`), and moves in parallel motions (`p`) over a static layer with repeated notes (`S1r`); b) K. 279.I m.35: in addition to a melodic layer with scale motive (`M1s`), a short (sparse, '_') homorhythmic layer appears, without affecting the vertical global density whose value remains 1 on the overall measure (`1[...]`); c) K. 279.I m.6: a typical example of Alberti bass, described as `HS1` harmonic and static layer of density equal to one, and a possible division into two sublayers; d) K. 279.III m.157-158: a concluding formula with melodic (horizontal movements), harmonic (verticality) and static (regularity and emphasis), high vertical density, with an octave motion (`o`) optionally detailed in the sublayer decomposition. A comma separates the dense texture from the contrasting silence ('_') in the last measure.

been the subject of consequent corpus annotations, due to the variety of musical features it involves and the rarity of formal specification.

1.3 Motivation and outline

In order to provide to the community consistent data to study symbolic texture in Western classical piano music, we release a dataset of manual annotations describing symbolic texture at each bar of 9 movements of Mozart Piano Sonatas, totaling a set of 1164 annotated measures. The corpus and the annotation process are detailed in Section 2. Section 3 provides statistics on the textural labels annotated in the dataset. Finally Section 4 presents preliminary results of texture prediction by a machine learning model.

2. PRESENTATION OF THE DATASET

2.1 Syntax for the annotations of symbolic texture

We follow the syntax proposed in [7] to describe textural properties of piano music. More precisely, the texture of a score region is annotated by a text label expressing a set of features with the following conventions:

Diversity. The overall texture is split into independent textural layers that are described individually, separated by a `/`. They are ordered by descending register.

Example: The label `M1/H2` includes two layers, as in examples a, b and c in Figure 1.

Function. The *function* of each layer is expressed with a combination of three specific labels being `M` for *melodic* function, `H` for *harmonic* function, and `S` for *static* function like pedals and ostinati.

Example: The label `M1/H2` includes one layer with a melodic function and one layer with an harmonic function. As an other example, the label `HS1` includes one single layer having both a harmonic and a static function like a persistent arpeggio.

Density. The *density* of a layer, also called *thickness* [1], corresponds to the number of voices it includes, expressed by an integer right after the function.

Example: The label `M1/H2` includes one melodic layer with one voice and one harmonic layer including two voices.

Global density. The *global density* of a region corresponds to its global number of voices and is indicated before brackets surrounding the whole label. It is an approximation of the average number of notes perceived simultaneously. In most cases, the global density is equal to the sum of the layer's density, in which cases its notation is optional because redundant.

Example: The label `3[M1/H2]` indicates a global density of 3 and can be simplified into `M1/H2`. However, the label `2[M1/H2]`, which can occur in certain types of sparse regions, cannot be simplified.

Internal organization of a layer. Additional elements can indicate the presence of relationships between voices: homorhythmy (`h`), parallel motion (`p`), octave (`o`) or characteristic musical figures: sustained notes (`t`), repeated notes (`r`), oscillations (`b`) or scales (`s`).

Example: The label `M2p/HS3hr` can describe a melody doubled at the third accompanied by repeated three-note chords.

Sublayer decomposition. Each layer can optionally be decomposed into sublayers notated between parentheses.

Example: In the label `M1/HS1(S1/M1)` (see Figure 1.c), the second layer `HS1` itself includes one *static* sublayer and one *melodic* sublayer enabling for example the expression of a single voice accompaniment consisting in an alternation between a single repeated pitch and a moving melodic line.

Sparsity. When a layer does not last during the full measure (as in Figure 1.b), or has too low horizontal density, it is considered as *sparse* and notated with '_'. This symbol is also used to annotate empty bars.