Guatemalan Cathedral Choirbooks: From Manuscript to Digital Images to Edited Symbolic Scores

Martha Eladia María Thomae Elías
(Martha E. Thomae)

Music Technology Area
Department of Music Research
Schulich School of Music
McGill University, Montreal

August 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Abstract

The legacy of colonialism is vast. Spain and Portugal established the first colonial empires in the Americas during the 16th century. Latin-American nation-states trace their origins to the administrative divisions of these colonial empires. Most of these countries are developing nations, and the poorest of them do not have the resources to preserve and disseminate their cultural legacy, including the early music produced within their territories as the result of cultural transmission from colonizer to colonized. Despite being known by musicologists and being part of the nation's history and cultural identity, these musical works remain largely inaccessible because they survive in unique manuscripts stored in archives that are difficult to access and in an earlier form of music notation readable only by experts. I address these issues using affordable digitization and music-encoding tools.

Choral music of the sixteenth century was written using mensural notation and a separate-parts layout. This tradition was carried into the newly acquired colonies of the Americas, where it persisted for over two centuries. The separate-parts layout of polyphonic books, where each voice is contained in a separate area of the book opening or in a separate partbook, does not show the vertical sonorities. Transcribing the music into a score makes it easier to analyze by modern scholars. Still, it is difficult and time-consuming, due to the complexities of the notation regarding rhythmic interpretation. Computers can simplify the process, transcribing the music into machine-readable scores, also known as symbolic scores. Symbolic scores have many advantages over physical scores: when rendered, they can be performed from, just like traditional paper scores, but they also allow for content-based search and computational research on early music. Furthermore, they enable the preservation of the music content on relatively small files and increase access to it

as the machine can render the encoded music in notation, play it back, and automatically transcribe it into modern notation.

In this dissertation, I present a workflow that allows for semiautomatically retrieving digital images and symbolic scores from physical sources. I test the workflow with one of the choirbooks of polyphonic sacred music from the Archdiocesan Historical Archive of Guatemala—a collection of Catholic sacred works composed from the 16th to the 18th centuries by composers from Europe and the Americas. Anyone interested in retrieving symbolic scores of mensural sources can use this workflow because it relies on do-it-yourself (DIY) processes and free online technologies, reducing barriers to accessibility for Latin-American musicologists and archives with limited resources. The workflow consists of three stages: (1) **digitization**, creating digital images of each of the pages of the manuscript using a *DIY book scanner*; (2) **optical music recognition (OMR)**, that encodes the music symbols in each of the voices of the choirbook; and (3) **automatic voice alignment and editorial correction**, that interprets the rhythmic values of the notes in mensural notation, combines them in a correctly aligned score, and provides a symbolic file that can be used for study or performance. I also describe how the time spent in the last stage is reduced to make the entire process economical. The results will be made publicly available.

The encoding tools presented in this dissertation can be used on all kinds of music in mensural notation, allowing for the semi-automatic transcription of mensural sources into symbolic scores. The workflow presented has the potential to significantly expand the access to music of the 16th century. This dissertation will enable other institutions to preserve and disseminate their musical heritage, allowing musicologists to understand the transmission of music from Europe to the Americas, or to other countries with a colonial legacy.

Résumé

L’héritage du colonialisme est vaste. L’Espagne et le Portugal ont établi les premiers empires coloniaux dans les Amériques pendant le 16e siècle. Les origines des états-nations latino-américains remontent aux divisions administratives de ces empires coloniaux. La plupart de ces pays sont aujourd'hui en développement. Les plus pauvres d’entre eux n’ont pas les ressources pour préserver et diffuser leur héritage culturel, incluant la musique ancienne produite sur leurs territoires résultant de la transmission culturelle de la nation colonisatrice à la nation colonisée. Malgré qu'elles soient reconnues par les musicologues comme faisant partie de l'identité historique et culturelle de la nation, ces œuvres musicales demeurent largement inaccessibles, survivant dans des manuscrits uniques entreposés dans des archives difficiles d'accès et écrites dans une forme obsolète de notation musicale qui ne peut être lue que par les experts. J'aborde ces problèmes en utilisant des outils de numérisation et d'encodage de la musique.

La musique polyphonique du 16e siècle était écrite en notation mensurelle et en parties séparées. Cette tradition a été transmises aux nouvelles colonies des Amériques, où elle a perduré pendant plus de deux siècles. La disposition en parties séparées dans les livres de chœur polyphoniques, où chaque voix se trouve dans un endroit différent sur la page ou dans un autre livre entièrement, ne montre pas les sonorités verticales. La musique est plus facile à analyser par les experts d'aujourd'hui si elle est transcrite en une partition complète. Toutefois, ce processus est long et difficile, parce que la notation rythmique est complexe. Un ordinateur peut simplifier le processus en transcrivant la musique en partitions lisibles par une machine, appelées "partitions symboliques" (*symbolic scores*). Ces partitions présentent de nombreux avantages par rapport aux partitions physiques : lorsqu'elles sont converties, elles peuvent servir à la performance, tout comme des partitions papier traditionnelles, mais elles permettent aussi la recherche basée sur le

contenu et la recherche informatique sur la musique ancienne. De plus, elles permettent de préserver le contenu musical dans des fichiers relativement petits et facilitent leur accès, puisqu'une machine peut restituer la musique encodée en notation, la rejouer et la transcrire automatiquement en notation moderne.

Dans cette thèse, je présente un flux de travail qui permet de récupérer semi-automatiquement des images numériques et des partitions symboliques à partir de sources physiques. Je teste le flux de travail avec un des livres de chœur de musique polyphonique sacrée des Archives archidiocésaines historiques du Guatemala, une collection d'œuvres catholiques sacrées composées entre les 16^{e} et 18^{e} siècles par des compositeurs d'Europe et des Amériques. Quiconque est intéressé à créer des partitions symboliques de sources mensurelles peut utiliser ce flux de travail, parce qu'il fonctionne à l'aide de processus *do-it-yourself* et de technologies en ligne gratuites, améliorant l'accessibilité à ces ressources pour les musicologues latino-américains et pour les archives aux ressources limitées. Le flux de travail consiste en trois étapes : (1) la **numérisation**, qui crée des images numériques de chaque page du manuscrit à l'aide d'un *DIY book scanner* ; la **Reconnaissance optique de musique** (ROM), qui encode les symboles musicaux composant chaque voix du livre de chœur ; et (3) **l'alignement automatique des voix et la correction éditoriale**, qui interprète les valeurs rythmiques des notes en notation mensurelle, les assemble dans une partition correctement alignée et fournit un fichier informatique qui peut être utilisé pour l'étude ou pour la performance. Je décris aussi comment le temps de correction finale peut être réduit pour rentre le processus complet économique. Les résultats seront rendus disponibles au public.

Les outils d'encodage présentés dans cette thèse peuvent être utilisés avec toutes sortes de musique en notation mensurelle, permettant la transcription semi-automatique de sources mensurelles en partitions symboliques. Le flux de travail présenté a le potentiel d'améliorer de façon significative l'accès à la musique du 16e siècle. Cette thèse permettra à d'autres institutions de préserver et diffuser leur patrimoine musical, permettant ainsi aux musicologues de comprendre la transmission de la musique de l'Europe aux Amériques, ou vers d'autres pays ayant un héritage colonial.

Acknowledgements

I would like to thank all my colleagues and friends for their support during this dissertation. I have been fortunate to be part of the Distributed Digital Music Archives and Libraries (DDMAL) Laboratory during my work and count on the expertise of its members. Special thanks go to Geneviève Gates-Panneton for the translation of the abstract into French and for her awesome job during the experiment, and to my colleagues and friends Gabriel Vigliensoni, Emily Hopkins, Néstor Nápoles López, and Timothy de Reuse for the revision of different sections of this dissertation or publications based on this research. I extend my gratefulness to my DDMAL colleague and friend Yaolong Ju, although he was not around the Lab anymore to read my dissertation, he has always provided me with words of encouragement and great chit-chat. Also, thanks to my McGill colleagues and friends Matan Gover, Samuel Howes, and Sylvain Margot for their support during my PhD studies, especially for their help with many counterpoint questions when preparing for my comprehensive exams and my dissertation defence. Finally, thanks to my wonderful husband, Andrés Lou Ríos, for his editorial work on the English version of the abstract.

Special thanks to my supervisors, Julie Cumming and Ichiro Fujinaga, without whose guidance and support this work and my studies would not have been possible. They have provided me with so many opportunities to deepen my knowledge of digital musicology and music encoding. I would also like to thank Peter Schubert for his advice on counterpoint-related questions.

This project would not have been possible without the support of many people involved in the different stages of this multi-disciplinary project. I want to thank everyone involved in the digitization part of the project. First, I want to thank the Archivo Histórico Arquidiocesano de Guatemala (AHAG), especially the chancellor of the Ecclesiastical Curia of Santiago de

Guatemala and director of the AHAG, Father Eddy René Calvillo, for his approval of the project. No work could have been done without Father Calvillo's and the AHAG's support. I would also like to thank musicologist Omar Morales Abril and archivist Alejandro Conde for providing information about the Guatemalan choirbook collection and the AHAG, besides great advice along the way.

I am also grateful to all the institutions and people that helped putting the digitization equipment together. Thanks to the DDMAL Lab and Professor Fujinaga, its Principal Investigator and my supervisor, for providing the lighting equipment, reference targets, hard drives, memory cards, and camera charger. Thanks to the Marvin Duchow Library and specially to Cathy Martin and Gabrielle Kern for lending the camera used to photograph the manuscript. Thanks to my father, José German Thomae Villela, for building the book cradle according to my specifications. And thanks to Darryl Cameron, Chief Electronics Technician in the Music Technology area of McGill, for lending me the copystand used.

I also want to thank all the experts that provided advice prior to or during the digitization process. I had the support of experts from three institutions involved in digitization projects of special collections, cultural heritage material, and music manuscripts. These are the McGill Library's Digital Initiatives, the Bibliothèque et Archives nationales du Québec (BAnQ), and the Digital Image Archive of Medieval Music (DIAMM). I want to thank Gregory Houston, from McGill's New Media & Digitization Administrator of the Digital Initiatives Department, for his suggestion regarding websites where I could obtain archival supplies (e.g., snake weights) and suggestions about how to improve the book scanner in the future. I will be always grateful for the support of three BAnQ experts: Jessica Régimbald, conservator at the direction du dépôt légal et de la conservation des collections patrimoniales; Marie-Chantal Anctil, coordinator of the section de la reproduction et des ateliers audiovisuels and direction de la numérisation; and Michel Legendre, photographer of the direction de la numérisation and section de la reproduction. The three of them received me on multiple ocassions at the BAnQ for consulting on conservation- and digitization-related questions. Régimbald advised me on how to handle the manuscript during digitization to avoid any damage, gave me a list of common conservation treatments prior to digitization, and provided me with information about platens for me to make an informed decision on whether to use them. I received advice from Anctil about the most important things to consider when building a do-it-yourself (DIY) book scanner. Legendre and Anctil were the ones who proposed the open-at-one-side cradle configuration to facilitate keeping the photographed page flat, which is the configuration that I ended up using. Legendre also advised me about the light selection, approving the lighting equipment chosen. Finally, even though I could not have a face-to-face communication with the experts from DIAMM, I could not have conducted this project without their advice. I am so thankful to DIAMM's project manager, Julia Craig-McFeely, and the lead photographer, Lynda Sayce, for all the email exchanges during the project. Most of the things I did during digitization are based on their advice, including the use of a neutral background (the black felt sheet), the placement of the color patch in a stack of cardboard pieces (painted black to match the background), and the lack of a platen. Thanks to Professor Eun Park for her great preservation management course at McGill, through which I learned a lot and I got in contact with the BAnQ and McGill Library personnel, and in which I met wonderful people that provided me with further advice and encouragement.

I also want to thank the Centro de Rescate, Estudios y Análisis Científico del Arte (CREA) for their evaluation of the conservation status of the choirbooks, which helped me choose the book I would digitize in this pilot project, and for giving conservation treatment to that choirbook prior

to digitization. I am also thankful to the Rozas-Botrán foundation from which CREA is a part of (this foundation manages social development projects in health and culture in Central America and Panama). And special thanks to CREA's lead conservator, Lucy González Muñoz, who enthusiastically accepted the project. I would not have been able to find out about CREA's paper-conservation work if it was not for various contacts in Guatemala that proactively made calls and consulted with other experts. So I am extending my gratitude to the stained glass restorer at the Dirección General del Patrimonio Cultural y Natural from the Ministerio de Cultura y Deportes in Guatemala City, Virginia Golcher; my aunt, Ana Miriam López, who was doing a BA in the Restoration of Artefacts and consulted various of her professors at the Universidad de San Carlos de Guatemala; to my aunt-in-law, Karla Lou, who put me in contact with Thelma Castillo, Art and Culture consultant at the Rozas-Botrán foundation; and to Thelma Castillo herself for contacting CREA's lead conservator (Lucy González Muñoz) and informing her about the project before my visit.

Finally, I want to thank the Centro de Investigaciones Regionales de Mesoamérica (CIRMA), the institution that holds the microfilms of many of the music assets of the AHAG. Special thanks to CIRMA's general director, Guisela Asensio Lueg; director of CIRMA's historical archive, Thelma Porres Morfin; and its photo library assistant, María de los Ángeles Ávila de León. Porres Morfin show me the microfilms that contain part of the music collection of the AHAG and provided me with background information about their history. Photographer Ávila de León offered to help me for free during the digitization process. At the end, due to scheduling issues, she was not the photographer that assisted me. However, CIRMA's general director (Asensio Lueg) was the one who provided me with the contact of the photographer that I hired for the project: Daniel Hernández-Salazar. Hernández-Salazar set up the camera parameters during

the first session and provided me with guidance and useful advice (e.g., including the verso number in the photograph itself by printing the number on a small piece of paper placed at one side of the color patch).

After the digitization process was completed, I performed optical music recognition on the images using MuRET. I want to thank everyone behind MuRET's development. Special thanks to David Rizo, its lead developer; Jorge Calvo-Zaragoza and Antonio Ríos-Vila, for the preparation of the symbol recognition models; Francisco (Paco) Castellanos, for the preparation of the document analysis model; and José Manuel Iñesta as the principal investigator. I was in constant communication with David and Jorge when finding issues, and I appreciate all the support they gave me, both in terms of technical support and friendship.

After concluding the OMR, I used the MP Editor to automatically lined-up the voices and provide editorial corrections in the score. I am thankful to the team behind the MP Editor: lead investigator Karen Desmond, lead developer (and friend of mine) Juliette Regimbal, Laurent Pugin, Andrew Hankinson, and Craig Sapp. Special thanks to Juliette and Craig as they worked in close collaboration with me and with David Rizo (from MuRET) to provide support for features needed in the MP Editor to conduct this research. I am also thankful to Alex Morgan for his development of humlib's Renaissance dissonance filter, which I used to correct more easily the results of the scored-up OMR files.

I am also grateful to the Music Encoding Initiative (MEI) community for their support with my questions and pull requests, and for always being a very friendly and welcoming community. Special thanks to Anna Plaksin, Giuliano Di Bacco, Stefan Münnich, Elsa De Luca, David Lewis, Claire Arthur, Anna Kijas, Johannes Kepper, Laurent Pugin, and Andrew Hankinson for their kind words. Their kindness, cheering, and friendship were very much welcome and needed in this pandemic time. I extend my gratitude to my friend and McGill colleague Behrad Madahi.

Thanks to the people that (unaware of it) guided me to McGill and to this line of research: Juan Pablo Pira Martínez, who introduced me to mensural notation when I wanted to write an interdisciplinary Undergraduate thesis for my BS in Math; and Dr. Raúl Rafael González Álvarez (R.I.P.), who helped me achieve this goal and introduced me to Pira. I am where I am thanks to you!

Last, but not least, thanks to my family for their support, my mom Ana Patricia Elías López, my father José German Thomae Villela, and my sister Anneliese María Thomae Elías. In addition to their encouragement and love, they all helped me during the digitization work. While my father built the book cradle, my mother and sister recorded the videos that report the configurations of the cradle and the digitization process (now available in a YouTube playlist). Finally, all the thanks and love in the world to my husband Andrés Lou Ríos for his support and infinite patience during this doctorate; and to my research companion, my kitty-cat Ludmila, who kept me company over many coding, reading, and writing nights.

This research has been supported by Fonds de recherche du Québec – Société et culture (FRQSC), Bourse au doctorat en recherche (13D - Musique) 2019-B2Z- 261749.

Preface

This dissertation, and the research to which it refers, is the candidate's own original work except for commonly understood and accepted ideas, or where explicit reference to the work of other people, published or otherwise, is made. This dissertation is organized as a monograph, presented in six chapters containing some content previously published in peer-reviewed papers at academic conferences on music encoding, digital musicology, and music information retrieval, and a peer-reviewed chapter in a musicology book (all of them cited in this dissertation):

- **Book chapter:** Thomae, Martha E. 2021. "The Guatemalan Choirbooks: Facilitating Preservation, Performance, and Study of the Colonial Repertoire." In *Christian Music Traditions in the Americas*, edited by Andrew Shenton and Joanna Smolko. New York: Rowman & Littlefield.

- Two contributions in the published **panel session**: Desmond, Karen, Laurent Pugin, Juliette Regimbal, David Rizo, Craig Sapp, and Martha E. Thomae. 2021. "Encoding Polyphony from Medieval Manuscripts Notated in Mensural Notation." In *Proceedings of the Music Encoding Conference*, edited by Stefan Münnich and David Rizo, 197–219. Alicante, Spain (online): Humanities Commons. https://doi.org/10.17613/tf2j-x697.

  These **contributions** are:

  - Thomae, Martha E., David Rizo, and Juliette Regimbal. "From OMR to the Measuring Polyphony Editor: Possibilities of Inter-change."

  - Thomae, Martha E., Craig Sapp, and Juliette Regimbal. "Interaction between Measuring Polyphony MEI Scores and Humdrum Analysis Tools."

- **Paper:** Thomae, Martha E., Julie E. Cumming, and Ichiro Fujinaga. 2022a. "Digitization of Choirbooks in Guatemala." In *Proceedings of the 9th International Conference on Digital Libraries for Musicologists (DLfM)*, 19–26. Prague, Czech Republic: ACM. https://doi.org/10.1145/3543882.3543885.
- **Paper:** Thomae, Martha E., Julie E. Cumming, and Ichiro Fujinaga. 2022b. "Counterpoint Error-Detection Tools for Optical Music Recognition of Renaissance Polyphonic Music." In *Proceedings of the 23rd International Society for Music Information Retrieval Conference (ISMIR)*. Bengaluru, India.

The candidate was responsible for every step involved in designing the experiment of this dissertation and analyzing the results, while the experiment was carried out by research assistant Geneviève Gates-Panneton. The candidate led the work done in tools like MuRET, the MP Editor, Verovio, and the Humlib library during this dissertation. While the candidate implemented some of the necessary changes into the MP Editor, Verovio, and the Humlib library, most of the implementation of this work was done by the respective developers of these tools: David Rizo (for MuRET), Juliette Regimbal (for the MP Editor), Laurent Pugin (for Verovio), and Craig Sapp (for Humlib). The dissertation text makes it clear who implemented which changes in the tools' code base. The candidate also received guidance from multiple experts and institutions for the digitization part of the dissertation as indicated in the Acknowledgements. The text that contains the suggestions of these experts (mostly in Section 2.2.2) clearly acknowledges the expert involved.

Professors Ichiro Fujinaga and Julie Cumming, the candidate's supervisors, provided guidance in different aspects of this dissertation, including experimental planning, data analysis, and interpretation of results. Professor Peter Schubert, an expert in modal counterpoint, also provided guidance on the interpretation of the results of the experiment. Prof. Fujinaga, the candidate's main supervisor, granted advice and resources for the digitization process, and supplied additional funding for some of the digitization equipment and for presenting the research outcome at various conferences. Both supervisors (Cumming and Fujinaga) provided proofreading of the chapters contained in this dissertation.

The candidate prepared the manuscripts of the book chapter (Thomae 2021) and the two papers listed above—by Thomae, Cumming, and Fujinaga (2022a; 2022b) for the Digital Libraries for Musicology (DLfM) and the International Society for Music Information Retrieval (ISMIR) conferences—with proofreading from supervisors Cumming and Fujinaga. The panel publication above (Desmond et al. 2021) contains various contributions, each prepared by a different set of authors and proofread by Karen Desmond; the manuscripts of the two contributions listed within this panel publication were mainly prepared by the candidate, with the first manuscript wrote in conjunction with David Rizo. This contribution by Rizo is reflected in the first half of Section 3.4 about the interoperability of MuRET and the MP Editor, with the text about IIIF manifests and URIs based on Rizo's contribution to Desmond et al. (2021) panel publication.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

During the colonial period (sixteenth to nineteenth centuries), the identity of the countries of Latin America was shaped by the transmission of culture from Europe. This cultural transmission is reflected in their musical heritage, from where many early music sources survive today. Preserving and disseminating the content of these colonial music sources is important, as they are part of the history and identity of a nation. However, the poorest of the former colonies do not have the resources for such enterprise, especially considering that early music books tend to be physically large and written in notation only accessible to experts. In this dissertation, I am providing a workflow consisting of affordable tools to create and retrieve digital images and encoded scores to preserve, study, and increase access to the Early Music heritage of these countries while also reducing the handling of the valuable original sources. I am focusing on the polyphonic early music heritage of the Metropolitan Cathedral of Guatemala City.

1.1 Motivation and Background on Guatemala City's Cathedral Polyphonic Choirbooks

Six large polyphonic choirbooks were in use in Guatemala City's Cathedral between c. 1600 and c. 1800. This collection, known by the sigla GuatC 1–6,[1] is part of Guatemala's colonial past. It was copied in the seventeenth and eighteenth centuries and contains mostly sixteenth-century polyphonic music written in mensural notation (used in Europe c. 1280–1600). It includes works by composers from both Europe and the Americas. This corpus documents a continuous performance tradition of sacred choral music from the Renaissance through the 1800s and, therefore, it is a valuable resource for studying the transmission of music from Europe to Latin America.

[1] GuatC: Guatemala. Guatemala City. Catedral, Archivo Capitular. Other sigla include GCA-Gc.

Currently, the choirbooks are held at the Archivo Histórico Arquidiocesano de Guatemala (Archdiocesan Historical Archive of Guatemala, AHAG). Although the first three choirbooks were inventoried by Pujol (1965), Stevenson (1970), and Snow (1996) and a general overview of the repertoire and history of the whole collection has been provided by Morales Abril (2015), access to the musical contents of these sources is difficult, with the notable exception of the fourth choirbook, transcribed in full into modern notation by Snow (1996). Digital images of the first three books made from microfilms are available; however, these images are of low quality with some pages missing and others cropped, plus there is some controversies about how these digital images were obtained. Access to the books at the AHAG requires special permission. Furthermore, the use of mensural notation restricts performance of the music to experts on the rhythmic intricacies of the notation. Guatemalan musicologists Dieter Lehnhoff and Omar Morales Abril have made efforts to preserve and disseminate these sources. Lehnhoff and Morales Abril have transcribed some of the pieces in this corpus into modern notation (Lehnhoff 1984; 2001; 2008; Morales et al. 2001; Morales Abril 2003) preserving a few pieces of the GuatC collection and facilitating their performance. Both musicologists have also performed this music by conducting a few concerts helping to the dissemination of this music—there is also one recording that includes some of the pieces of this corpus (*Coros de Catedral* 1995). However, these efforts cover only a small part of the corpus.

Transcription of this music into modern scores—as done by Snow, Lehnhoff, and Morales—is important for their dissemination and study by modern scholars. The original mensural sources tend to use a separate-parts layout, where each voice lies on a different area of the page (e.g., in choirbook format) or different book (e.g., in partbooks). Scholars nowadays are not used to deal with music that is in parts. A modern score, on the other hand, facilitates visualizing both the melodic information (also present in the separate-parts layout) and the vertical sonorities, which allow modern scholars to study counterpoint. However, transcribing mensural music into scores is also a time-consuming task, especially given the context-dependent duration of the note values in triple meter. Machines can help in this transcription task, producing symbolic scores. These are machine-readable scores, which have many advantages over physical scores. In addition to allow for performance (when rendered) as a physical score does, a symbolic score allows for content-based search and computational music analysis. It also allows for preserving the music in a small file and increasing access to the music as the machine can read the file and render it in notation, play it back, automatically transcribe it into modern notation, and compare it to the same piece in concordant sources—provided these are encoded too—among other things.

My goal is to preserve the GuatC choirbooks and increase access to them through digital images, metadata, and symbolic scores. In the past, I have worked on projects involving the processing of mensural notation with my Master's thesis focusing on the development of a scoring-up tool to interpret mensural notation and automatically present the voices in a score layout, and other projects concerning the retrieval of symbolic files from digital images of music documents by preparing and using document analysis and symbol recognition models for chant manuscripts as part of my Research Assistantship at the Distributed Digital Music Archives and Libraries (DDMAL) Lab at McGill University. I intend to apply the acquired knowledge in these projects for the preservation and access of the Metropolitan Cathedral choirbooks in Guatemala (my home country). In this dissertation, I am focusing on the first choirbook of the collection (GuatC 1).

I will digitize the folios of the GuatC 1 (obtaining high-quality, colored digital images), create symbolic scores, and edit them to include corrections for scribal errors. The symbolic scores can be semiautomatically generated from the digital images using Music Information Retrieval

(MIR) technologies, including optical music recognition (OMR), automatic voice alignment for mensural notation, editorial correction software, and computational error detection as will be expanded on the coming sections. The resultant symbolic files will contain all the information for this music to be rendered as a score with its voices lined up—allowing a musician or music analyst to visualize the vertical sonorities occluded in the original layout—and played back. These files will also facilitate scholarly research into the works' musical style and transmission.

1.2 Overview on Digitization, OMR, and Scoring Up Technologies for Early Music Books of Polyphony

Given the antiquity, physical size, and binding of the choirbooks, special equipment and care are necessary for their digitization. According to the Federal Agencies Digital Guidelines Initiative (2016), these books should be digitized from above with a book scanner of the appropriate dimensions. Such a book scanner is not readily available in Guatemala. However, an alternative to be considered is the use of do-it-yourself book scanners.[2]

OMR is commonly used to retrieve the music symbols in digital images of music documents. Currently, there are four primary OMR frameworks for early music notations: Rodan (Hankinson 2014), OMMR4All (Wick and Puppe 2019a), Aruspix (Pugin 2009), and MuRET (Rizo, Calvo-Zaragoza, and Iñesta 2018). MuRET provides the best support for mensural notation in manuscripts. It includes a mensural-symbol-recognition model trained on handwritten Spanish mensural notation—similar to the Guatemalan mensural notation—and tools for correcting the results (Calvo-Zaragoza, Toselli, and Vidal 2019).

[2] http://diybookscanner.org

When dealing with mensural notation, it is not enough to have all the music symbols (rest, pitches, and note shapes) recognized by the OMR system. One still needs to interpret the duration of mensural notes and arrange the parts into a score. My automatic scoring-up tool for mensural notation is a heuristic process that computes the value of each note based on mensural rules extracted from historical music treatises and presents the piece in score layout (Thomae 2017). While other tools that automatically interpret mensural notation have been developed (Huang et al. 2015; Rizo et al. 2017), they do not meet the requirements to transcribe the music in this manuscript.

I have integrated my scoring-up tool into the Measuring Polyphony Editor (MP Editor), an online editor for mensural notation that allows editorial corrections (Desmond, Hankinson, et al. 2020). Editorial corrections—that is, correction of scribal errors—are needed to provide a correct score. These corrections favor motivic repetition, voice imitation (common in Renaissance music), and legal counterpoint. Detection of counterpoint errors (i.e., errors in vertical sonorities and voice leading) can help in revealing scribal errors. Among various computation music analysis tools, only the VIS Framework and the humlib software library have support for analyzing counterpoint. Both the Framework and the library include a tool to label dissonances following the rules of Renaissance counterpoint outlined by Schubert (2008). The tool is known as a "dissonance indexer" in the VIS Framework and as a "dissonant filter" in the humlib library, both were developed by Alexander Morgan.[3] Humlib's dissonance filter has a larger set of dissonance labels than VIS's dissonance indexer, as it was developed later in Morgan's career.

[3] Information about VIS Framework's dissonance indexer can be found at https://vis-framework.readthedocs.io/en/v3.0.5/vis.analyzers.indexers.html#dissonance-module. Information about humlib's dissonant filter can be found in the Verovio Humdrum Viewer page at https://doc.verovio.humdrum.org/filter/dissonant/. And for the Humlib library: https://humlib.humdrum.org/.

1.3 Project Overview: Digitization and Encoding Workflow

I will use the following three-stage workflow and tools to obtain the digital images and symbolic scores:

1. **Digitization.** This stage consists of photographing the pages of the choirbooks. For this task, I will use a *DIY book scanner* following the specifications of international organizations with experience in the digitization of old documents (DIAMM, McGill Library, BAnQ, and FADGI).

2. **OMR.** This stage consists of applying OMR to the digital images to create symbolic files that encode the music symbols of each of the voices in the image. For this purpose, I will use *MuRET* for identifying Guatemalan mensural symbols. The results will be corrected within the MuRET interface, and the symbolic music file encoding the parts (voices) of the piece will be used by the *MP Editor* software in the next stage of the workflow.

3. **Automatic voice alignment and editorial correction.** This stage interprets the duration of the mensural notes, presents the piece as a score, and allows the user to enter editorial corrections. Editorial corrections refer to corrections of scribal errors, which, although less common than OMR errors, tend to be harder to identify, especially when the piece is not displayed as a score. For this stage, I will use the *MP Editor*. I will also implement a counterpoint error-checker to aid in the editorial correction process. I will integrate the humlib's *Renaissance dissonance filter* into the MP Editor and conduct an experiment to evaluate whether its use makes the correction of the scored-up OMR piece more efficient.

1.4 Structure of the Dissertation

The content of this dissertation is divided into six chapters, including this introduction. Chapter Chapter 2 presents the detailed background information to conduct this project. It contains a literature review on the Guatemalan Cathedral polyphonic choirbook collection (the GuatC collection) and their historical background; followed by a summary on the digitization of cultural heritage documents with special focus on large manuscript volumes; then a section on optical music recognition technologies for early music notation; and an introduction to mensural notation, the voice alignment issue, and developed technologies to address this issue.

The methodology is divided into two chapters. While Chapter Chapter 3 presents the digitization and encoding methodology (this is, the workflow summarized in Section 1.3), Chapter Chapter 4 presents the experiment's methodology. Chapter Chapter 3 presents the technologies chosen for completing each stage of the workflow and the work done to encode the collection. It opens with three sections, each related to a stage of the workflow. The first section contains information on the survey for digitization equipment and personnel, and the DIY book scanning technology. The second section focuses on the OMR and provides information about the MuRET OMR framework. The third section presents the MP Editor as solution for the voice alignment issue. The chapter ends with two sections presenting my work to allow for the interaction of these tools. The first of these, the fourth section, focuses on the interaction between MuRET and the MP Editor, which allows to use the output of MuRET as the input of the MP Editor. Finally, the fifth section presents the details about the integration of humlib's Renaissance dissonance filter into the MP Editor, which allows to study one of the key aspects of this dissertation: determining whether the use of a counterpoint error detector makes the process of correcting the scored-up OMR file more efficient. This key aspect is the focus of this dissertation's experiment, presented in Chapters Chapter 4 and Chapter 5.

Chapter Chapter 4 focuses on the preparation of the experiment. It has two sections, one about data preparation, and another summarizing the evaluation procedures and experimenter's tasks. Chapter Chapter 5 is the experiment chapter. It details the outcome of each of the stages of the workflow, with three sections that present the results and discussion of the digitization, OMR, and voice-alignment stages, respectively. The last of these sections focuses on the experiment with the dissonance filter.

Finally, Chapter Chapter 6 presents the conclusions of this dissertation, summarizing the work done and the experiment implications. It also provides details on what worked and what could be improved in each of the stages of the workflow, and it summarizes updates and recommendations that could speed up the process in the future. The chapter ends with a list of contributions and future work. Additional documents, such as the archive's letter of permission to conduct this project and the conservation report of the GuatC books, are provided in the Appendix.

1.5 Research Contributions

This research will help my Guatemalan community preserve and promote its musical heritage by producing digital surrogates of the original manuscripts, including both images and symbolic files, thus reducing the handling of the valuable original sources. The images provide high-quality access to the manuscript source, and the machine-readable symbolic files preserve the music content on relatively small files and increase access to it by allowing for playback, render in original notation, content-based search, computational music analysis, and automatic translation into modern notation for performance. I am designing and testing an economical methodology for digitization and encoding of mensural musical heritage that can be used by other archives in Latin America or anywhere in the world. To make this methodology as practical and affordable as possible, the MIR includes only free, online, open-source software, and I have documented the details of the DIY book scanner setup.

Any interested scholar can use the encoding tools of this workflow to semi-automatically transcribe the music of all kinds of mensural sources into symbolic scores. This workflow has the potential to massively expand the access to sixteenth-century music. The applicability of this workflow to other mensural music sources will enable the study of the transmission of the Western music tradition across the former European colonies and identify local composition practices. The project will contribute to preserving a historical musical heritage that otherwise may be lost, damaged, or forgotten.

Chapter 2 Background

This chapter contains the literature review of this dissertation, divided into four sections. Section 2.1 presents the corpus I am working with, the Guatemalan Metropolitan Cathedral polyphonic choirbooks. It provides the historical background of the Cathedral and the choirbook collection (GuatC), followed by details on scholarly research of this collection, and specifics about the first choirbook (GuatC 1), which is the focus of this project. The following three sections present the literature review of each of the three stages involved in the digitization and music-encoding workflow introduced in Section 1.3. Section 2.2 focuses on the digitization of cultural heritage documents, starting with the basic concepts and finalizing with a summary of recommendations and imaging technologies suitable for rarebooks of large dimensions. Section 2.3 focuses on optical music recognition of early music, presenting the basics of OMR, recent developments on the area, and a summary of OMR tools for early music. Finally, Section 2.4 presents the basics of mensural notation, the voice alignment issue that characterize this notation, with computational tools developed to address this issue, and a summary of computational music analysis tools.

2.1 The Guatemalan Cathedral Choirbooks

This section presents the choirbook collection that is the focus of this dissertation, the Guatemalan Cathedral choirbooks (GuatC).[4] Before getting into the details of the collection, Section 2.1.1 provides a brief history of the Guatemalan Metropolitan Cathedral and the use of sacred music within its walls. Section 2.1.2 presents the GuatC choirbook collection, covering previous work and detailing its repertoire, copyists, and composers. Finally, Section 2.1.3 focuses on the first of

[4] The *GuatC* abbreviation comes from the Census Catalogue (Call et al. 1979–1988, 1:259). The *GCA-Gc* is another abbreviation used for this manuscript collection, as can be found in the Digital Images Archive of Medieval Music (DIAMM) website: https://www.diamm.ac.uk/search/?q=guatemala.

these choirbooks, the book of masses GuatC 1, that I digitized and encoded as part of this dissertation.

2.1.1 Brief History of Guatemala City's Cathedral and its Music

Guatemala's European sacred music tradition began in its second capital, *Santiago de los Caballeros de Guatemala* (today *Ciudad Vieja* or Old City), founded in 1527 in the Almolonga Valley on the outskirts of the *Volcán de Agua* (Water Volcano).[5,6] This started with the construction of the church, eventually promoted to the status of cathedral. In 1530, the Spanish priest Francisco Marroquín arrived in *Santiago* from Mexico, having been convinced to come by conquistador Pedro de Alvarado (Pérez de Antón 1999). The priest oversaw the construction of the church, which was finished in 1534 (Lehnhoff 1986, 32–33). In the same year, Pope Paul III issued the papal bull that established the diocese of Guatemala and raised the parish church in Alvarado's capital to the dignity of cathedral, appointing Marroquín as its first bishop ("Libro de Actas Capitulares 1" 1531, 109; Lehnhoff 1986, 34; Snow 1996, 2).[7] However, Marroquín was not consecrated as bishop until 1537 by the Archbishop of Mexico, Juan de Zumárraga (Sáenz de Santa María 1964, 35). He then established the *cabildo eclesiástico* (cathedral chapter, the governing board of the church) and defined the roles of *chantre* (cantor) and organist, among other

---

[5] According to Recinos (1980, 102), soon after the Spanish conquest of Guatemala began in 1524, conquistador Pedro de Alvarado declared the Kaqchikel Maya capital Iximche as the first capital of Guatemala on 25 July—although Schele and Mathews (1999, 297) indicate that this event took place on 27 July. Alvarado called the new capital "Santiago de los Caballeros de Guatemala" ("St. James of the Knights of Guatemala"). The date the capital was founded (25 July) was also the Día de Santiago (Feast of Saint James, the Great), Patron Saint of Spain. After continuous Kaqchikel attacks, the Spanish refounded the capital in 1527 in the Almolonga Valley near the Volcán de Agua (Recinos 1980, 105).

[6] It is worth mentioning that this dissertation focuses on cathedral music, which was destined for Spanish and criollos. The catechization of Guatemalan indigenous people was conducted outside the cathedral, in the countryside. Music codices found in the department of Huehuetenango (at northwestern Guatemala)—specifically from the towns of Santa Eulalia, San Juan Ixcoy, San Mateo Ixtatán, and Jacaltenango—prove the development of the European music practice in the indigenous communities. The codices from Santa Eulalia, San Juan Ixcoy, and San Mateo Ixtatán are now at the Lilly Library at Indiana University. The PDFs of these manuscripts' images can be found online at http://webapp1.dlib.indiana.edu/findingaids/view?docId=InU-Li-VAD6654&doc.view=items.

[7] A cathedral serves as the central church of a diocese (i.e., an ecclesiastical district under the jurisdiction of a bishop); it contains the bishop's "cathedra" or throne ("Cathedral" 2021; "Diocese" 2021).

positions (“Libro de Actas Capitulares 1” 1531, 111; Lehnhoff 1986, 34).[8] This event laid the foundation for the introduction and development of sacred and polyphonic music in the cathedral of Santiago de Guatemala (Lehnhoff 1986, 35).

An inventory from 9 May 1542 shows the presence of polyphonic books in the Almolonga cathedral (Lehnhoff 1986, 36–37; Snow 1996, 5).[9] It lists the books used in the celebration of the liturgy as follows:

> “cinco missales sevillanos…, dos manuales y quatro processionales, vn pontificial digo libro pontifical (sic), vn antiphonario, libro antiphonario Romano, vn breviario Romano, otro missal sevillano de pergamino, vn psalterio romano, vn passionario…, *vn libro grande de Canto d’Organo, otros quatro libros pequeños de música*” (“Libro de Actas Capitulares 1” 1531, 18).

> Five missals from Seville …, two manuals and four processionals, a pontifical titled libro pontifical (sic), an antiphoner, a Roman antiphoner, a Roman breviary, another missal according to the use of Seville of parchment, a Roman psalter, a passion book … *a big book of polyphonic music, as well as four small music books.*[10]

The last two entries (in italics) refer to books of polyphony: “a big book of polyphonic music and four small music books.” Snow (1996, 5–6) suggests that the big book of music is a choirbook containing fifteen polyphonic masses, and the four small music books could be a collection of motets *a 4* in partbooks. The presence of polyphonic books does not necessarily indicate that polyphony was already sung during the mass since the cathedral office of *chantre* seemed to be vacant at that time (Snow 1996, 6). Lehnhoff (1986, 36) points out that there are no certificates of

[8] The definition of the role of the *chantre* can be found in the original text by Marroquín (in Latin), in the manuscript “Libro de Actas Capitulares 1” (1531, 111). The translation in English by Snow (1996, 3) reads: “the cantor had to be learned and skilled in *musica*, that is, polyphony, or at least in *cantus planus*, or plainchant, and was required to preside over the cathedral clergy personally and without the possibility of a substitute whenever they chanted the liturgy, and to sing from the music lectern or wherever else was appropriate. He was to teach, correct, and coordinate everyone in everything pertaining to the celebration of the liturgy in chant.”
[9] The inventory contains the properties of the Almolonga cathedral that were moved to the next location of the capital.
[10] The complete list of music books can be found in Snow (1996, 5). Fiorentino (2016, 505–506) explains the term “canto d’organo” (canto de órgano) as written polyphonic music, as opposed to the term “contrapunto” that refers to improvised polyphony.

appointment or receipts of payment associated with any member of the choir service like the ones found in the years following the city's change of location.

The capital was refounded in the Panchoy Valley in 1543 due to a devastating landslide that came down the Volcán de Agua on 10 September 1541 that destroyed the old city. The new city (today *Antigua Guatemala* or Old Guatemala) received the same name as the old capital, Santiago de los Caballeros de Guatemala. The new Santiago was the capital of colonial Guatemala through the rest of the sixteenth, seventeenth, and most of the eighteenth centuries (until the 1773 earthquake severely damaged it).

Given the years in which the GuatC choirbooks were copied (indicated in the next section), they must have been copied in this new Santiago. Here, the musical life of the cathedral continued to flourish under the leadership of Francisco Marroquín who led the development of sacred music in Santiago de Guatemala's cathedral until his death in 1563. The *actas del cabildo eclesiástico* (minutes of the cathedral chapter) provide evidence of this, from the erection of the cathedral to the hiring of musicians, organists, choir boys, *chantres* and *sochantre*s, and *maestros de capilla*, as well as the acquisition of plainchant and polyphonic books (Lehnhoff 1986, 55).[11] Shortly after his death, and due to his accomplishments, famous figures in the history of polyphonic music in the Americas start appearing as *maestros de capilla* in the new Santiago de los Caballeros cathedral. These include Hernando Franco, Pedro Bermúdez, and Gaspar Fernández (Catalyne and Brill 2001; Snow 2001; Stevenson 2001).[12] As will be seen later, these three are among the

[11] The *maestro de capilla*'s responsibilities included directing the polyphonic choir, choosing the repertoire, and composing. On the other hand, the *sochantre* (the assistant of the *chantre*) was in charge of the plainchant choir.
[12] The Grove article on "Gaspar Fernandes" contains outdated research. Morales Abril (2013) disproved various assumptions regarding Gaspar Fernández's provenance, job as organist, and period as chapelmaster in Guatemala's Metropolitan Cathedral. Here I translate the main points of Morales Abril's article. While previously it was believed that Fernández was born in Évora (Portugal), Morales Abril points out that Guatemala's chapelmaster and Évora's Gaspar Fernandes are two different people. According to the AHAG *libranzas de pago*, Gaspar Fernández was already working at the Metropolitan Cathedral of Guatemala in 1596, time at which Évora's Gaspar Fernandes was

composers whose music is most frequently found in various Mesoamerican collections, including the GuatC choirbooks which are the focus of this dissertation.

The GuatC choirbook collection was copied in this third capital city of Guatemala. Part of the repertoire of the GuatC collection was copied by Gaspar Fernández at the beginning of the seventeenth century, before his departure to Puebla in 1606. This repertoire was copied from the books held by the cathedral since Almolonga times (Lehnhoff 1986, 52; Snow 1996, 17). The first inventory that includes a specific section about the "libros de canto de órgano" (books of polyphonic music) is an inventory from 1633 (Morales Abril 2015, 104; "Cabildo Catedral, Inventarios" 1633, 17v–18r). Morales Abril (2015, 104–105) points out that three of the choirbooks in the GuatC collection are already present in this inventory. GuatC 4 is mentioned as a "libro de las Salves" (book of Salves). GuatC 2 is mentioned as a "libro de himnos y magnificats" (book of hymns and magnificats). A "libro de misas, que llaman el libro blanco" (book of masses that they call the white book) is also mentioned here. Morales Abril suspects that this book is the one mentioned on the title page of GuatC 1 ("Libro de missas copiado de el que escrivió el P. Gaspar Fernandes el año de 1602") as a book of masses that was recopied into GuatC 1, as will be seen in Sections 2.1.2 and 2.1.3. These three manuscripts were all copied by Gaspar Fernández.[13] Morales

still active in Évora according to the *acuerdos capitulares* from the Cathedral of Évora. Moreover, the fact that Gaspar Fernández was *primer colegial* in the *Colegio Seminario de la Asunción* and this college was requested to give preference to the descendants of discoverers and conquerors of Guatemala (according to the AHAG, *Colegio Seminario, Reales cédulas*, June 22nd, 1592) sustains Morales Abril's claim that there is a high probability that Gaspar Fernández was born in Guatemala. Morales Abril also amends the year when Fernández was hired as chapelmaster to 1603 (previously believed to be 1602, based on the data of the title page of GuatC 1) and he points out that Fernández was never hired as organist (this previous assumption was based on information contained on the *actas capitulares, libro 2, folio 16v*, but this information referred to "Gaspar Martínez" and not "Gaspar Fernández").

[13] The other choirbooks mentioned in this inventory are three. The first, a "libro de salmos" (psalm book) is no longer in existence, but much of its content is probably preserved in PueblaC 6 (choirbook 6 from the Cathedral of Puebla, Mexico). And there are two books of polyphonic masses by Guerrero and Morales, a "libro de Guerrero que tiene toda música" and a "libro de Morales, de misas," that are no longer in existence (Morales Abril 2015, 105).

Abril also believes that choirbook 3 already appears in an inventory from 1704 (Morales Abril 2015, 105–106; "Cabildo Catedral, Inventarios" 1633, 76v–77r).

The capital was moved to its current location in 1775, with the name *Nueva Guatemala de la Asunción* (New Guatemala of the Assumption), now known as Guatemala City. The choirbooks were moved to this new location. They are currently held at the music vault of the *Archivo Histórico Arquidiocesano de Guatemala* (Archdiocesan Historical Archive of Guatemala, AHAG), an archive located at the Archiepiscopal Palace beside the Metropolitan Cathedral in Guatemala City.[14]

2.1.2 The GuatC Choirbook Collection

The collection of choirbooks from Guatemala's cathedral (GuatC) consists of five large bound books and a sixth small unbound one. The repertoire of this collection, consisting mostly of Renaissance polyphonic music, was copied during the seventeenth and eighteenth centuries and remained in use until at least the beginning of the nineteenth century (Morales Abril 2015, 103).

The music of this collection is mostly written in mensural notation—a notation system used during the Renaissance for polyphonic music—and in choirbook format (Figure 2-1). In choirbook format, each voice is presented as a self-contained block occupying its own space on the page or opening. One opening can include an entire section of a work, or part of a section. The page turns are coordinated in all voices. For four-part texture, as is the case for most pieces in GuatC 1, each voice occupies one of the four quadrants of the opening (Figure 2-1). The actual distribution of the voices in these quadrants depends on the source (Schmidt, Leitmeir, and Gumbert 2018, 32–4).

[14] The Archivo Histórico Arquidiocesano de Guatemala "Francisco de Paula García Peláez" is located within the Palacio Arzobispal (7ª avenida 6-21 zona 1). In addition to the choirbook collection, the music vault of the archive stores chant books and sheets of sacred music.

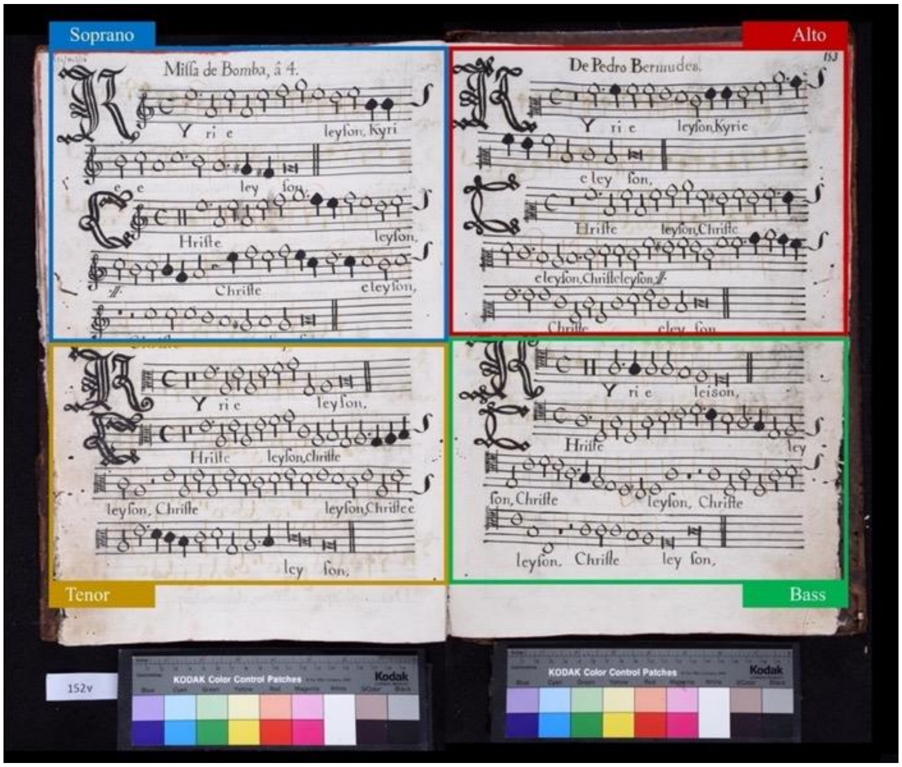


Figure 2-1: Example of choirbook format for a four-voice section in GuatC 1.

Three of the six GuatC choirbooks have been known by scholars for a few decades. GuatC 1–3 were inventoried by Pujol (1965), Stevenson (1970), and Snow (1996). GuatC 4 was rediscovered after the 1976 earthquake (Snow 1996, ix), and it was studied and transcribed in its entirety by Snow. The last two choirbooks were discovered in 2010 in the AHAG. The fifth book was found by Omar Morales Abril and Javier Marín López in July 2010 (Morales Abril 2015, 97). In October of the same year, Morales Abril found, among unclassified music sheets, what he considered to be a small unbound choirbook (Morales Abril 2015, 97). This sixth book consists of a few pages attached together with three pieces of thread. Despite its small dimensions (21.5 cm x 30 cm) and number of folios (19 folios), the parts are written in choirbook format.

In addition to the inventories mentioned above and the transcription of choirbook 4 by Snow,[15] (Guatemalan) musicologists Dieter Lehnhoff and Omar Morales Abril have transcribed and performed music from the GuatC collection. Transcriptions include Lehnhoff (1984; 2001; 2008), Morales [Abril] et al. (2001), and Morales Abril (2003); and recordings include the one by Lehnhoff with the Ensemble Millennium and Schola Cantorum (1995). Moreover, imaging work has been done with part of the collection. The first three choirbooks, along with other music in the archive, were microfilmed in 1984.[16] These microfilms are held at the *Centro de Investigaciones Regionales de Mesoamérica* (CIRMA), an archive in *Antigua Guatemala* (Old Guatemala). The images from the choirbooks GuatC 1–3 are on reel 9. Unfortunately, the images of the choirbooks are missing a few folios, and other folios seem cropped—which is not the case when looking at the original manuscript. CIRMA is currently generating PDF files from these microfilms.[17] Digital images (DjVu files) from these microfilms were obtained by the Miami University Libraries in 2004 and can be consulted online.[18] According to the Música Colonial Archive website,[19] the microfilms were borrowed from CIRMA and brought temporarily to the Miami University Libraries in 2002, where they were digitized into TIFF (in 2003) and then converted into DjVu files (in 2004) to be shared online (Goodliffe 2020).[20] Selected items from the microfilms

[15] Snow has also transcribed other pieces contained in the Guatemalan choirbooks, either in full or incipits. Snow's "Obras completas de Rodrigo Ceballos" (1995) and "Gaspar Fernandes, obras sacras" (1990) anthologies include transcriptions of some of the GuatC pieces by Ceballos and Fernández. And in his book "The Extant Music of Rodrigo de Ceballos and Its Sources" (1980), he provides incipits of some Ceballo's pieces found in GuatC1 and 2.
[16] The microfilms are held at the archive of the *Centro de Investigaciones Regionales de Mesoamérica* (CIRMA) in Antigua Guatemala, Guatemala. CIRMA made them in 1984 with the authorization of the Chancellor of the Ecclesiastical Curia, Monsignor Efraín Hernández (Thelma Porres, director of CIRMA's historical archive, email communication, March 2020). Guatemalan musicologist Dieter Lehnhoff conducted the microfilming process, starting in June 1984, with staff and equipment provided by William Swezey, current director of CIRMA (Lehnhoff 1994, 3–4).
[17] Thelma Porres, director of CIRMA's historical archive, personal communication, November 2018.
[18] The DjVu files can be consulted at http://conan.lib.miamioh.edu/musica/data/. The names of the files belonging to the 9th reel (the ones corresponding to the GuatC collection) start with the number "9."
[19] https://www.cpdl.org/wiki/index.php/M%C3%BAsica_Colonial_Archive
[20] Unfortunately, the staff at CIRMA seems to be unaware of this transaction.

(including some of the items in the choirbooks) have been transcribed into modern scores and are available in the Choral Public Domain Library (CPDL) “Música colonial scores” site, but they represent only a tiny portion of the music contained in the microfilms (Goodliffe 2020).[21] Most of the works in CPDL’s Musica colonial scores have been transcribed from the microfilms, while others have been transcribed from concordant sources (Goodliffe 2020).

GuatC 1 is a book of masses, GuatC 2 is a book of hymns and Magnificats, GuatC 3 and 5 are books with miscellaneous contents, GuatC 4 is a book for Holy Week and Salve services, and GuatC 6 presents an Office for the Dead. Table 2-1 provides general information about each of the six books in the collection.

[21] The transcribed scores can be found at http://www.cpdl.org/wiki/index.php/Category:Música_colonial_scores.

Table 2-1: General information for each of the choirbooks GuatC 1–6 (Morales Abril 2015, 96–103).

<table>
<tr><th colspan="2">Choirbook Number</th><th>Type</th><th>General Information and Copyists</th></tr>
<tr><td colspan="2">1</td><td>Book of masses (and short liturgical works)</td><td>Copied from Gaspar Fernández's *libro de misas* (1602) by Manuel José de Quiroz c. 1760–65</td></tr>
<tr><td rowspan="2">2</td><td>A</td><td>Hymns</td><td rowspan="2">Copied by Gaspar Fernández and obtained by the cathedral in 1606</td></tr>
<tr><td>B</td><td>Magnificats (and Benedicamus settings)</td></tr>
<tr><td colspan="2">3</td><td>Many motets and varied content (hymns, sequences, litanies, and masses)</td><td>Intervention in the eighteenth century: new folios were pasted over old ones either to hide a piece or to substitute it with another one. The binding was not restored and, eventually, it became infested by silverfish. This caused both the old and new folios to be detached, resulting in the disorganized/incomplete aspect of the book.

Nicolás Márquez Tamariz (1627–93), maestro de capilla from November 1669 to October 1693, copied most of the music of the original folios.

Juan Fernández de Leon (1671–1731), originally from Oaxaca (México), was active in Guatemala beginning in 1696 and served as maestro de capilla between 1717–31. He added two pieces to the last folio.

Manuel José de Quirós (fl. 1694–1765), maestro de capilla and Guatemalan composer, copied the music of the new folios during the mid-eighteenth century.</td></tr>
<tr><td colspan="2">4</td><td>Holy Week and Salve services</td><td>Gaspar Fernández (in 1605)</td></tr>
<tr><td colspan="2">5</td><td>Diverse content</td><td>Copied by different scribes during the mid-seventeenth to mid-eighteenth centuries:

Juan José Guerrero, who was active as a composer in the Metropolitan Cathedral of Guatemala since 1644 and served as maestro de capilla from July 1658 to October 1669. He copied and composed a set of polychoral psalms for Vespers.

Nicolás Márquez Tamariz, who was Guerrero's successor. He copied two motets, an invitatory, and a psalm for Vespers.

The Oaxacan musician Juan Fernández de León copied two responsories.

An unknown scribe copied another invitatory.

Manuel José de Quirós copied the last ten pieces, which consist of a Benedicamus Domino, a Christmas invitatory, and various sections of the mass ordinary.</td></tr>
<tr><td colspan="2">6</td><td>Office of the Dead</td><td>Probably copied during the first half of the seventeenth century, perhaps by Diego de Galvez Prado (fl. 1597–1648), maestro de capilla 1636–48.</td></tr>
</table>

The choirbooks contain pieces by composers active in different regions (see Figure 2-2). Of the twenty-seven composers present in the GuatC corpus, ten of them worked in the territory of the Viceroyalty of New Spain (nine in Guatemala and one in Mexico), some were European composers that visited this territory and others were born in it. There are also ten composers from the Iberian Peninsula who never set foot in the Americas, together with four non-Hispanic composers, and three composers whose origin remains to be uncovered. Non-Hispanic composers include names such as Palestrina and Lassus, while Peninsular composers include the triumvirate of Spanish polyphonists Tomás Luis de Victoria, Cristóbal de Morales, and Francisco Guerrero, among others. The most prolific composers in the GuatC collection who lived in Guatemala include Hernando Franco, Pedro Bermúdez, and Gaspar Fernández, who were maestros de capilla at the Metropolitan Cathedral of Guatemala in 1569–1574, 1601–1603, and 1603–1606, respectively (Gembero Ustárroz 2005; Morales Abril 2007; 2013). These three are among the most frequently found composers in more than one archive from the Mesoamerican region (Lemmon 1993). Gaspar Fernández was also the scribe of choirbooks 2, 4, and of a *libro de misas* on which choirbook 1 was based. Out of the ten composers active in the Viceroyalty of New Spain, five of them were born or trained in the Americas: Gaspar Fernández,[22] Juan José Guerrero,[23] Manuel José de Quirós,[24] Fray Francisco de Quirós,[25] and Juan Matías de Rivera;[26] the first four in Guatemala and the last one in Mexico (Morales Abril 2015, 115).

[22] While it was previously believed that Fernández was born in Évora (Portugal), Morales Abril (2013) provides evidence that Guatemala's maestro de capilla and Évora's Gaspar Fernandes are two different people.
[23] Juan José Guerrero is identified as "Criollo"—a person of Spanish descent born in the American colonies—in his entry in the Books of Hispanic Polyphony database. Javier Marín-López, "Guerrero, Juan José", Books of Hispanic Polyphony, ed. E. Ros-Fábregas (access date: 04 Dec 2019), https://hispanicpolyphony.eu/person/24640. Submitted: 29 Jun 2017; last modified: 11 Oct 2017.
[24] Manuel José de Quirós was born in Guatemala as indicated by Stevenson (1980, 42) and Lemmon (1984).
[25] Fray Francisco de Quirós was the brother of Manuel José de Quirós, as indicated in the header of GuatC 3 *Sancta Maria succurre*, and he was also a native-born Guatemalan (Stevenson 1980, 42).
[26] Juan Matías de Rivera is identified as "Indian" (i.e., native) born in Oaxaca (Mexico) in his entry in the Books of Hispanic Polyphony database. Javier Marín-López, "Mathías Ribera, Juan", Books of Hispanic Polyphony, ed. E.

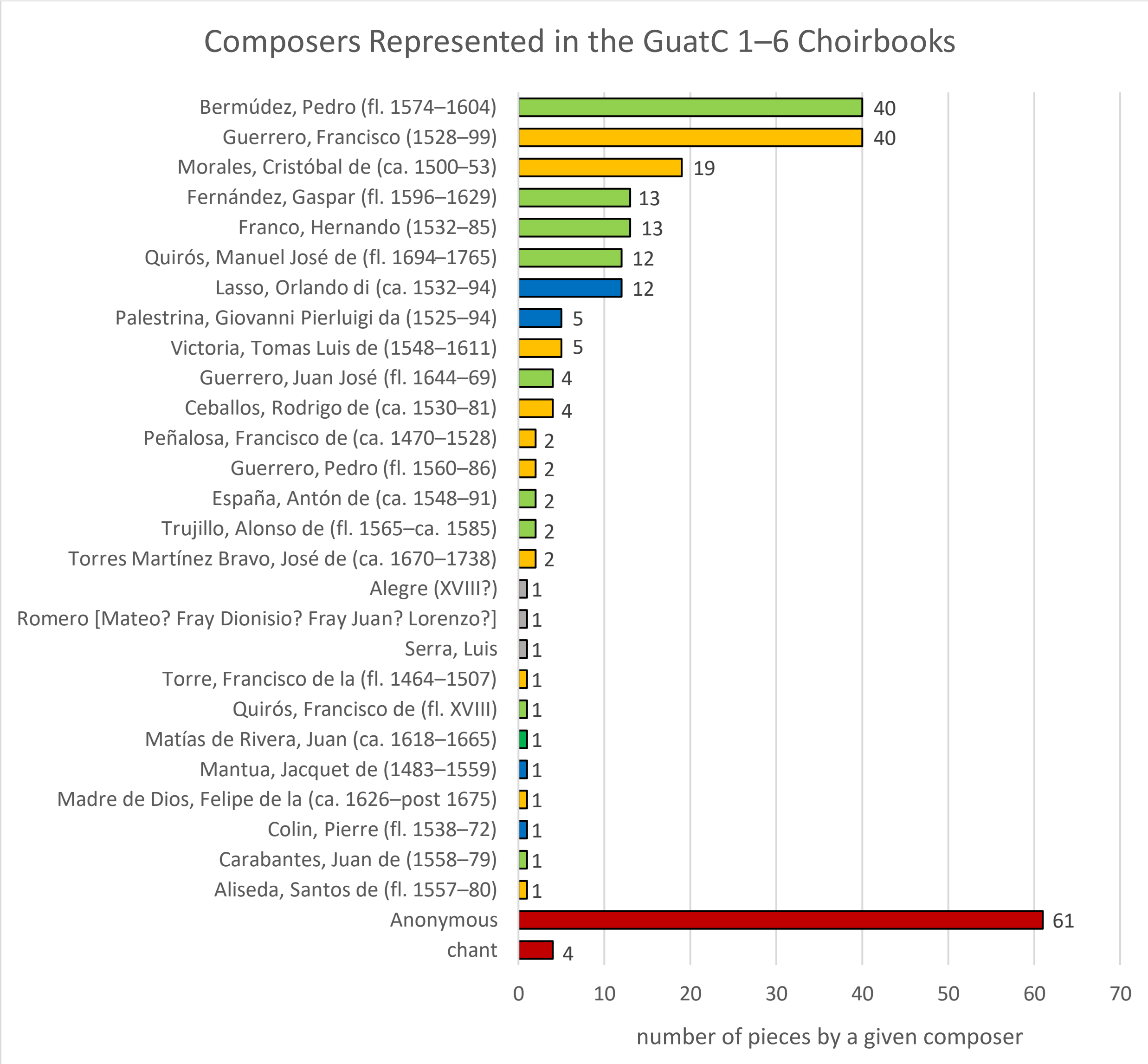


Figure 2-2: Composers present in the GuatC choirbooks and number of pieces written by each of them. This is a graphic representation of the information provided by Morales Abril (2015, 118–121).

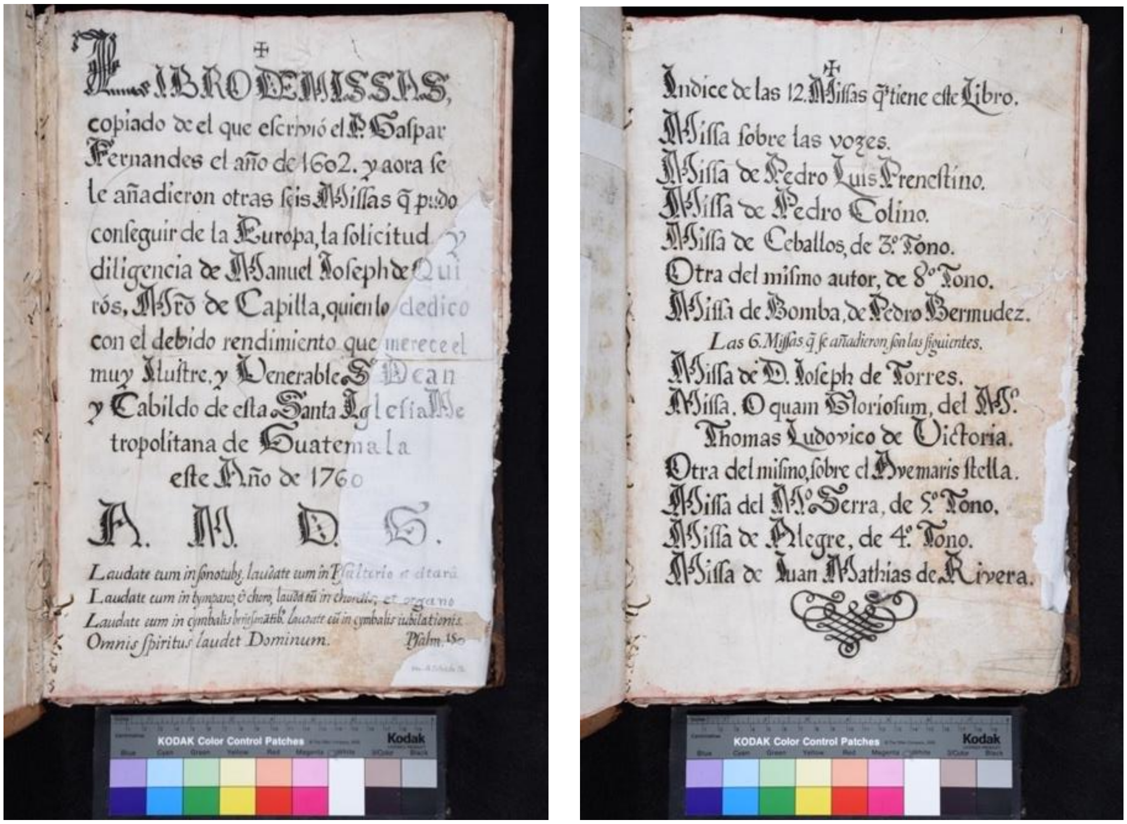
Key: light green for composers active in Guatemala, dark green for composers active in other parts of the Viceroyalty of New Spain (here only Juan Matías de Rivera, in Oaxaca), yellow for Iberian Peninsular composers, gray for Iberoamerican composers whose actual region of activity is undetermined, blue for composers from other parts of Europe, and red for Anonymous and chant.

Ros-Fábregas (access date: 10 Dec 2019), https://hispanicpolyphony.eu/person/24266. Submitted: 23 Jun 2017; last modified: 11 Oct 2017.

2.1.3 **The GuatC 1 Choirbook**

While a general overview of the choirbook collection is presented by Morales Abril (2015), this section focuses on GuatC 1, the choirbook digitized and encoded in this dissertation. The inventory of GuatC 1 contents is provided in Table 2-2.[27] The choirbook contains twelve masses and fifteen brief liturgical pieces. Six of the masses were copied from a *libro de misas* copied by Gaspar Fernández in 1602 (as mentioned on the index page), and the other six were added by Manuel José de Quirós in the eighteenth century (see Figure 2-3). The fifteen polyphonic pieces, most of them ascribed to sixteenth-century composers (Bermúdez, Guerrero, and Franco), are also an eighteenth-century addition (Morales Abril 2015, 98). For a list of the GuatC 1 pieces transcribed in CPDL's Música colonial scores, please consult the Appendix.

[27] The information in this table comes from the index page of GuatC 1 (which separates the 12 masses into two groups, the ones copied from the *libro de misas* and the masses added when this manuscript was copied) and the inventories by Pujol (1965, 4), Stevenson (1970, 65–71), Snow (1996, 25), and Morales Abril (unpublished). Most of the information regarding genre was obtained from Pujol, although he divides piece 21 into two invitatories (*Surrexit Dominus vere* and *Alleluia, alleluia, alleluia*) and identifies piece 26 as the motet *Agnus redemit*. The latter has been corrected by Snow to *Dic nobis Maria* from *Victimae paschali laudes*, which is a sequence. The composer attributions come from Morales Abril. Information regarding the folios and number of voices has been double-checked against the manuscript.



LIBRO DE MISSAS,
copiado de el que escrivió el P. Gaspar
Fernandes el año de 1602. y aora se
le añadieron otras seis Missas q pudo
conseguir de la Europa, la solicitud y
diligencia de Manuel Joseph de Qui
rós, Mro de Capilla, quien lo dedico
con el debido rendimiento que merece el
muy Ilustre, y Venerable S. Dean
y Cabildo de esta Santa Iglesia Me
tropolitana de Guatemala
este Año de 1760

A. M. D. G.

Laudate eum in sono tubæ, laudate eum in Psalterio et cithara.
Laudate eum in tympano, et choro, laudate eum in chordis, et organo
Laudate eum in cymbalis benesonantibus, laudate eum in cymbalis iubilationis.
Omnis spiritus laudet Dominum. Psalm. 150

Indice de las 12. Missas q tiene este Libro.
Missa sobre las vozes.
Missa de Pedro Luis Prenestino.
Missa de Pedro Colino.
Missa de Ceballos, de 3º Tono.
Otra del mismo autor, de 8º Tono.
Missa de Bomba, de Pedro Bermudez.
Las 6. Missas q se añadieron son las siguientes.
Missa de D. Joseph de Torres.
Missa. O quam Gloriosum, del Mº Thomas Ludovico de Victoria.
Otra del mismo, sobre el Avemaris stella.
Missa del Mº Serra, de 5º Tono.
Missa de Alegre, de 4º Tono.
Missa de Juan Mathias de Rivera.

Figure 2-3: The title page (left) indicates that this book of masses was copied from a previous book copied by Gaspar Fernandes in 1602, and to which six new masses were added by Manuel Joseph de Quirós in 176[0]. The index page (right) lists the twelve masses in the two groups, the ones from the 1602 book and the six masses added later.

Table 2-2: Inventory of the GuatC 1 choirbook (Morales Abril unpublished).

No.	Folios	Piece	Composer	Genre	Voices	Information
1	1v-4r	Asperges me	[¿Pedro Bermúdez?]	antiphon	4	Added in the eighteenth century
2	4v-7r	Asperges me	Anonymous	antiphon	5	
3	7v-11r	Vidi aquam	Pedro Bermúdez	antiphon	4	
4	11v-14r	Vidi aquam	[Hernando Franco]	antiphon	4	
5	14v-17r	Asperges me	[Alonso de Trujillo]	antiphon	4	
6	17v-31r	Missa sobre las vozes	Cristóbal de Morales	mass	4	From the 1602 book. This is Morales's hexachord mass.
7	31v-43r	Missa sine nomine	Giovanni Pierluigi da Palestrina	mass	4	From the 1602 book
8	43v-54r	Missa Pere de nous	Pierre Colin	mass	4, Agnus II a 7	From the 1602 book
9	54v-71r	Missa de 3º tono	Rodrigo de Ceballos	mass	4, Agnus II a 5	From the 1602 book
10	71v-85r	Missa sine nomine	José de Torres y Martínez Bravo	mass	4, Agnus II a 8	Added in the eighteenth century
11	85v-95r	Missa O quam gloriosum	Tomás Luis de Victoria	mass	4	Added in the eighteenth century
12	95v-109r	Missa Ave maris stella	Tomás Luis de Victoria	mass	4	Added in the eighteenth century
13	109v-119r	Missa de 5º tono	Luis Serra	mass	4	Added in the eighteenth century
14	119v-133r	Missa de 4º tono	Alegre	mass	4	Added in the eighteenth century
15	133v-152r	Missa de 8º tono	Rodrigo de Ceballos	mass	4, Agnus 2 pars a 5	From the 1602 book. Ceballos' Missa Simile est regnum cælorum, parody mass based on a motet by Morales.
16	152v-169r	Missa de Bomba	Pedro Bermúdez	mass	4	From the 1602 book. Parody mass based on Mateo Flecha's (el Viejo) ensalada *la Bomba*.
17	169v-180r	Missa sine nomine	Juan Matías de Rivera	mass	4	Added in the eighteenth century
18	180v-181r	Christus natus est	Pedro Bermúdez	invitatory	4	Added in the eighteenth century
19	181v-183r	Christus natus est	Pedro Bermúdez	invitatory	8	
20	183v-184r	Christus natus est	Anonymous	invitatory	5	
21	184v-186r	Surrexit Dominus vere	Anonymous	invitatory	5	
22	186v-187r	Lumen ad revelationem	Hernando Franco	antiphon	5	
23	187v-188r	Lumen ad revelationem	Pedro Bermúdez	antiphon	5	
24	188v-189r	Lumen ad revelationem	Pedro Bermúdez	antiphon	4	
25	189v-190r	Surrexit Dominus vere	Anonymous	invitatory	4	
26	190v-192r	Victimæ Paschali laudes	Francisco Guerrero	sequence	4, 2a Pars a 5	
27	192v-193r	Tantum ergo	[Pedro Bermúdez]	hymn	4	

Regarding the masses from the original *libro de misas*, Snow provides evidence that all of them were composed in Europe, except for Bermúdez's *Missa de Bomba*. The Morales and Ceballos masses were probably brought to Guatemala by Bermúdez, since he had served at Granada Cathedral (Snow 1996, 19). Snow lists European concordant sources for the Morales, Palestrina, Colin, and one of the Ceballos masses (his popular *Missa tertii toni*).[28] The *Missa de 8º Tono* by Ceballos is his *Missa Simile est regnum caelorum*, an imitation mass based on a motet by Morales. *The Missa de Bomba* by Pedro Bermúdez, Guatemalan *maestro de capilla* between 1601 and 1603, is an imitation mass based on the ensalada *La Bomba* by Mateo Flecha el Viejo (the Elder). Of the group of six masses added in the eighteenth century the ones by Torres, Serra, Alegre,[29] and Matías de Rivera appear to be unique to this manuscript, according to Snow (1996, 19).

The GuatC choirbook collection documents a continuous performance tradition of sacred choral music from the Renaissance until at least the beginning of the nineteenth century, allowing to study the transmission of music from Europe to Latin America. Given the limited access to the manuscript sources, it is of utmost importance to digitize and encode them to preserve and increase access to this corpus. I expect this pilot project with GuatC 1 to be a significant step in this direction.

---

[28] The text in Snow (1996, 19) reads as follows on the concordances of these masses: "The *Missa sobre las vozes* proves to be the hexachord mass by Morales, preserved in only three other sources: a manuscript at the Capilla Real in Granada, the manuscript formerly known as Medinaceli 607 and now owned by Bartolomé March Servera, and Tarazona 5. … The mass by Palestrina is the *Missa sine nomine* first published in the composer´s *Missarum liber secundus* in 1567 (RISM P 660), which was dedicated to Philip II, and the Colin setting is his *Missa Pere de nous*, published at Lyon in 1546 by Jacques Moderne in his *Liturgicon musicarum duodecim missarum* (RISM C 3310). The two masses by Ceballos are his popular *Missa tertii toni*, which is preserved in at least ten other sources, and his *Missa Simile est regnum caelorum*, referred to in Guatemala 1 as *Missa octavi toni* but actually a parody mass based on a motet by Morales."

[29] According to Morales Abril (2015, 98), the "Alegre" mentioned in GuatC 1 is not be the same person as Gregorio Allegri.

2.2 Digitization of Cultural Heritage Documents

The goal of this dissertation is to preserve and increment access to the Guatemalan choirbooks through digital images and symbolic scores, which are generated from the images through optical music recognition and other music information retrieval technologies. With this goal in mind, this section provides an overview on the best practices and recommendations for obtaining raster images within the context of digitization projects of cultural heritage documents, compiled from various documents available online or personal communication with personnel from institutions with experience on the subject.[30] A *raster* or *bitmapped image* is a type of digital image that is composed of pixels—in opposition to other digital images, called vector images, that are described through mathematical equations. A raster image is a two-dimensional matrix of pixels. It is obtained by either sampling an analog item at regular intervals and recording the color value for each sample when scanning or by capturing the samples and their colors in one shot when photographing (Besser 2003, 3). Within the context of digitization projects, the images obtained at capture are called master images (or master files). Master images are obtained at the best possible quality, and they are used for preservation and archival purposes (Besser 2003, 44). From these master images, other lower-quality and (therefore) more easily distributed images are derived; these are called access images (or access files) (Besser 2003, 3).

Section 2.2.1 presents the basic concepts of bitmapped images, including details about resolution, bit depth, compression, file formats, file size, and color representation. Section 2.2.2 presents the best practices and recommendations of different institutions conducting digitization

[30] I found most of the online resources going through the bibliography from https://digitizationguidelines.gov which, although being a US site, pointed to both US and non-US resources as can be found at https://www.digitizationguidelines.gov/guidelines/Guidelines_Bibliography-2009rev.pdf. Some of these resources needed to be accessed through Wayback Machine (https://web.archive.org) as can be seen in the references. I also consulted international resources like the International Federation of Library Associations and Institutions (IFLA).

projects (especially focusing on cultural heritage institutions) and some of the standards regarding resolution, bit depth, file formats, and other concepts presented in the first section for master and access files.

2.2.1 Basic Concepts of Raster Images

As described above, a raster or bitmapped image is a file encoding the color values of a grid of pixels, i.e., the samples taken by a scanning device. The quality of a raster image depends on the sampling rate of the scanning device and the number of colors it can capture per sample. These two features are related to the concepts of resolution and bit depth.

Resolution is the sampling frequency, expressing how many samples are taken per inch. It is usually expressed in *pixels per inch* (*ppi*) for capture devices and monitors, and in *dots per inch* (*dpi*) for print devices (McIlwaine et al. 2002, 22).[31] To illustrate this, a $1\ in^2$ image obtained at $300\ ppi$ would contain a total of $90{,}000\ pixels$. As another example, an A4 document $(8.27\ in \times 11.69\ in)$ scanned at $300\ ppi$ would contain a total of $8.7\ Megapixels = 8.7 \times 10^6\ pixels = 2{,}490 \times 3{,}510\ pixels = (300\ pixels/in \times 8.27\ in) \times (300\ pixels/in \times 11.69\ in)$.

The tone value of the sample is encoded in bits. The **bit depth** or **dynamic range** provides the number of bits available to encode the value of the sample. For the different types of images, we have:

- On a black and white (B&W) image, the bit depth is equal to 1. Only one bit is needed to provide a binary value of black or white (0 or 1).

[31] A pixel is a much smaller physical unit than a dot (McIlwaine et al. 2002, 22).

- On a grayscale image, each sample needs to encode different gray tones. The preferred level of grayscale is that of 8 bits (or 1 byte) per pixel, which allows for encoding 256 different shades of gray (McIlwaine et al. 2002, 18).
- On a color image, red, green, and blue channels are used to encode different values of these three colors, which in combination provide other colors (see Section 2.2.1.1 on color representation). Also using 8 bits per channel here, which results in 24 bits (or 3 bytes) for the three channels, allows for encoding a total combination of 16.7 million colors. This 24-bit value is known as *true color* level (McIlwaine et al. 2002, 18).

The number of samples in the document and the bits used to encode each of those samples determine the file size as indicated by the following equation:

$$\begin{aligned} FileSize &= AreaPixels \times BitDepth \\ &= [(Resolution \times Length)\,(Resolution \times Width)] \times BitDepth \\ &= Resolution^2 \times (Length \times Width) \times BitDepth \\ &= Resolution^2 \times Area \times BitDepth \end{aligned}$$

The previous formula returns the file size in bits; its division by 8 returns the file size in bytes. Determining the file size of a scanned image (excluding the header information) is as simple as having the dimensions of the original document (in square inches) and the resolution and bit depth of the capturing device. See the following examples on how to calculate the file size for black-and-white, grayscale, and colored images for an A4 document captured at 300 ppi:

- Black and white (1-bit) image:
  $FileSize = (300\,ppi)^2(8.27\,in \times 11.69\,in)(1\ bit/pixel) = 8.7 \times 10^6\ bits$
  Or, equivalently in bytes, $1.1 \times 10^6\ bytes = \mathbf{1.1}\ \boldsymbol{MB}$

- Grayscale (8-bit) image:
  $FileSize = (300\,ppi)^2(8.27\,in \times 11.69\,in)(8\ bit/pixel) = 69.6 \times 10^6\ bits$
  Or, equivalently in bytes, $8.7 \times 10^6\ bytes = \mathbf{8.7}\ \boldsymbol{MB}$

- Colored (24-bit) image:

$FileSize = (300\ ppi)^2(8.27\ in \times 11.69\ in)(24\ bit/pixel) = 208.8 \times 10^6\ bits$
Or, equivalently in bytes, $26.1 \times 10^6\ bytes = \mathbf{26.1\ MB}$

The Table 2-3 below provides the results of the same calculations for different document sizes.

Table 2-3: Comparison of file size for documents of various dimensions and bit depths

Document Dimensions (in^2)	**File Size (MB)**		
	B&W (1-bit) Image	**Grayscale (8-bit) Image**	**Color (24-bit) Image**
A2 (16.54×23.36)	4.35	34.8	104.3
A3 (11.69×16.54)	2.2	17.4	52.2
A4 (8.27×11.69)	1.1	8.7	26.1

The file size of an image has implications in terms of storage and transmission. **Compression** can be applied to images to reduce the associated costs. “Compression is a method for reducing the size of a file by squeezing the same information into fewer bits” (Wong 2016, 41). The amount of compression is expressed by a ratio comparing the size of the compressed image against the uncompressed one (e.g., a ratio of 4:1 means that the image is compressed to one-fourth its original size) (Besser 2003, 19).

There are two types of data compression: lossless and lossy. The former reduces the bits to encode the image without losing any information. This is done, for example, by identifying and eliminating redundancies (e.g., pixels with identical color values) (Besser 2003, 19).[32] On the other hand, lossy compression algorithms do lose information present in the original image; however, these algorithms are usually designed so that the information lost is imperceptible to the human eye, provided one does not use an elevated compression ratio (Wong 2016, 41; Besser 2003, 19).

[32] An example of this is the run-length encoding (RLE), a lossless compression algorithm where a “sequence of the same repeated value is replaced by one instance of the value followed by the number of times it is repeated” (Wong 2016, 41). To illustrate this, I will take the example given in Wong (2016, 42). Suppose that the blue color given by the byte 00001010 is contained in a section of the sky on an image and that this section contains 100 pixels with that exact color. With no compression, this section would take 100 bytes. If using RLE, the same section would be encoded using only 2 bytes: the first indicating the instance of blue color (00001010) and the second encoding the number 100 in binary value (01100100).

For archival images, as will be shown in Section 2.2.2, it is recommended to use uncompressed files. However, if one needs to use compression for master images—due to insufficient storage space, for example—one should use lossless compression (Besser 2003, 20). Nonetheless, it is important to keep in mind that any type of compression (lossless or lossy) will introduce risks for long-term generational integrity—this is, the ability to produce perfect duplicates of the material over and over again, over many generations with no information loss after multiple cycles of compression and decompression (Besser 2003, 19). Therefore, storing master images in uncompressed file formats is always preferable when no level of degradation is acceptable.

Regarding file formats, many digital imaging projects have settled on the formula of using TIFF for master files, JPEG for access files, and GIF for thumbnail files (Besser 2003, 21). The following table presents these three formats and others that have been proposed for storing master and access images: JPEG2000 and PNG.

Table 2-4: File formats details as found in Besser (2003, 21–2), Wong (2016, 38), and various sources of documentation for each file format (CompuServe Incorporated 1990; Murray and Van Ryper 2008; W3C 2003; JPEG 2022; Federal Agencies Digitization Guidelines Initiative 2022; Library of Congress 2021b; 2021a, 1; 2021c; Adobe Developers Association 1992).[33]

File Format	Extension	General Information	Supported Features	Compression
GIF (Graphics Interchange Format)	.gif	Initial release in 1987 and latest release in 1989. Proprietary format (CompuServe).	• Pixel's depth from 1 to 8 bits. • Supports grayscale images. • Supports up to 8-bit indexed color (256-color palette). Its limited color palette is best for illustrations with large areas of solid color and clear division between these areas. • Allows transparency of one designated color.	Lossless compression method LZW. (LZW's patent expired in 2004.)
PNG (Portable Network Graphics)	.png	Designed as a free, open-source alternative to GIF. It was first released in 1996.	• Allows variable transparency. • Up to 8-bit for indexed color. • Supports 8- and 16-bit grayscale images (with 16- and 32-bit when including an alpha channel). • 24- and 48-bit for color images (with 32- and 64-bit when including an alpha channel).	Lossless compression.
JPEG (Joint Photographic Experts Group)	.jpg or .jpeg	Introduced in 1992 by the JPEG Committee, with its development starting in 1986. The format that most people know as JPEG is, in fact, JFIF (JPEG File Interchange Format), a public-domain storage format for JPEG compressed images.	• Its compression method was designed for photographic data, which means it works best for continuous-tone images. Sharp transitions create artifacts. • It does not support transparency. • It does not support embedded metadata.	Lossy compression method.
JPEG 2000	.jp2 or .jpx	Developed from 1997 to 2000. "The JPEG committee has striven to ensure that implementations of Part 1 [the core coding system] can be royalty and license-fee free."	• Image compression standard developed to supersede the original one. • Supports embedded XML metadata.	Lossless and lossy compression. In the lossy realm, discrete wavelet transforms (DWT) result in improved quality and compression ratios compared to that of "old" JPEG (JFIF).

[33] The 'alpha channel' is just another channel in the image (just as the red, green, and blue are channels of an RGB image). This extra alpha channel allows encoding the transparency or opacity of a pixel's color.
The term 'indexed color' is a way to deal with the encoding of the colors on an image when having a limited number of bits to represent them. The color information of the image is stored on a color lookup table (CLUT) or palette. The pixels in the image do not encode the actual color values but the index of the palette corresponding to that color.

TIFF (Tag Image File Format)	.tif or .tiff	Created by Aldus Corporation, with its initial release in 1986. There are many versions, and issues regarding applications being able to read some versions but not others. The most recent version is TIFF 6.0 (published in 1992). Earlier versions are valid TIFF 6.0 files. Copyright held by Adobe.	Latest version (TIFF 6.0): • Tag-based file format. This tagged internal structure allows for arbitrary additions (e.g., custom metadata fields). • Supports for bilevel, grayscale, palette-color (i.e., indexed color), and full-color images. • Supports alpha channel and layers (i.e., multipage).	Allows uncompressed, LZW compression (lossless), ZIP (lossless), and JPEG (lossy). (LZW's patent expired in 2004.)

2.2.1.1 *Color Representation*

Different devices have different color spaces. Examples of color spaces used in monitors include Adobe RGB and sRGB, while SWOP CMYK is used for printers (Wong 2016, 42). The **color space** provides the color gamut available to render an image (either displaying or printing it) on the device. A color space is always associated with a color model, which provides a mathematical representation for each color. **Color models** describe color numerically in terms of tuples, usually expressing the varying amounts of primary colors. Each model uses a different system and set of primaries (dimensions) to describe colors; the most common are RGB, CMYK, HSB, and CIE (Wong 2016, 42).

The RGB model describes colors using three primary colors: red, green, and blue; any other color is described as a combination of these three. This is the model usually used in monitors, where the composition of color is represented by the addition of red, green, and blue lights, which is similar to the way the human eye works. Adding red light and green light gives yellow; green and blue gives cyan; blue and red gives magenta; and adding full-intensity red, green, and blue lights gives white (see Figure 2-4). The RGB model uses triplets (r, g, b), where each element of the triplet has a value that describes the amount of red, green, and blue, respectively. These three dimensions are also known as channels. In a 24-bit image, the values of the coordinates r, g, and b go from 0 to 255—these are the 256 values generated by the eight bits available to each channel ($256 = 2^8$). Therefore, in a 24-bit image, a value of $(255, 0, 0)$ represents red, $(0, 255, 0)$ represents green, and $(0, 0, 255)$ represents blue. Varying amounts of these three primary colors are used to represent other colors; for example, $(255, 166, 38)$ is a light orange color.

The CMYK model describes colors using the three primary colors: cyan, yellow, and magenta. This is the model usually associated with printers. Mixing cyan with magenta gives blue;

magenta with yellow gives red; and yellow with cyan gives green. In theory, mixing the three primaries at full intensity $(100\%, 100\%, 100\%)$ would result in black. However, this is not true in practice. Due to imperfections on the ink, the color given is rather a brownish-black; because of this, the K (black) component is needed (Wong 2016, 47–48).

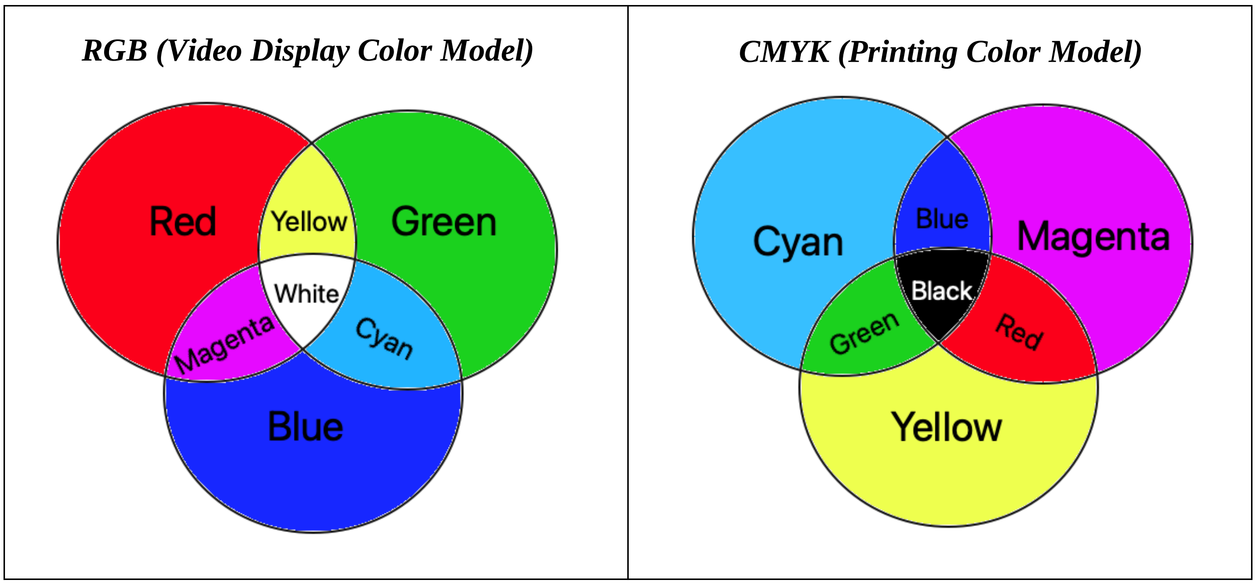


Figure 2-4: RGB and CMYK color models, their primaries, and the colors generated by their mix.

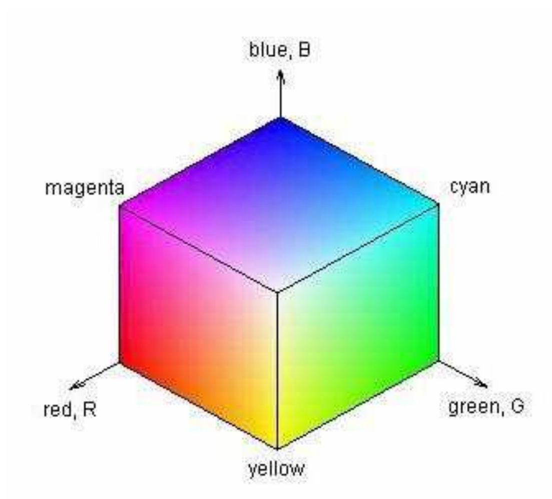


Figure 2-5: RGB color model mapped to a cube, with the x-axis as red (R), the y-axis as green (G), and the z-axis as blue (B). Image from Human-centered content-based image retrieval - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-RGB-color-space-visualized-as-a-cube_fig3_228719004 [accessed 2021-11-02].

The two models presented above express colors by using Cartesian coordinates (i.e., rectangular coordinates) that give the amounts of different primary colors, defining a rectangular space (as the one shown in Figure 2-5 for RGB). The HSB and CIE color models also produce a three-dimensional space of color, but the dimensions used to express these colors are not related to primary colors. The HSB (or HSV) color model uses a cylindrical coordinate system (rather than a rectangular/Cartesian one) to map colors in a cylindrical space given by the (h, s, b) values, where h represents the hue as an angle between $0°$ and $360°$, s represents saturation as a percentage in the radial axis, and b represents the brightness (or v for value) as a percentage in the vertical axis (see Figure 2-6). The CIE XYZ model's primaries (dimensions) do not represent any physical colors, but the model does use a rectangular set of coordinates (x, y, z). This color model was designed so that the color space associated with it encompasses the gamut of all colors visible to the human eye. The CIE color space is shown in Figure 2-7 (see the horseshoe encompassing all colors).

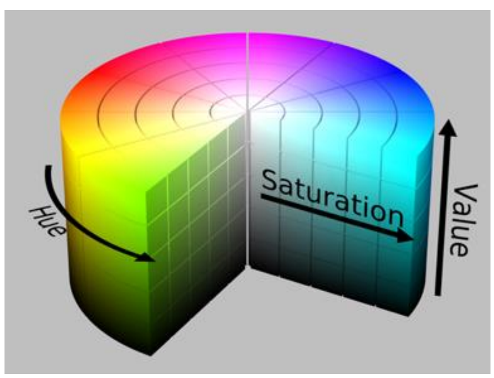


Figure 2-6: The HSV color model mapped to a cylinder. This image was obtained from Wikipedia, created by SharkD under the [CC-BY SA 3.0] license.

The RGB model (used in monitors) is related to the standard RGB (sRGB) and Adobe RGB color spaces mentioned at the beginning; on the other hand, the CMYK model (used in printers) is related to the SWOP CMYK color space. The color space in a device defines the colors available in that device to render the colors in the image as encoded by the color model. As shown in Figure 2-7, the SWOP CMYK color space of a printer does not cover the full red, full green, and full blue colors of an sRGB monitor (see the three corners of the triangle defined by the sRGB color space and compare to the SWOP CMYK polygon). This is why when printing an image, these colors tend to look duller (Wong 2016, 51). There are also colors covered by the SWOP CMYK color space that are not covered by the sRGB. Adobe RGB solves this. It was designed "to meet the demands for an RGB working space suited for print production," having "a color gamut larger than sRGB that encompasses typical press gamuts" (Adobe Systems Incorporated 2005), as shown in Figure 2-7.

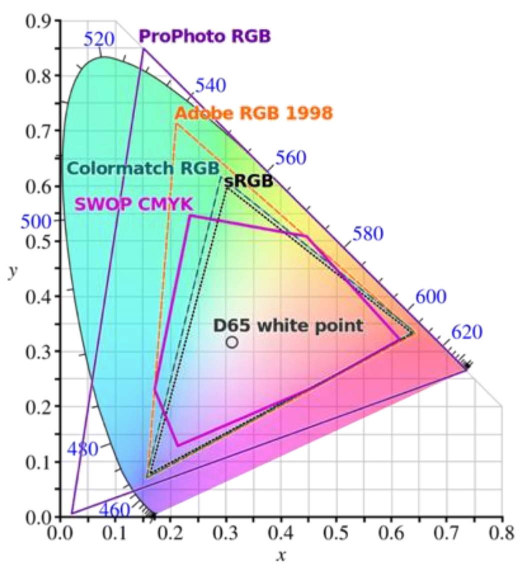


Figure 2-7: Comparison of some RGB and CMYK color gamut on a CIE XY chromaticity diagram. This image was obtained from Wikipedia, created by BenRG and cmglee and shared under CC BY-SA 3.0.

2.2.2 Digital Imaging Standards and Best Practices

There is no single standard for image digitization projects. Different institutions have shared their own standards, guidelines, and best practice recommendations for digitization. I found some of these resources by looking into the bibliography of the Federal Agencies Digital Guideline Imitative, which, although written by a US organization, points to both US and non-US resources.[34] I also looked for international guidelines (e.g., the International Federation of Library Associations and Institutions, IFLA).

Some examples of **guidelines and best-practice** documents for image digitization include: the Arts and Humanities Data Service's (AHDS) *Preservation handbook: bitmap (raster) images* (Eadie 2005) and *Digital images archiving study* (Anderson et al. 2006), the Digital Preservation Coalition's (DCP) *Digital Preservation Handbook (2015)*, the Minerva project's *Good Practice Handbook* (Drake, Justrell, and Tammaro 2003), and the International Federation of Library Associations and Institutions' (IFLA) *Guidelines for digitization* projects (McIlwaine et al. 2002). Examples of **standards** (e.g., resolution, bit depth, and file formats) shared by institutions include the National Library of Australia's (NLA) *image capture standards* webpage (2022), the OCLC's—previously known as the Online Computer Library Center—*Preparing Digital Surrogates for RLG Cultural Materials* webpage (2021), the Bibliographical Center for Research *BCR's CPD Digital Imaging Best Practices (2008)*, and the Federal Agencies Digital Guidelines Initiative (FADGI) Still Image Working Group's *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files (2016)*.[35] All of them indicate the image digitization standards for different types of documents (e.g., text, maps, and photographs), with only NLA and FADGI considering manuscript documents. In addition to the standards, the BCR's CPD and FADGI documents also provide rich documentation regarding guidelines and best practices.

[34] https://www.digitizationguidelines.gov/guidelines/Guidelines_Bibliography-2009rev.pdf

[35] A new draft of this FADGI document does exist (the 2022 revised guidelines) but was not used for this dissertation as it was not available when the digitization process started.

As indicated before, besides the NLA's and FADGI's, no other document makes explicit mention of the digitization of manuscript sources, and no document mentions "music manuscripts" at all. Because of this, I directly consulted with other institutions. I had email communication with project director Julia Craig-McFeely and lead photographer Lynda Sayce from the Digital Image Archive of Medieval Music (DIAMM), as they have experience digitizing music manuscripts. These (medieval) music manuscripts are normally written on parchment, which is not the case with the Guatemalan choirbooks, which are written on paper. Therefore, I also consulted with conservator Jessica Régimbald, coordinator of digitization Marie-Chantal Anctil, and photographer Michel Legendre from the Bibliothèque et Archives nationales du Québec (BAnQ), who have experience with digitization of cultural heritage material (from Quebec) and paper; and with Gregory Houston from the New Media & Digitization Administrator of the Digital Initiatives Department at the McGill Library, with experience on digitization of special collections and paper.

All documents listed in the first paragraph—guidelines, best practices, or standards—agree on taking the images at the highest quality one can afford and using file formats with no compression or lossless compression for master files (e.g., TIFF and JPEG2000), most of them advocating for uncompressed TIFF files.[36] The NLA and FADGI agree on resolutions of 400 ppi for manuscript sources larger than A4 (with FADGI also accepting 300 ppi values), and the use of RGB 8 bit (i.e., 24 bits per pixel) for non-color-critical images and RGB 16 bit (48 bits per pixel) for color-critical ones. DIAMM also uses a minimum capture resolution of 400 ppi. This minimal resolution value and the rest of DIAMM's baseline set of standards for minimum capture quality outlined on its "image quality" (2021) webpage are meant to ensure that a color-accurate, high-

[36] To be fair, at the time of the writing of most of the documents, JPEG2000 seemed like a promising format but was not fully recommended yet because it still had little support from leading software and web browsers. The most recent documents—DCP's and FADGI's—do include JPEG2000 as a recommendation.

quality, full-size facsimile can be reproduced from the archived image in the case the original source is lost or damaged. Riley and Fujinaga (2003) provide some recommendations for digital image capture of musical scores, with resolutions recommended for the successful OMR processing of these scores. They indicate that one needs a minimum of 3 pixels per detail for successful OMR (in the Gamera software, presented later in Section 2.3.3.2.1).[37] They found out that 600 ppi was enough to capture all significant details of most music notations, as it allows to capture details as small as 0.005in (0.0.27mm) with the required 3 pixels. They suggest that for larger printed notation, a 300 ppi value might be sufficient. They also indicate using at least 24 bits for color images and, although they do point out that TIFF is the de facto standard still used for master files in most imagining projects, they suggest the use of PNG for master images—given the many variations of the TIFF format, as pointed out in Table 2-4, which results on TIFF not being a 'true' but a 'de facto' standard.

Since only FADGI provides standards for the digitization of manuscripts together with guidelines and best-practices documentation, I took this document and the advice from DIAMM, BAnQ, and the McGill Library personnel as reference points. Table 2-5 summarizes the digital imaging specifications for bound volumes of rare and special collections found in the FADGI *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files* document. It is divided into four columns depending on the quality of the images: 4-star quality represents the state-of-the-art in image capture and 3-star quality results in very good professional

[37] Some deep-learning approaches for OMR are successful even with images that are not high resolution as is the case of the document analysis model (Castellanos, Calvo-Zaragoza, and Iñesta 2020) and the holistic staff-level symbol recognition models used in MuRET (Calvo-Zaragoza and Rizo 2018a; Calvo-Zaragoza, Toselli, and Vidal 2019), which will be introduced later in Sections 2.3.3.3 and 2.3.2, respectively. In MuRET, the images are reduced in size at the very beginning before any document analysis or symbol recognition process happen. This is also the case for the document analysis method by Calvo-Zaragoza, Vigliensoni, and Fujinaga (2017) that is used in Rodan, which will be discussed later in Sections 2.3.2 and 2.3.3.2.2. As indicated in end-to-end OMR documentation in Rodan, the image needs to be resized at the very beginning of the OMR process (http://ddmal.music.mcgill.ca/e2e-omr-documentation/hints.html#staff-size-height-and-training).

images for various uses including OCR, these two levels should be the ones aimed for to avoid re-scanning as indicated by FADGI (2016, 9); on the other hand, 2-star and 1-star quality result on images of very limited use, mostly taken for informational purposes (e.g., provide a reference to locate the original).

Table 2-5: FADGI standards for bound volumes of rare and special materials for the different performance levels (Federal Agencies Digital Guidelines Initiative 2016, 18).

	1 Star	2 Star	3 Star	4 Star
Master File Format		TIFF, JPEG 2000, PDF/A	TIFF, JPEG 2000, PDF/A	TIFF, JPEG 2000, PDF/A
Access File Formats		All	All	All
Resolution		300 ppi	300 ppi	400 ppi
Bit Depth		8	8 or 16	16
Color Space		Adobe 1998, ProPhoto, ECIRGBv2	Adobe 1998, ProPhoto, ECIRGBv2	Adobe 1998, ProPhoto, ECIRGBv2
Color		Color	Color	Color

The table shows the standards for resolution, bit depth, color space, and file formats. The FADGI Technical Guidelines is a document rich in technical details. The personnel in charge of digital image capture and quality control needs to have the appropriate technical background to achieve the different quality levels of a FADGI-compliant project; after all, "cultural heritage digitization is a specialization within the imaging field that requires specific skills and experience" (Federal Agencies Digital Guidelines Initiative 2016, 8). The institutions consulted in this project confirmed the need for the person in charge of digitization to be a professional photographer with experience in digitizing cultural heritage material: experience in photography and training on handling special collections.

2.2.2.1 ***Imaging Technology***

The imaging technologies recommended by FADGI (2016, 19) for bound volumes of rare or special collections are either manually operated book scanners or digital cameras with book cradles, both without glass or plastic platens. The scanning technologies not recommended by FADGI are flatbed scanners, automated page-turning book scanners, and linear scanning processes (i.e., digital scanning back cameras).[38] The guideline also recommend against using lighting sources that raise the surface temperature of the original by more than 2ºC (Federal Agencies Digital Guidelines Initiative 2016, 19).

Figure 2-8 shows the basic structure of a book scanner, with a book cradle, cameras, platen, and lighting. A V-shaped book cradle provides support for the back and spine of the book. It reduces the stress on the book's spine by allowing the book to be opened at an angle rather than open flat. A pair of cameras, each pointing at one of the faces of the cradle, photograph each of the two pages in the book opening. The platen—either glass or plastic—handles the page curvature; it falls on top of the pages allowing for capturing flat images. Finally, dedicated lighting is important to capture good images. The lights should be strong and even, and they must be positioned in a way that minimizes glare and reflections on the captured image.

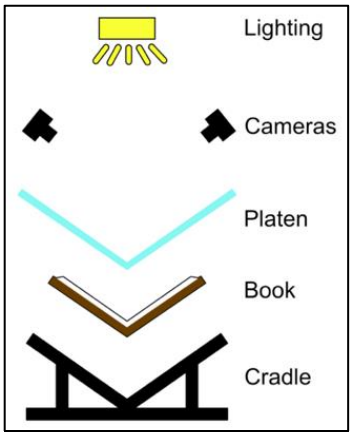


Figure 2-8: Common structure of a book scanner. This image was obtained from the DIY Book Scanner project at http://diybookscanner.org, and it is shared under [CC BY SA 4.0](#).

[38] Digital scan backs take an image by scanning it; this is, moving a sensor along an x axis rather than by capturing it in a single shot. A digital camera has a NxM matrix of (red, green, and blue) photosensors to capture the image. On the other hand, a digital scan back has a 1xN array of photosensors (a layered sensor stack detecting the red, green, and blue colors) which slides along an axis taking one exposure per point; by moving M positions, it captures an image of NxM.

While book scanners of different sizes and camera resolutions exist, with varying prices, information about how to make a do-it-yourself (DIY) book scanner is available on the internet.[39]

The use of a platen depends on the items to be digitized. In general, for bound volumes of rare and special collections, it is not recommended to use a platen unless a paper or book conservator approves it (Federal Agencies Digital Guidelines Initiative 2016, 19). The following points summarize all the feedback from DIAMM and BAnQ regarding the types of platens, implications about their use, and other options to flatten the pages of the book to be digitized. These points are marked with a 'P' for 'platen-related' advice. I provide all the advice received for others embarking on similar projects to be able to decide whether to use a platen or not knowing all the options and implications:

P1. There are two options for platens: glass or plexiglass. Régimbald, BAnQ's conservator, pointed out the pros and cons of each. A plexiglass platen is lighter than a glass one, so there is less risk of damaging the document due to the weight of the platen. On the other hand, plexiglass gets scratched while glass does not. If using a plexiglass platen, the technician should have a few sheets handy (Jessica Régimbald, email message to author, 3 January 2019).

P2. According to FADGI, using platens is generally not recommended for rare books: "special collections materials should not be placed in contact with glass or other materials in an effort to hold originals flat while imaging without the approval of a paper or book conservator. This technique can lead to physical damage to the original" (Federal Agencies Digital Guidelines Initiative 2016, 19), as can be seen in points P3 and P4.

[39] Various designs can be seen in the DIY book scanner website at http://diybookscanner.org.

P3. Avoid the use of a glass or plastic platen when working with parchment. As pointed out by DIAMM's project manager, Craig-McFeely, in documents written on a parchment surface, the ink sits on top of the surface. Because of this, when using a glass or plastic platen to flat the page, the ink sticks to the platen and is lifted off the parchment when taking the platen off (Julia Craig-McFeely, email message to author, 4 January 2019).

P4. On paper, this "ink sticking to the platen" situation is less of an issue, and one can use a platen to flatten the pages (Jessica Régimbald, email message to author, 3 January 2019). However, the issue can still happen, so the technician should verify if this is the case before deciding on using the platen (Julia Craig-McFeely, email message to author, 4 January 2019).

P5. DIAMM's photographer also advised against the use of platens reporting issues with glare/reflections, as well as distortion and color cast caused by the platen (Lynda Sayce, email message to author, 5 January 2019).

P6. Spatulas or other implements can be used to assist in holding pages flat for imaging as long as they do not obscure informational content (Federal Agencies Digital Guidelines Initiative 2016, 19).[40] Craig-McFeely pointed out that the implement used should also avoid occluding the page edges since many scholars examine these for information about ruling and trimming (email message to author, 4 January 2019). Craig-McFeely has used large sewing needles glued to round metal washers for easier hold. The needles are flexible, allowing for good pressure control, and they are thin, which makes them almost invisible

[40] Some links to spatulas for archives and preservation from University Libraries: *Indiana University Bloomington Preservation Lab* https://parkslibrarypreservation.wordpress.com/2012/03/13/the-tools-that-we-use-micro-spatulas/ and *Iowa State University* https://libraries.indiana.edu/microspatula-any-other-name….
Spatulas can be bought in Amazon or preservation/archival supplies websites like Gaylord Archival (https://www.gaylord.com/?site=gaylord) and Carr McLean (https://www.carrmclean.ca) in Canada.

in the picture (and easily removable for access files). She also pointed to the use of Porcupine quills by a colleague. Both needles and quills are thin enough to not occlude the edges of the page being digitized. As indicated in FADGI (2016, 19), if any of these implements are used, they should not be edited out of the master files.

P7. If deciding on using a glass/plastic platen, stick the color patch on the underside of the platen so that any color cast caused by it can be detected by referring to the color scale (Federal Agencies Digital Guidelines Initiative 2016, 73; Julia Craig-McFeely, email messages to author, 4–5 January 2019). As reported by the DIAMM experts, glass and regular plexiglass often cause a grey effect on a picture; although, there are plexiglass sheets specifically designed for photography.

P8. If using a plastic or glass platen, carefully verify the images to make sure that there are no glares coming from the lights reflecting off the shiny surface (Julia Craig-McFeely, email message to author, January 4, 2019).

P9. To clean the platen, use a solution of water and pure ethanol (at a ratio of 1:1) and a microfiber or a delicate-task clean wiper (e.g., Kimwipes). Make sure the platen is completely dry before putting it back in contact with the document. Since pure ethanol is difficult to obtain, one could use isopropyl alcohol at 99% purity (obtainable at pharmacies) and mix it with water at the given ratio (Jessica Régimbald, email message to author, 3 January 2019).

In addition to the implements presented in P6, an alternative configuration of the book scanner can help to keep the page as flat as possible without the use of a platen. In this configuration, the book cradle opens at one side so that the photographed page lies on a horizontal surface (see Figure 2-9).

Figure 2-9: Opening the book at one side for digitization of one page at a time. The camera is placed on top of the page to be digitized using a sturdy tripod and a horizontal arm. Lights are positioned at each side of the page for even distribution and at an angle to avoid reflections. Image from https://www.bsb-muenchen.de/einblicke/vom-buch-zum-byte-digitalisierung-an-der-bayerischen-staatsbibliothek/.

This configuration of the book cradle allows for digitizing only one page at a time, the one lying in the horizontal face of the cradle, instead of digitizing two at a time as in the configuration shown in Figure 2-8. In this case, the technician would photograph all the recto (or versos) folios first, then turn the book around and photograph all the versos (or rectos). The camera can be held perpendicular to the page digitized by using a copystand (Figure 2-10) or a sturdy tripod with a horizontal tube holding the camera (Figure 2-9). Both methods allow moving the camera up/down, keeping the distance between the camera and photographed page constant. Regarding the lighting, in this configuration, the same amount of light should be provided on each side of the page. This allows for even distribution of the light over the page to be photographed, each light source balancing the gradient of the light source at the other side. The lights should be at 45º so that their reflection is not captured by the camera.[41]

Figure 2-10: Copystand.

2.2.2.2 *Summary of recommendations*

This section contains a list of recommendations that I put together as part of the planning of this digitization project. The points are marked with a 'G' for 'general' advice and are divided into categories, starting with the staff.

Although recommendations about staff have been mentioned before, here the two main points are summarized:

[41] All the details of this book cradle configuration and their implications were presented by Marie-Chantal Anctil and Michel Legendre, coordinator and photographer of the digitization direction of Bibliothèque et Archives nationales du Québec (BAnQ), respectively, during a visit to BAnQ in 2018.

G1. "To be FADGI compliant, all imaging performed on special collections material must be done by personnel with advanced training and experience in the handling and care of special collections." (Federal Agencies Digital Guidelines Initiative 2016, 19).

G2. Furthermore, "for those working on digital image capture and quality control for images, a basic foundation in photography and imaging is essential" (Federal Agencies Digital Guidelines Initiative 2016, 8). Achieving the quality levels defined by FADGI is problematic without a staff that has a solid technical background.

Regarding the files:

G3. According to the FADGI specifications (2016, 20), image retouching is not allowed in master files, but image processing techniques can be used in the creation of the access files.

DIAMM indicates that "as long as the equipment used to view or print the image is correctly calibrated, and makes correct use of the embedded profile of the capture equipment, the user will see correct colour either on-screen or when the digital image is printed" ("Image Capture" 2021), having the "policy that if ANY post-process activity is necessary in order to make the image correct, then the image has been incorrectly taken, and must be taken again with the problem corrected in order to create an archive-quality image." ("Image Quality" 2021)

G4. "All images should have the colour profile of the capture device embedded at the point of capture" ("Image Quality" 2021).

G5. "The image must be saved in uncompressed TIF format, with no JPG or other compression format used at any point during the capture process" (Craig-McFeely 2008, 13).

Regarding the imaging technology and process:

G6. According to FADGI (2016, 19), single-exposure, total-area capture imaging systems are the most appropriate technologies for special collections, but other digital imaging technologies can be used if deemed more fitting.

DIAMM has two methods for image capture: single-shot camera and digital scanning back.[42] The former allows for a maximum resolution of 39 megapixels, obtaining archive quality images for small documents (up to A4 size) or for undamaged documents up to A3 size ("Imaging" 2021).[43] The digital scanning back is used when higher resolution is needed (132 megapixels) to digitize a larger document—A3 size and upwards—or a damaged document for which imaging is intended to provide a surrogate for digital restoration ("Imaging" 2021).[44]

G7. The book cradle angle used in the archives consulted (BAnQ and McGill Library) is about 110°. BAnQ's conservator indicated that an angle between 100° and 120° is enough if the book cannot be opened flat (Jessica Régimbald, personal communication, 26 October 2018). Of course, "bound materials must not be opened beyond the point where the binding is stressed. In some cases, that may mean that the volume cannot be opened sufficiently to image" (Federal Agencies Digital Guidelines Initiative 2016, 20).

---

[42] See footnote 38.

[43] DIAMM uses a PhaseOne P45+ as its single-shot camera. A 'single-shot' camera means that the whole image is taken in one action, working as a normal digital SLR would. The difference relies on the size of the images: a high-end consumer digital SLR captures images at 8–11 megapixels, while the P45+ captures these at 39 megapixels ("Imaging" 2021; Craig-McFeely 2008, 22).

[44] DIAMM uses a PhaseOne PowerPhase FX scanning back, mounted on a custom-built focus box supplied by ICAM Archives Systems (Craig-McFeely 2008, 22–4). The scanning back has a double row of sensors: one row of 24,000 alternating red and green sensors, and a second row of 24,000 alternating green and blue sensors. Each group of four sensors (red-green/blue-green) captures the information corresponding to one pixel. Therefore, the scanning back takes a strip image of 12,000 pixels wide. The equipment acquires an image by taking multiple strips of pictures in the same way a flatbed scanner would, with the difference that the document has no direct contact with the equipment. Obtaining a single picture with this technology takes 5–6 minutes; however, a single-shot camera for obtaining similar resolution would be highly expensive ("Imaging" 2021).

G8. The camera(s) should be perpendicular to the folio(s) to be photographed.

G9. Lenses are designed and optimized for different applications. According to FADGI (2016, 53), "generally, apochromatic and macro lenses outperform other lenses for cultural heritage close focusing applications, and lenses designed specifically for digital flat field imaging are best. It is important to assure that the image circle of the lens is appropriate for the imaging sensor used and that the lens is designed to be used at the working magnifications needed."

G10. As indicated by the digitization experts at BAnQ, Marie-Chantal Anctil and Michel Legendre, the camera parameters should be set by a professional photographer before each imaging session (including operations like white balancing).

DIAMM provides documentation as to how they set the camera focus: "The focus is adjusted minutely making no perceptible difference to the operator, and the same very small segment is captured and enlarged. The process is repeated until the enlarged result is as sharp as possible: this gives a finesse of focal length that is a considerable improvement on by-eye focus" ("Image Quality" 2021).

G11. "Lighting should ideally mimic natural daylight, with an even and stable light-spread across the source. Flash and natural light are not desirable" ("Image Quality" 2021). DIAMM's mobile studio is set up in a darkened room ("Image Capture" 2021), avoiding light sources other than the ones of the equipment.

G12. DIAMM's photographer indicated that, if not shooting tethered (i.e., shooting with the camera connected to a computer or other device), one needs to consider a way to record what the images are. Rectos tend to be numbered, which facilitates image identification;

however, versos are not (Lynda Sayce, email message to author, 5 January 2019). Shooting tethered also allows for evaluating the quality of the images in a computer rather than in the small camera monitor.

G13. Shooting with a remote control or with the computer itself reduces the camera manipulation once it is set up and will eliminate the vibration introduced by pressing the shutter button of the camera when taking the photo.

G14. Use of black felt for a neutral background (Julia Craig-McFeely, email message to author, 4 January 2019). Regardless of which material is being used for the background (the black felt or a backing sheet), it should extend beyond the edge of the page to the end of the image on all open sides of the page (Federal Agencies Digital Guidelines Initiative 2016, 19).

G15. "A tool kit consisting of gloves, spatulas, thin mylar strips, weighted hold down aids, lens wipes, air duster, soft brushes, etc. should be available at each imaging station" (Federal Agencies Digital Guidelines Initiative 2016, 17).

Regarding reference targets (for color and dimension):

G16. FADGI (2016, 72) suggests using "reference targets in each image of originals being scanned including, at a minimum: a gray scale, color reference, and an accurate dimensional scale… All types of tone and color targets should be replaced on a routine basis. As the targets are used they will accumulate dirt, scratches, and other surface marks that reduce their usability."[45]

[45] DIAMM's paper (Craig-McFeely 2008, 13) indicates that a color reference and dimensional scale must be included in each image; while DIAMM's website ("Image Quality" 2021) indicates that it should be a gray scale and the dimensional scale, with the color scale optional. A photographer at the Centro de Investigaciones Regionales de Mesoamérica (CIRMA)—the archive holding the microfilms of some of the music sources of the AHAG—indicated that the fact that the color reference includes the black and white colors would be enough for colored images. On the

FADGI (2016, 72) also provides an alternative approach for high production environments: "In a high production environment, it may be more efficient to scan targets separately and do it once for each batch of originals. The one target per batch approach is acceptable as long as all settings and operation of the equipment remains consistent for the entire batch, and any image processing is applied consistently to all the images."

G17. Keep the reference targets at the same level as the image plane (Federal Agencies Digital Guidelines Initiative 2016, 73). DIAMM's photographer, Sayce, suggested two ways to keep the color patch at the same level as the photographed page depending on the use of a platen or not. If there is a platen, stick the color patch underneath it (see suggestion P7). If there is no platen, stick the color patch at the top of a stack of cardboard sheets (painted black to match the background suggested in G14). When moving through the pages of the book, remove/add a cardboard sheet from/to the stack to change the height of the color patch and, thus, guaranteeing that it remains at the same level as the photographed folio (Lynda Sayce, email message to author, 5 January 2019).

G18. If scanning (rather than photographing), place the targets along the short dimension of the originals to avoid increasing the dimensions of the scanned area and, therefore, the file size (Federal Agencies Digital Guidelines Initiative 2016, 72).

G19. Smaller targets can be created by cutting down the full-size targes, this will allow getting the camera closer resulting in a higher resolution (Federal Agencies Digital Guidelines Initiative 2016, 73; Julia Craig-McFeely, email message to author, 4 January 2019).

other hand, Craig-McFeely indicated that the gray scale is also useful, and even more than the color chart (email message to author, 4 January 2019).

G20. Place the target close to but distinctly separated from the original so that it is easy to crop it out of the image in the access files (Federal Agencies Digital Guidelines Initiative 2016, 72).

Other considerations related to conservation:

G21. Drake, Justrell, and Tammaro (2003, 36–37) provide some advice as to how to minimize the negative effects of digitization over rare or valuable items; this includes bringing the digitization equipment to the source item rather than transporting the item itself, avoiding unbinding of bounded books, and using the appropriate scanning technology (this is, book cradle and digital cameras instead of a flatbed scanner as discussed in Section 2.2.2.1).

G22. Preparation of the material for digitization includes conservation assessment and treatment (if necessary) (Chapman 2000, 10). Items to be digitized should be reviewed by the conservation staff to determine whether or not digitization is safe (Federal Agencies Digital Guidelines Initiative 2016, 55–56), and any requirements for special handling or conservation of the source should be examined carefully prior to digitization (McIlwaine et al. 2002, 20).

BAnQ's conservator gave a list of minimal (paper) conservation treatment activities they conduct in the context of digitization; these are: surface cleaning to remove surface soil from the document, tear repair to stabilize the document for handling during digitization, flattening to give access to information that would be otherwise obfuscated from a fold (for example), and minimal repair of the structure of the book for stabilization purposes so that the book can be digitized. The documents are only treated to allow for digitization; so, if the document is stable and the information is available as is, the document is not treated (Jessica Régimbald, email message to author, 29 October 2018).

Last points:

G23. Digitization and the activities involved in it can be conducted in-house or can be outsourced. "If the organization does not have the resources, it may be best to outsource the project. Also, avoid the trap of assuming doing the work in-house will costs less. Insourcing may cost more than outsourcing" (Federal Agencies Digital Guidelines Initiative 2016, 54).

G24. For textual materials, it is common to see recommendations among the guidelines cited for OCR and the use of TEI for encoding the text content. The equivalent in music would be to use OMR for the recognition of the music symbols and MEI for their encoding. More details about these two technologies will be presented in Section 2.3.

G25. Regarding long-term preservation of digital images, as indicated in Besser (2003, 62), no one knows yet what the best preservation strategy is. Therefore, it is crucial to keep redundant copies of digital assets on different media, under archival conditions, and at different locations (Besser 2003, 28), keeping the master images for the long term and under maximum protection (Besser 2003, 62). It is also necessary to run regular checkups (yearly) on data integrity and media stability and be prepared to enter into a migration process (approximately every five years) (Besser 2003, 62). "Constant vigilance and the consistent use of open standards and system-independent formats, where possible, will be the best guarantee of the long-term viability of a digital image collection" (Besser 2003, 62).

2.3 Optical Music Recognition for Early Music

Optical music recognition (OMR) is the process of converting the digital image of a music document into a machine-readable file that encodes the music content of that image in a symbolic format (e.g., MusicXML, MIDI, MEI, or Humdrum). This process is similar to optical character recognition (OCR), used to make digital text documents readable by the computer.

The traditional approach for OMR involves a series of steps. Section 2.3.1 presents the traditional OMR workflow with each of its steps. State-of-the-art OMR approaches involving deep learning techniques are making some of these traditional steps unnecessary (e.g., staff line removal and binarization) and are also allowing to combine multiple steps into one. Section 2.3.2 presents these state-of-the-art OMR techniques that reduce the number of steps involved in the OMR workflow.

While there are several commercial OMR applications for printed and typeset music documents written in common Western music notation (CWMN),[46] there are only just a few early music notation OMR systems. Early music documents impose additional challenges for OMR that these commercial OMR applications cannot handle. Some of these challenges include the degradation of the sources, the handwritten nature of the music, the presence of different hands in a single source, and the early notation styles used. Section 2.3.3 presents OMR applications that have been used to process early music documents. These applications coordinate the different steps involved in the OMR workflow and allow for user interaction to correct the results at various stages. They have a learning-based approach, where the process improves its results through the input of the expert users. Some of these applications are designed for a specific type of early music source (e.g., mensural typeset partbooks). Others can adapt to the particular requirements of the different types of early music sources by learning to deal with the wide variety of handwritten symbols and the conservation status of the source through labeled examples.

[46] A few of them include PhotoScore by Neuratron (used in Sibelius, https://www.avid.com/photoscore-and-notateme-lite), SmartScore by Musiktek (used in Finale, https://www.musitek.com), and SharpEye (http://www.visiv.co.uk).

2.3.1 OMR Traditional Workflow

Given the complexity of the OMR task—retrieving the symbolic representation of a digital score—researchers initially approached the OMR process by dividing it into smaller achievable tasks (Calvo-Zaragoza, Hajič, and Pacha 2020). Over time, the OMR workflow described by Bainbridge and Bell (2001) and refined by Rebelo et al. (2012) became the standard. The traditional OMR workflow involves the four main phases shown in Figure 2-11, each of them including several subprocesses.

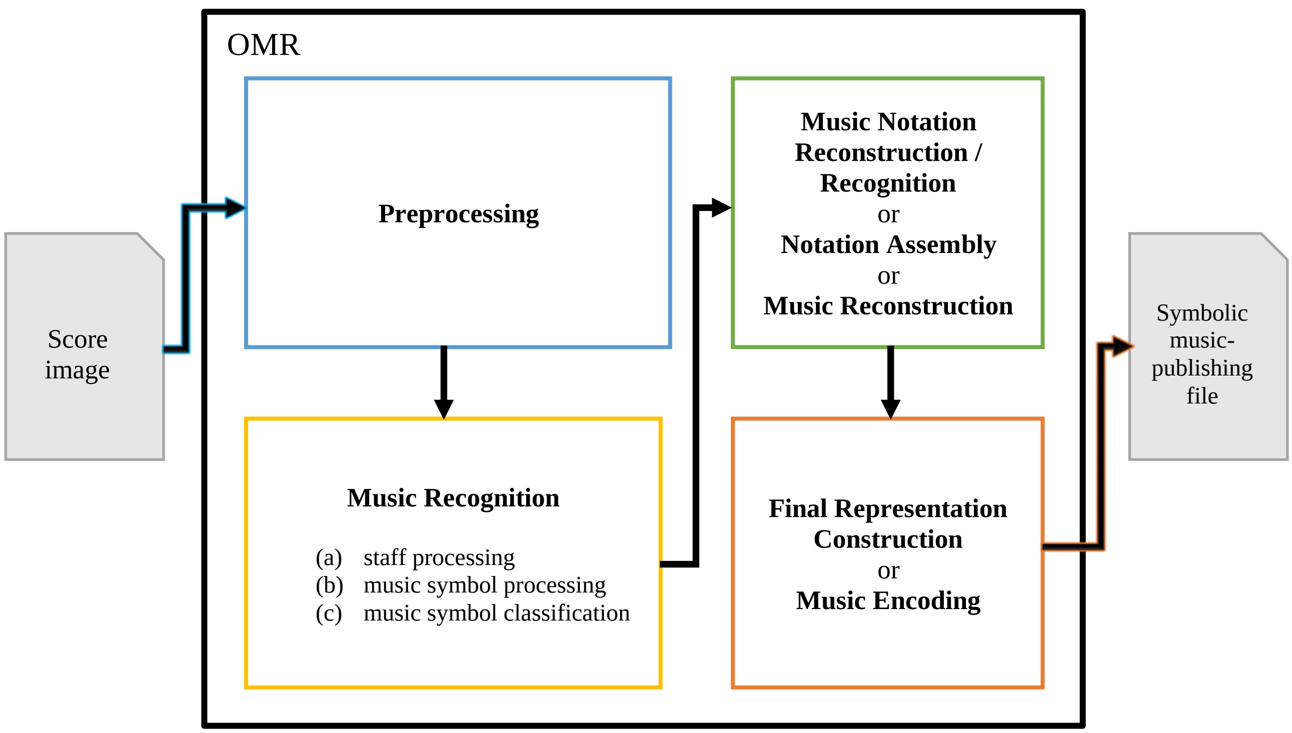


Figure 2-11: General OMR workflow architecture, based on Bainbridge and Bell (2001), Rebelo et al. (2012), and Calvo-Zaragoza, Hajič, and Pacha (2020).

The **preprocessing** phase consists of image processing operations such as noise removal, blurring, deskewing, contrast adjustment, sharpening, morphological operations, and binarization. The afore-mentioned operations are meant to facilitate the music recognition phase. The material digitized may exhibit different levels of degradation—this is especially true for cultural heritage material—that can affect the performance of the following stages of the OMR workflow. For example, broken lines on the symbols due to degradation will affect the music symbol processing step. In this case, applying blurring or morphological operations (e.g., dilation) can help overcome this issue (Droettboom et al. 2002). Since the goal of these operations is to facilitate the music recognition phase, they are not to be applied to the image if they are not deemed necessary (e.g., deskewing is not needed if the image was taken correctly in the first place). Until recently, binarization was one crucial operation within the preprocessing stage. It was applied to the image at the beginning to segment it into two layers: background (which does not contain any pertinent information for OMR purposes) and foreground (where all the music and text content of the image lies). Recent developments, explained in more detail in Section 2.3.2, made this preprocessing operation optional.

The **music recognition** phase is divided into three steps. The first one is the **staff processing**, which consists of detecting and removing the staff lines to allow for the isolation of the music symbol layer used in the following two steps. The second step is the **symbol processing**. In Rebelo et al. (2012) and Bainbridge and Bell (2001), this step involves the segmentation of symbols into primitive elements (e.g., note heads, stems, and tails), which will be assembled afterward during the music notation reconstruction stage. In other studies, the symbol processing step consists of the use of connected component analysis to isolate the different symbols by automatically grouping the connected components into glyphs (Huang et al. 2015; Pugin et al. 2008). The third and last step is the **symbol recognition/classification**. Here a classifier is trained on a set of labeled examples to recognize the different classes of music symbols in the document. These last two steps, symbol processing and classification, are usually carried out together.

The last two phases in the OMR workflow, music notation reconstruction and final representation construction, are intertwined. The **music notation reconstruction** or **notation assembly** phase consists of extracting the musical semantics from the recognized symbols by defining the musical structure and interpreting the spatial relationships between symbols. *Defining the musical structure* involves detecting music phrases, measures, systems, and other structural elements. *Interpreting the spatial relations between symbols* requires the reincorporation of the staff lines. Together with the music symbols, the staff lines allow for: encoding the pitch of each note, determining the function of a given accidental (which notes are affected by its presence), distinguishing ties from slurs (when a curved line happens between consecutive notes of the same pitch it is a tie rather than a slur), etc. Another example of the spatial relation between symbols revealing their meaning is the case of dots accompanying a note in CWMN. When a dot is placed at the top or bottom of a note head, it indicates a performance instruction (staccato). On the other hand, when a dot is placed at the right side of the note head, it indicates a change in the note's value (a dotted note).

Finally, in the **final representation construction** or **music encoding** phase, the derived notation representation obtained in the previous stage is converted into a machine-readable encoding format, translating the information into the standardized structure of a symbolic music format. Music symbolic formats that can be used for this phase include MIDI, MusicXML, MEI, etc. The Music Encoding Initiative (MEI) tends to be the chosen symbolic format for encoding early music, as will be seen in the OMR applications presented in Section 2.3.3. MEI is a symbolic format developed by a community of scholars—including librarians, musicologists, and technologists—interested in encoding a wide variety of music documents (Hankinson, Roland, and Fujinaga 2011; Roland, Hankinson, and Pugin 2014). The symbolic format supports encoding various notations, including common Western musical notation (CWMN), mensural notation, neume notation, and tablature. It also has the flexibility to be expanded to encode other repertoires. Due to its diverse community, MEI supports encoding a wide variety of information through its different modules. In addition to various Western notations, MEI allows for encoding very rich metadata, scholar editions, and relations between the encoded music and the facsimile or recording containing the music (relating the notes to the specific region they appear on the page or to their timestamp in an audio recording). The facsimile functionality of MEI has been used in the OMR systems presented in Section 2.3.3 to highlight the regions of the recognized music symbols in the digital image.

2.3.2 Recent Developments: Combination of Multiple Steps into One

Until recently, binarization and staff line removal were considered necessary operations prior to the music symbol processing and classification. These two processes are now deemed unnecessary due to the introduction of a deep learning *document-segmentation algorithm* that allows separating the background, staff-line, text, and music-symbol layers of an image in a single step (Calvo-Zaragoza et al. 2018; Calvo-Zaragoza, Vigliensoni, and Fujinaga 2017). The algorithm uses a Convolutional Neural Network (CNN) trained on labeled pixel data. When the trained CNN receives a new image, it labels its pixels as belonging to one of the four categories: background, staff line, text, or music symbol. Later on, Castellanos et al. (2018) substituted the CNN approach with a Selectional Auto-Encoder (SAE), which increased the efficiency of the classification process (from hours to minutes). These document-segmentation algorithms are considered preprocessing operations, leaving the music recognition phase with only two steps: music symbol processing and music symbol classification. This is the case in the OMR workflow used by the *Single Interface for Music Score Searching and Analysis* (SIMSSA) project (Figure 2-12), a project that attempts to devise tools and techniques for searching and analyzing digitized music documents (Fujinaga and Hankinson 2015).[47]

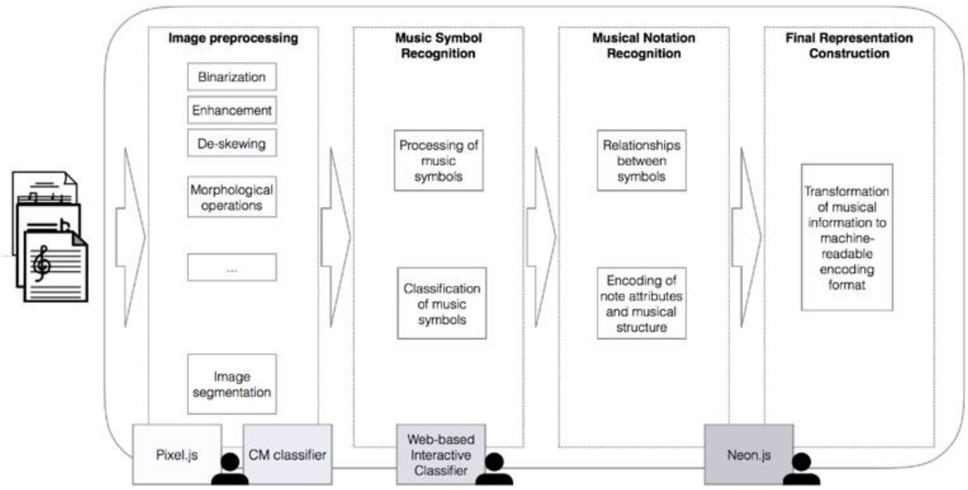


Figure 2-12: Overview of the phases of the OMR workflow from the SIMSSA project and the corresponding user interfaces to correct the output at each stage. The user interfaces are indicated by the bottom boxes with a human silhouette on their right side (Vigliensoni, Calvo-Zaragoza, and Fujinaga 2018).

Other deep learning algorithms that combine several steps of the OMR workflow into one have been introduced in recent years. As shown in Figure 2-11, the music recognition stage typically involved staff line removal because the presence of staff lines makes it hard to separate isolated symbols through connected component analysis. However, deep learning models with CNNs have proven successful in dealing with the music recognition stage holistically without removing the staff lines (Calvo-Zaragoza, Hajič, and Pacha 2020). Two of these holistic methods are the staff-level recognition method and the page-level recognition method, both tested over Spanish manuscripts in mensural notation.

The holistic staff-level recognition method, introduced by Calvo-Zaragoza and Rizo (2018b), recognizes the sequence of symbols in a staff as a whole, provided that there is only one

[47] https://simssa.ca/

voice per staff. This method was applied to printed CWMN (Calvo-Zaragoza and Rizo 2018b) and handwritten mensural notation (Calvo-Zaragoza, Toselli, and Vidal 2019). Both staff-level holistic algorithms use a convolutional recurrent neural network (CRNN), which consists of a block of convolutional neural network (CNN) layers followed by a block of recurrent neural network (RNN) layers. The convolutional layers are in charge of extracting relevant image features, and the recurrent block interprets these features in terms of sequences of music symbols. The sequence of symbols in the staff outputted by these aforementioned CRNN algorithms provides two pieces of information per symbol: the symbol class and its position—numbered line/space—in the staff (see Figure 2-13). The training data for these holistic staff-level algorithms consists of pairs of staff region images and the transcript of their music sequence.

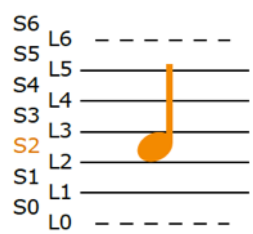


Figure 2-13: Encoding vertical staff line positions into discrete categories (Pacha and Calvo-Zaragoza 2018).

The page-level recognition method, introduced by Pacha and Calvo-Zaragoza (2018), recognizes all the symbols present on a page, providing both the symbol class and its position in the staff (see Figure 2-13). This page-level approach uses a region-based convolutional neural network (R-CNN), initially designed for object detection. In Pacha and Calvo Zaragoza (2018), the R-CNN is trained for the direct detection and categorization of mensural music symbols. The training data consists of score images with their corresponding set of symbols and regions (i.e., bounding boxes) in which they appear.

These two holistic methods (staff and page level) reduce the steps needed in the OMR workflow by removing the necessity of some of these steps and collapsing multiple steps into one. None of these two recognition methods require staff-line removal (they work at the staff or the page level with the staff lines intact). The music symbol processing step, the music symbol classification step, and most of the pitch detection step, which is a fundamental part of the music reconstruction stage, are collapsed into a single (bigger) step.

2.3.3 Frameworks and Workflow Managers for OMR

Given the various steps of the OMR workflow, a workflow manager or an OMR framework is needed to conduct the workflow in a semiautomatic way. A few OMR applications exist that can deal with early music. These are **Aruspix**, the **Gamera** document recognition framework, the **Rodan** workflow manager, the **MuRET** online OMR framework, and the **OMMR4all** online OMR framework.

2.3.3.1 *The Aruspix Project*

Aruspix is an open-source desktop application for the OMR, superimposition, and collation of early typographic music prints (i.e., music printed mainly during the sixteenth and seventeenth centuries with movable typefaces) to prepare digital editions.[48] It was developed by Laurent Pugin in close collaboration with the Marenzio Online Digital Edition (MODE) and the Music Encoding Initiative (MEI).[49] Pugin initially designed Aruspix to help editors prepare printed critical editions

[48] Aruspix can be downloaded as a desktop application for Mac OS X (https://www.aruspix.net/downloads.html). It also has a command-line version that can be set up in either Mac OS X or Linux (https://github.com/DDMAL/aruspix/wiki/03-%E2%80%93-Aruspix-Command-line).

[49] The Marenzio Online Digital Edition (MODE) is an online critical edition project of Luca Marenzio's (ca. 1553–1599) secular music. The project started in 2006. It uses Aruspix to collate the music sources. MODE encodes the contents of these sources following the standards of the Music Encoding Initiative and presents the encoded music online through Verovio (an open-source engraving library). http://www.marenzio.org/index.xhtml

The Music Encoding Initiative (MEI) is a scholarly community-based effort to define a system for encoding a wide variety of music documents in a machine-readable structure. https://music-encoding.org/

by facilitating the exhaustive comparison of different sources using Aruspix's OMR, superimposition, and collation functionalities.[50] Eventually, Aruspix became a digital edition project (Pugin 2009).

Aruspix's OMR functionality has two steps: (1) the preprocessing or document segmentation step, and (2) the symbol recognition step. Aruspix's preprocessing step consists of segmenting the document into different layers: an in-staff music symbol layer (in black), a title elements layer (in yellow), a lyrics of inter-staves music symbols layer (in orange), an ornate letters layer (in dark green), an in-staff text elements layer (in light green), and a blank layer (in white). The user can correct the results of the document-segmentation process within the Aruspix app (see Figure 2-14). After completing the preprocessing step, the user can run the symbol recognition step and use the editor and superimposition feature of Aruspix to correct the recognized symbols (correcting either the note shape identified or its position in the staff, see Figure 2-15).

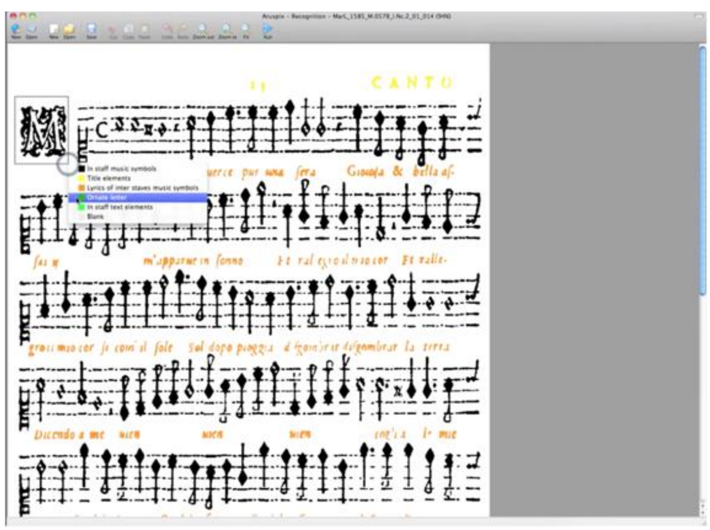


Figure 2-14: Results of the automatic preprocessing step in Aruspix. In this example, the user is correcting the results by selecting the ornamental letter—currently classified as part of the 'in staff music symbols' layer (in black)—and assigning it to the correct layer (ornate letter layer in dark green). *https://www.aruspix.net/about.html*

[50] Aruspix allows for comparing different copies of an edition through its superimposition functionality. It also allows for comparing different re-editions of the same work through its OMR and collation features (Pugin 2009).

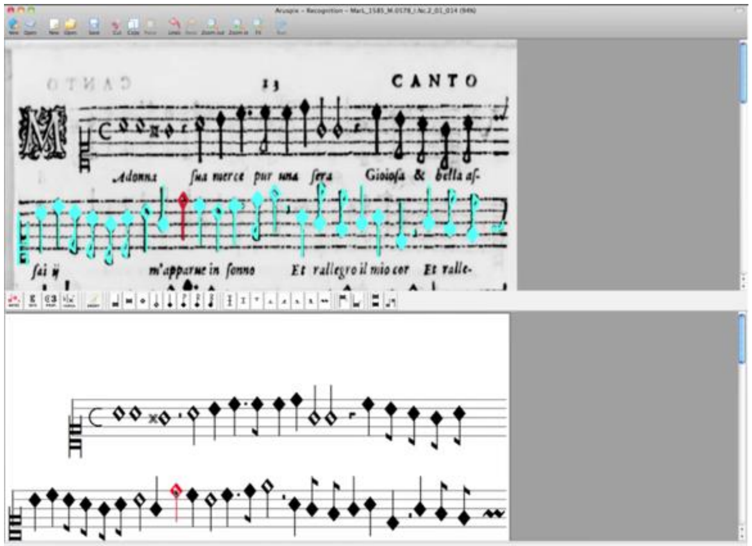


Figure 2-15: Results of the automatic symbol recognition step in Aruspix. The user can edit the results within the application, activating Aruspix's superimposition feature to facilitate the correction process (see second system in the top panel). *https://www.aruspix.net/about.html*

Aruspix's music recognition algorithm is based on Hidden Markov Models (HMMs), used to model sequential data. The use of HMMs in OMR was inspired by their use in OCR (Pugin 2006; Rebelo, Capela, and Cardoso 2010). Aruspix uses the typeface—the music symbol with its staff lines—as the recognition unit; each typeface is presented by an HMM (Pugin 2006). This allows to process a music region without staff removal and to predict the symbols and pitches in one step. As soon as a page is processed and corrected by the user, Aruspix dynamically learns and improves itself.

Aruspix has been used in the MODE project for encoding the secular music of Luca Marenzio (ca. 1553–99) printed in white mensural notation (Pugin 2018). It has also been used on the Early Music Online (EMO) collection—324 volumes of sixteenth-century printed music held at the British Library, most of them partbooks in white mensural notation—which has been digitized from microfilms (Pugin and Crawford 2013).

2.3.3.2 *The Gamera Framework and the Rodan Workflow Manager*

2.3.3.2.1 The Gamera Document Recognition Framework

Gamera is a framework for building custom document recognition systems. The goal is to create custom applications that fit the document being processed rather than attempting to meet the diverse demands of different documents with a monolithic application (Droettboom, MacMillan, and Fujinaga 2003). Gamera is an open-source, cross-platform desktop application.[51] It has a modular architecture, where each module focuses on a specific stage of the workflow: preprocessing, document segmentation and analysis (e.g., staff line identification and removal for music documents, and figure removal for text documents), symbol segmentation and classification, syntactical or structural analysis, and output. These modules can be expanded through a plugin system. Gamera has a workflow system that allows the user to tie tasks together to build a complete recognition system for the document at hand. It also has support for batch processing.

The core of the Gamera system is the segmentation, feature extraction, and classification of symbols. Different classifiers can be plugged in. These classifiers have two flavors: non-interactive and interactive. In the latter, the user can correct the results and provide more training data for the classifier to improve its performance. The included classifier is a k-nearest neighbor (k-NN) classifier. In a k-NN classifier, a glyph is assigned the most common class among its k nearest neighbors in a multidimensional space. In Gamera, the weights of the features that define the multidimensional space can be optimized for classification using a genetic algorithm (GA).

For OMR purposes, Gamera has been used for building a lute tablature recognition system (Dalitz and Karsten 2005), a measure recognition system for CWMN (Vigliensoni, Burlet, and Fujinaga 2013), and a CWMN OMR system for the Lester Levy collection (MacMillan,

[51] It can be downloaded at https://gamera.informatik.hsnr.de/.

Droettboom, and Fujinaga 2001; Choudhury et al. 2000). It has also been used to process the handwritten adiastematic neume notation found in the St. Gallen manuscript 390 / 391 (Hartker's Antiphoner) in the context of the Optical Neume Recognition Project (Helsen and Behrendt 2016).[52]

2.3.3.2.2 The Rodan Workflow Manager

Rodan is a workflow management system developed by Andrew Hankinson (2014) for OMR, inspired by the IMPACT Project workflow manager for OCR. Rodan can integrate different recognition systems, such as Aruspix and Gamera, with other OMR-related systems (Fujinaga, Hankinson, and Cumming 2014). As will be seen later, several tools from Gamera have been adapted to work within Rodan. Nowadays, Rodan can manage other kinds of workflows besides OMR (e.g., music analysis workflows).

Rodan is a management system for designing and executing cloud-based workflows. While the user can manage its document recognition workflow from any device connected to the Internet, the computationally intensive recognition tasks are performed on the server side (Fujinaga and Hankinson 2015). In Rodan, users define workflows by chaining together a set of available processing jobs (Hankinson 2014). Once the workflow has been defined, Rodan follows the progress of the image document through the different stages of the workflow.

Rodan allows for interactive jobs—these are jobs that require human input/feedback—which are useful in certain stages of the OMR process to either provide data or correct mistakes before proceeding to the next step (Hankinson 2014). Users are notified when interactive jobs require their intervention. Figure 2-12 showed a few of the interactive jobs included in Rodan.

[52] https://www.cs.bham.ac.uk/~aps/research/projects/neumes/

Rodan can include any software/algorithm as a job (a step) on the workflow as long as a Rodan wrapper is provided.[53] Rodan wrappers have been written already for several OMR steps. Figure 2-16 presents the complete OMR workflow in Rodan as used by the SIMSSA project.[54]

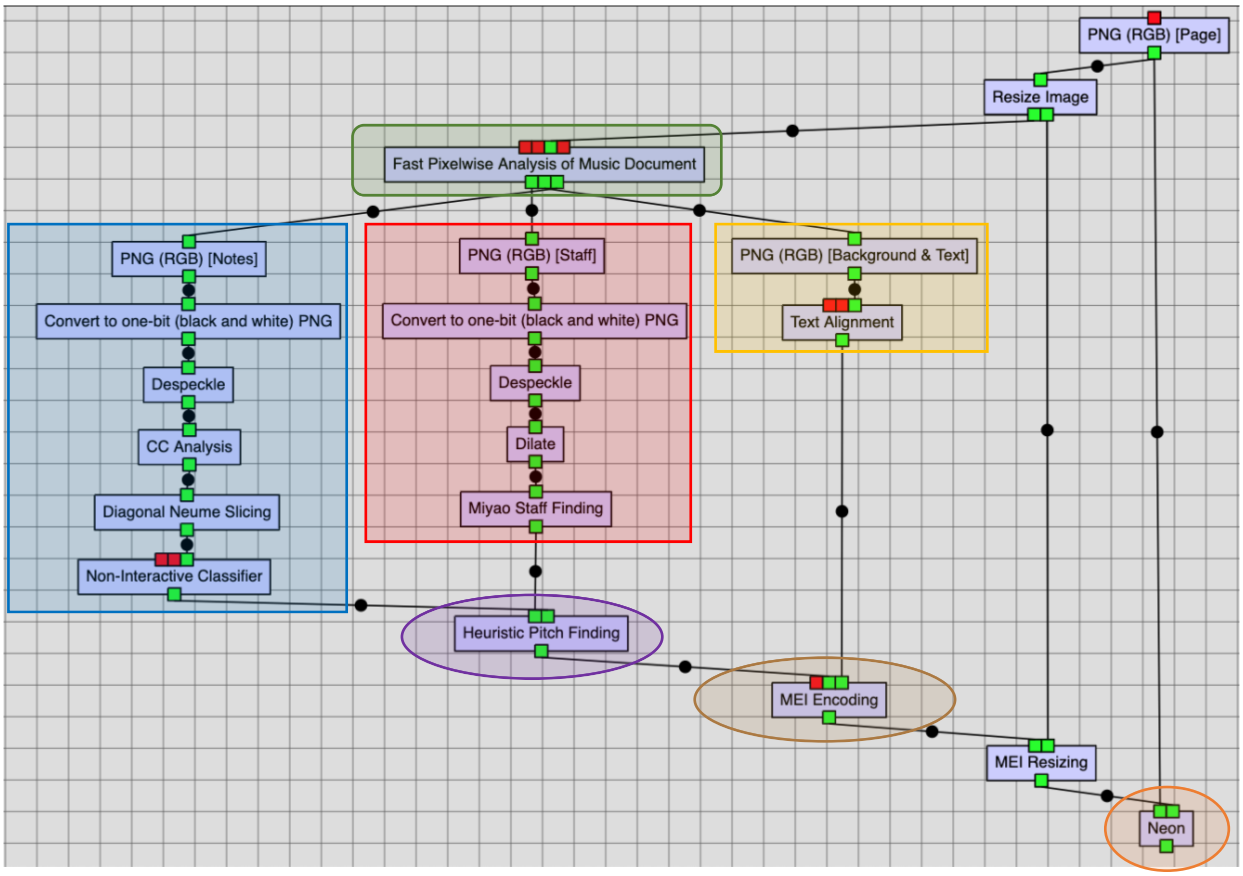


Figure 2-16: SIMSSA's full OMR workflow used to process neume square notation documents. Most of the jobs involved are adaptable to other music notations.

Key: The document analysis (green box) segments the image in three layers: music symbol, staff-line, and other. Each layer is processed by a different path in the OMR workflow (see blue, red, and yellow boxes). The blue box shows some preprocessing operations and the music recognition stage performed over the music symbol layer. The red box shows some preprocessing operations performed over the staff-line layer, followed by a staff-line tracking job. The output of the blue and red paths is combined in the pitch finding job (purple oval), which is part of the music reconstruction OMR stage. The text is extracted from the 'other' layer (yellow box). All this information is put together on an MEI file (brown oval), which can be corrected using the Neon job (orange oval), completing the final reconstruction stage. http://ddmal.music.mcgill.ca/e2e-omr-documentation/hints.html

[53] https://github.com/DDMAL/Rodan/wiki/Write-a-Rodan-job-package

[54] Details about the individual jobs involved in the OMR workflow can be found at http://ddmal.music.mcgill.ca/e2e-omr-documentation/.

The document analysis stage relies on pixel-wise labeling, which is performed automatically by the SAE neural network presented by Castellanos et al. (2018), the results of which can be manually corrected with the Pixel.js job (Saleh et al. 2017).[55] Pixel.js is also used to prepare the ground truth to train the document-segmentation algorithm (see Figure 2-17).

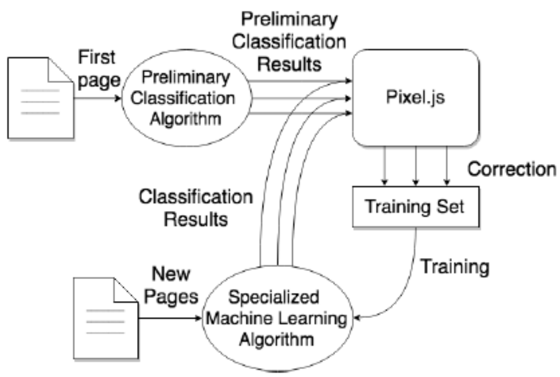


Figure 2-17: Suggested human-aided workflow using Pixel.js (Calvo-Zaragoza et al. 2017).

Each of the layers in which the image is segmented—music symbols, staff lines, and other (which includes the text, illuminations, and background)—is processed by a different path in the OMR workflow, as shown in Figure 2-16. The music symbols layer is processed by the path shown in the blue box. The music recognition stage is performed by the last three jobs in the blue box. The CC (Connected Component) Analysis job—standing for connected components analysis, which groups the connected pixels into glyphs—and the optional Diagonal Neume Slicing job perform the music symbol processing step.[56] This is followed by the music symbol classification step, performed by the k-NN symbol classifier adapted from the Gamera framework.[57] The

[55] The SAE is divided into two jobs, a training job and a classification job. There are also training and classification jobs implementing the CNN approach presented by Calvo-Zaragoza, Vigliensoni, and Fujinaga (2017). The CNN approach requires less training data to obtain a good model, but the classification time per page is in the range of hours, while the SAE classification time per page is in the range of minutes (Castellanos et al. 2018).
[56] The Diagonal Neume Slicing job is useful when dealing with neumes in square notation to divide the neumes into their individual components, reducing the number of classes of symbols to classify in the next step.
[57] Several Gamera functionalities have been wrapped as Rodan jobs. This includes image preprocessing functions such as binarization and morphological operations (e.g., dilation and despeckle), and music symbol processing and classification functions such as CC Analysis and the k-NN classifier. The *Miyao staff finder* job included in Rodan

classifier job has two flavors: non-interactive and interactive. In the latter, the user can correct the results and provide more training data for the classifier to improve its performance.

The staff line layer is processed by the path shown in the red box in Figure 2-16, which includes some preprocessing operations and a job to track the trajectory of each staff line (to be used later for pitch finding). The 'other' layer is processed by the path shown in the yellow box. The Text Alignment job requires the user to provide the transcript text and an OCR model to perform transcript alignment (i.e., to determine where each syllable of the transcript text appears on the image). The OCR model retrieves an inaccurate transcript from the text in the image, but it contains the correct positions for most characters. A global sequence alignment method is used to match up the known correct transcript with the imperfect OCRed text, which results in transcript alignment (de Reuse and Fujinaga 2019).

Moving into the last two stages of the OMR workflow (music reconstruction and music encoding), the classified music symbols and the staff lines are combined to determine the pitch of the individual notes using a set of heuristics presented by Vigliensoni et al. (2011). The MEI Encoding job, currently designed for square neume notation, combines the (pitched) neume components into neumes and the neumes into syllables following the MEI schema for encoding neume notation. The resulting MEI file can be corrected using the Neume Editor Online (Neon) job, a web app that allows for the superimposition of the encoded MEI neumes over the original image, facilitating the correction of the OMR results (Burlet et al. 2012; Regimbal et al. 2019; 2020).

Rodan has been used to perform OMR on square neume notation sources, such as the Salzinnes Antiphonal manuscript and the printed Liber Usualis (Fujinaga and Hankinson 2015;

for staff-line detection (Figure 2-16, red box) is part of the algorithms provided by the Gamera toolkits (specifically, the MusicStaves Toolkit).

Hankinson 2014).[58] Most jobs in the Rodan OMR workflow are notation-independent. The machine-learning jobs of document-segmentation and the Gamera-based classifier are trained on different documents to work with various sources and glyphs. Most other jobs are independent of the content of the image (e.g., preprocessing jobs like binarization, despeckling, and dilation). The only jobs in the Rodan workflow that are notation-dependent are the MEI Encoding job and Neon, designed for square neume notation.[59]

2.3.3.3 *The MuRET OMR Framework*

The Music Recognition Encoding and Transcription (MuRET) is an online OMR framework designed as a machine-learning-based research tool, allowing different OMR algorithms to be plugged in. MuRET is developed by David Rizo in the context of the HISPAMUS project, which focuses on preserving handwritten Spanish music heritage through automatic transcription (Iñesta et al. 2018). The outcomes of MuRET are both the symbolic files of the transcribed music and the data for the evaluation of the machine-learning algorithms involved in the transcription process (Rizo, Calvo-Zaragoza, and Iñesta 2018; Iñesta, Rizo, and Calvo-Zaragoza 2019).

All processing and recognition stages used in MuRET employ state-of-the-art machine learning techniques. The document segmentation stage is performed using a Selectional Auto-Encoder (SAE) that directly detects staff regions, as presented by Castellanos, Calvo-Zaragoza, and Iñesta (2020). The symbol detection and recognition stage is performed in an end-to-end (holistic) fashion, staff by staff, using the holistic staff-level recognition methods mentioned in

[58] The results of the OMR process on the Liber Usualis can be found at http://liber.simssa.ca/, while the Salzinnes Antiphonal can be found in the Cantus Ultimus website at https://cantus.simssa.ca/manuscript/133/. Both books are searchable as a result of the OMR process.

[59] A legacy version of Neon had support for staffless neume notation. The GitHub repository of this version has been archived: https://github.com/DDMAL/Neon.js-Legacy-. A demo of this legacy version of Neon can be found at http://neon.simssa.ca/.

Section 2.3.2. These are the methods introduced by Calvo-Zaragoza and Rizo (2018b) for CMN, and the one by Calvo-Zaragoza, Toselli, and Vidal (2019) for mensural notation.

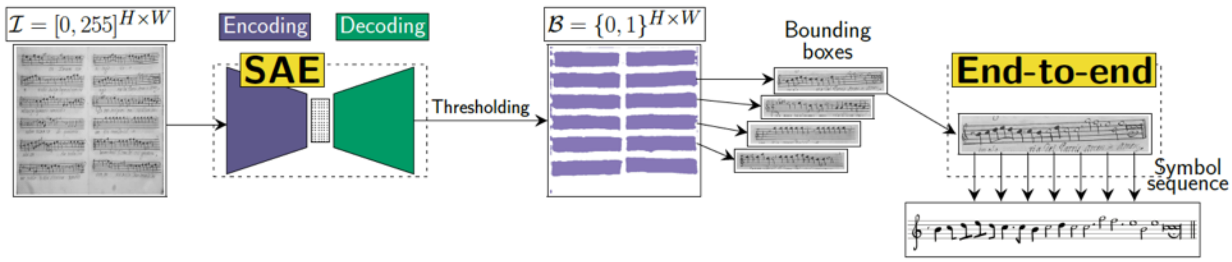


Figure 2-18: Staff-region retrieval by SAE and following symbol recognition by holistic (end-to-end) staff-level recognition method. Image from Castellanos, Calvo-Zaragoza, and Iñesta (2020).

The user also has the option to use a symbol-level recognition method. By tracing the shape of the music symbol with an e-pen or drawing a bounding box around the target symbol, the classifier will identify the symbol class and its vertical position in the staff using two CNNs—one for each classification task (Iñesta, Rizo, and Calvo-Zaragoza 2019). This symbol-level recognition method allows the user to add any symbols missed by the holistic staff-level recognition process. Both the staff-level and the symbol-level recognition algorithms provide two pieces of information for each glyph: symbol class and vertical position (numbered line/space) in the staff.

The sequence of symbols in a staff obtained by the symbol recognition stage represents an “agnostic sequence.” The term was introduced in the OMR context by Calvo-Zaragoza and Rizo (2018a; 2018b) to refer to a sequence containing symbols characterized by their graphic value (glyph class) and position in the staff, but with no musical meaning. For example, in an agnostic sequence, a series of sharps is just a sequence of symbols, whereas their role as part of a key signature is not represented. The music reconstruction stage in MuRET transforms these agnostic sequences into semantic ones (encoded in **kern format for CMN and in **mens format for mensural notation).[60] This step is performed using machine translation techniques (Ríos-Vila et al. 2021). While **kern and **mens are already standard symbolic formats, MuRET can convert these into MEI or MusicXML encodings.

Section 3.2 presents the MuRET interface for conducting each of the OMR stages—document analysis, symbol recognition, and semantic encoding—and correcting their output. As mentioned before, MuRET was conceived as a research tool—allowing one to plug in different algorithms, test them, and compare them with appropriate metrics—not as a publishing software; therefore, no batch processing is implemented (Iñesta, Rizo, and Calvo-Zaragoza 2019). At this moment, MuRET does not include any functionality for processing lyrics.

MuRET has been used to process some mensural notation sources, as is the case with a manuscript from Zaragoza (Calvo-Zaragoza, Rizo, and Iñesta 2016; Rizo, Calvo-Zaragoza, and Iñesta 2020) and, more recently, a few music books at the Biblioteca Nacional de España (BNE) (Rizo et al. 2022).

2.3.3.4 *The OMMR4all Online OMR Framework*

OMMR4all is an online OMR and correction framework for Medieval neume notations that uses state-of-the-art deep learning algorithms (Wick and Puppe 2019b). It has been used to accelerate the transcription of monophonic pieces in the Corpus Monodicum Project (Wick, Hartelt, and Puppe 2019).[61] The OMR workflow in OMMR4all is shown in Figure 2-19. It requires the user to provide the image of a page to be processed and the transcribed lyrics present on the page. The

[60] **kern and **mens are Humdrum formats to encode common Western music notation and mensural notation, respectively.

[61] The Corpus Monodicum project of the University of Würzburg is dedicated to the research and edition of European Middle Ages monophonic music with Latin text. https://corpus-monodicum.de/.

output is an MEI file encoding the music and lyric content of the images according to the MEI specifications for neume notation.

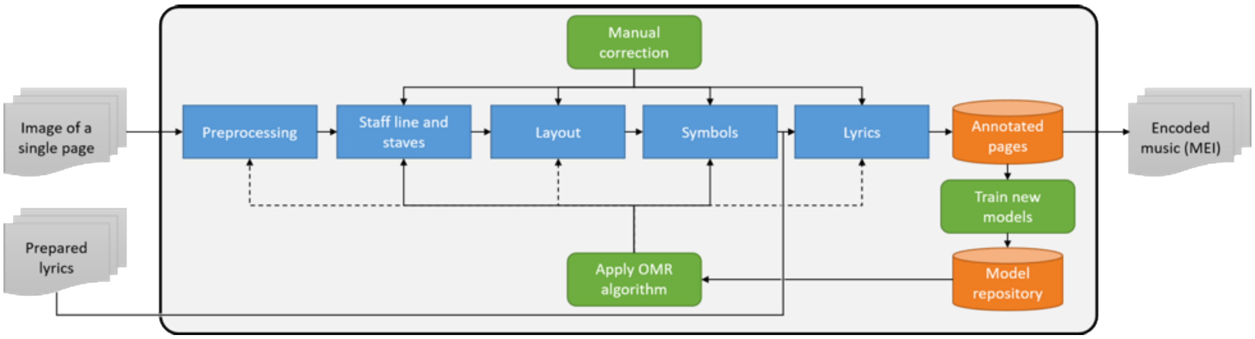


Figure 2-19: OMR4all workflow. The steps of the workflow are in blue, while the places involving human interaction are shown in green. The orange cylinders show the storage for annotated pages or trained models. Dashed arrows indicate that these algorithms currently do not rely on a model from the 'model repository' (Wick and Puppe 2019b).

The preprocessing step applies OCRopus's binarization and deskewing algorithms.[62] The next step identifies the staff regions. This is done by identifying the pixels belonging to a staff line using a fully convolutional network (FCN), combining them into a polyline representing a single staff line, and combining these lines into individual staves based on their relative distances (Wick, Hartelt, and Puppe 2019). The layout analysis step identifies the regions containing lyrics by assuming that the area between two staves belongs to a lyrics' region. The symbol detection (and recognition) step also makes use of an FCN, which acts at the staff level (Wick, Hartelt, and Puppe 2019). The syllable assignment step assigns the transcribed lyrics to groups of one or more neumes using the output of Calamari's OCR (Wick, Reul, and Puppe 2018). The OMMR4all web app provides support for correcting the output of any of these steps through an overlay editor (see Figure 2-20).

[62] OCRopus is a free document analysis and optical character recognition (OCR) software. https://github.com/ocropus.

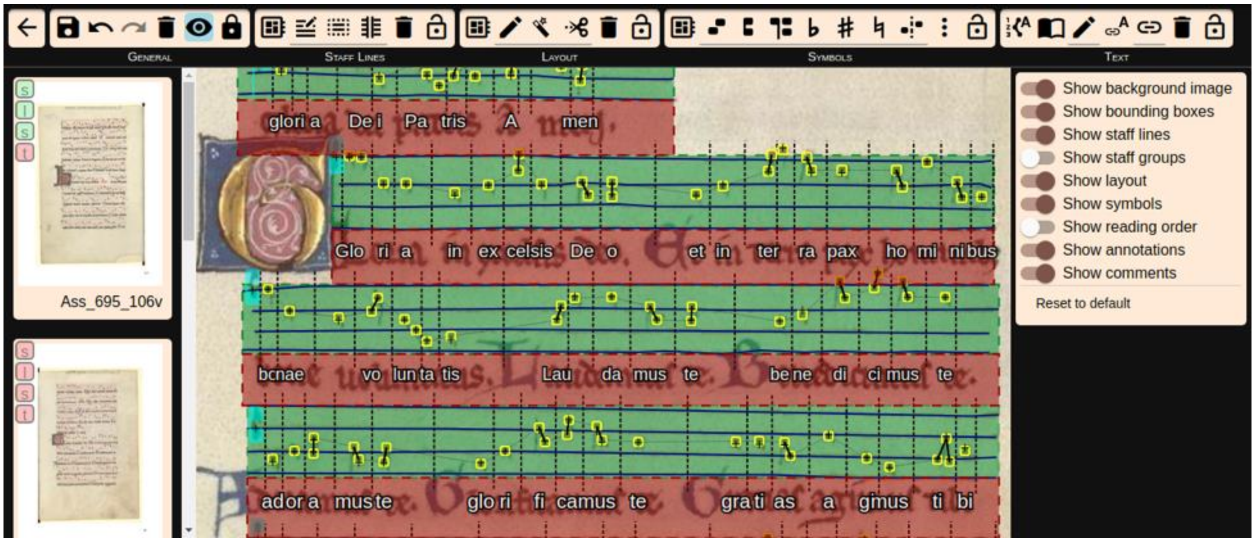


Figure 2-20: OMMR4all overlay editor for corrections. The various buttons of the toolbar define tools to correct the annotations or launch the automated tools (Wick and Puppe 2019b).

Key: staves (green regions), lyrics (red regions), individual note/neume components (yellow boxes), graphical connections for individual neumes (solid lines connecting note components), start of a new neume (vertical lines), and clefs (cyan boxes).

Similar to MuRET, OMMR4all's symbol recognition stage works at the staff level, and its document-segmentation process also involves identifying these staff regions (also the case in Aruspix). The treatment of lyrics is similar to that of the Text Alignment job used in the Rodan OMR workflow, requiring a prepared transcript of the lyrics and the output of an OCR algorithm to align the text with the music. Similar to Neon and Aruspix, the superimposition feature of the interface facilitates the correction of the OMR results. As in other OMR applications presented here (e.g., the Gamera symbol classifier and the document-segmentation algorithm in Rodan), OMMR4all allows to iteratively train its models (for staff-region detection and symbol recognition) to adapt them to unseen notation styles or music documents. As is the case with Rodan, OMMR4all allows the user to manage its document recognition workflow from any device connected to the Internet, while the computationally intensive tasks (e.g., the training of models) are performed on the server side. The web apps of MuRET, OMMR4all, and the Rodan workflow manager do not require any installation, which lowers the barrier for musicologists to use these tools. All these applications use MEI as the final symbolic format for encoding the recognized music since MEI supports the encoding of early music notation values (both neumes and mensural) and the encoding of the facsimile regions of the recognized symbols.

2.4 Mensural Notation and the Voice Alignment Problem

2.4.1 Mensural Notation: A Summary

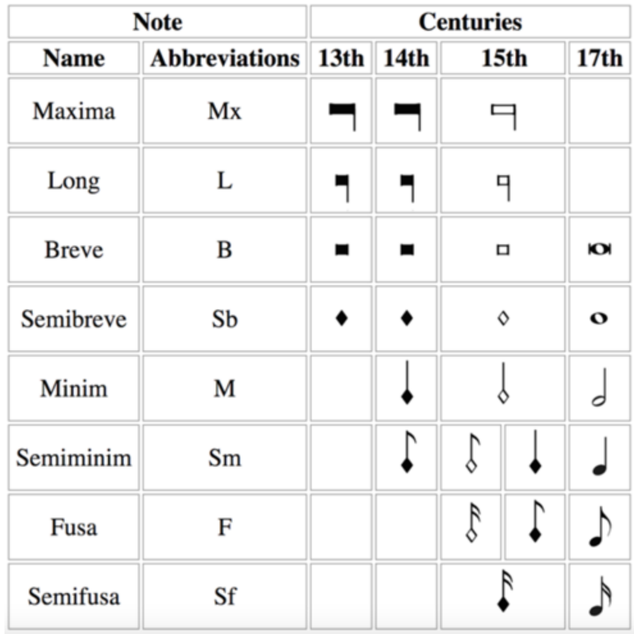
This section includes a summarizes the key concepts of mensural notation: note values, mensuration, imperfection and alteration, dots of division and addition, and coloration. The content of this section comes from my Master's thesis (Thomae 2017) and a paper (Thomae, Cumming, and Fujinaga 2019).

Mensural notation was used during the Late Middle Ages and the Renaissance to notate vocal polyphonic music. Mensural notation began around the mid-thirteenth century. Before that rhythm was conveyed by patterns of ligatures (i.e., groups of two or more notes joined together), known as "modal notation" or the system of the rhythmic modes (Roesner 2001). Mensural notation introduced the use of different note shapes to represent different note values (as in Common Western Music Notation, CWMN) (Parrish 1978, 108–9). At the same time the use of distinct note shapes was introduced, the score arrangement was abandoned in favor of notation of individual voices on different parts of the page (Bent 2001; Scholes, Nagley, and Grier 2011; Apel 1953, 271).

Table 2-6 shows the note shapes used during the mensural notation period and how they evolved into the CWMN note shapes in use since the seventeenth century. As seen in the table, in the fifteenth century, black note heads became "void" or white instead of black (Kelly 2015, 204;

Boone 1999, 763). Mensural notation in the Renaissance using these hollow notes was called *white mensural notation*—in opposition to the *black mensural notation* used before.

Table 2-6: Note shapes used in mensural notation (from the thirteenth to the end of the sixteenth century) and in common Western music notation (since the seventeenth century). Table from Thomae (2017).



Note		**Centuries**			
Name	**Abbreviations**	**13th**	**14th**	**15th**	**17th**
Maxima	Mx				
Long	L				
Breve	B				
Semibreve	Sb				
Minim	M				
Semiminim	Sm				
Fusa	F				
Semifusa	Sf				

The four largest notes (maxima, long, breve, and semibreve) can have two possible values: a ternary value called *perfect*, and a binary value called *imperfect* (see Table 2-7). All the smaller notes have a binary (imperfect) value unless dotted (just as in CWMN). In addition to the note shape, two other factors should be considered in order to determine the actual duration of the maxima, long, breve, and semibreve. These factors are mensuration and context.

Mensuration defines the relation between a note and that of the next smaller degree (Apel 1953, 96). There are four types of mensuration, each defining the relation of the notes at different note levels: *modus major* (or *maximodus*) for the maxima-long relation, *modus minor* (or *modus*) for the long-breve relation, *tempus* for the breve-semibreve relation, and *prolatio* for the semibreve-minim relation (see Table 2-7). Since the fifteenth century, mensuration signs for the four combinations of (perfect and imperfect) tempus and (major and minor) prolation—already introduced by Vitry in the fourteenth century—were commonly used (see Table 2-8).

Table 2-7: Mensuration at the different note levels (Thomae 2017).

Mensuration	**Defined Relation**	**Mensuration Value**	**Relative Duration**
Modus Major	Maxima - Long	Perfect	Perfect maxima
		Imperfect	Imperfect maxima
Modus Minor	Long - Breve	Perfect	Perfect longa
		Imperfect	Imperfect longa
Tempus	Breve - Semibreve	Perfect	Perfect breve
		Imperfect	Imperfect breve
Prolation	Semibreve - Minim	Major	Perfect semibreve
		Minor	Imperfect semibreve

Table 2-8: The four combinations of tempus and prolation, their modern names, signs, and implications on the values of the breve (B) and semibreve (Sb). The terminology in the last column comes from Apel (1953, 98).

Sign	**Mensuration (modern name)**	**Value of the notes [B, Sb]**
𝇇	perfect tempus and major prolation	*[3,3]*
𝇈	perfect tempus and minor prolation	*[3,2]*
𝇉	imperfect tempus and major prolation	*[2,3]*
𝇊	imperfect tempus and minor prolation	*[2,2]*

Even though notes have a default perfect or imperfect duration given by the mensuration, in the case of triple meter (i.e., perfect mensuration), the *context* can modify the default durational value of the note. Context refers to the notes preceding or following a particular note. There are two types of modifications due to context:

- *Imperfection*, when a perfect note loses one-third of its value and becomes imperfect.
- *Alteration*, when a note's value is doubled.

The principles of imperfection and alteration, enumerated by Franco of Cologne's in his *Ars cantus mensurabilis* (ca. 1280), describe the different scenarios (contexts) in which these modifications must be applied.[63] According to Parrish (1978, 110), the principles are driven by the idea of applying these modifications (imperfection or alteration) to keep the structure of the music in triple meter (i.e., perfect mensuration), as shown in the examples of Figure 2-21.[64]

[63] For an English translation of portions of this treatise, consult Strunk and Treitler (1998, 2:229–32).
[64] The actual text in Parrish (1978, 110) reads: "the logic behind these principles is that the underlying movement of the music is by a series of perfections, each of which is three breves long." The "three breves long" part refers to the fact that Franco's principles of imperfection and alteration are written in terms of the long-breve relation.

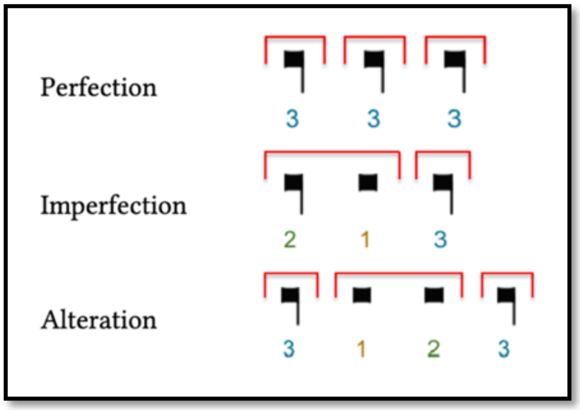


Figure 2-21: Examples of the different principles. The numbers below the notes indicate the duration of these in terms of the breve, and the red brackets mark the groups of notes that form a perfection. These brackets illustrate the use of imperfection and alteration to keep the structure of the music in triple meter. Image from Thomae, Cumming, and Fujinaga (2019).

The *dot of division* is used to indicate a different grouping of perfections. For this reason, it can only be used in perfect mensuration. The rearrangement of the perfect groupings, due to the presence of a dot of division, results in a different interpretation of the sequence of notes (see Figure 2-22).

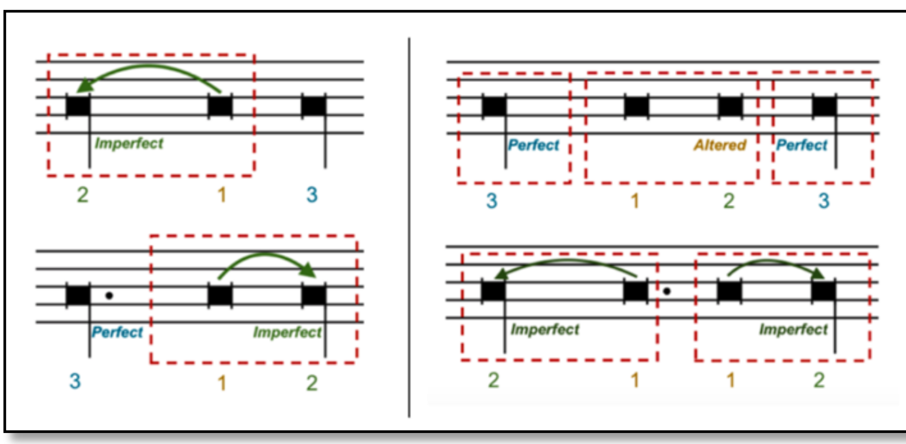


Figure 2-22: Change in the interpretation and perfect groupings (dotted red boxes) of the sequences in Figure 2-21 (middle and bottom) by the use of a dot of division. Image from Thomae, Cumming, and Fujinaga (2019).

In addition to contextual modifications to the note values, there are also non-contextual modifications that are implemented using other features. These are dots of augmentation and coloration. A *dot of augmentation* or *dot of addition* is a dot that increments an imperfect note by half of its value—just like a regular CWMN dotted note—resulting in a ternary duration. *Coloration* consists of writing the notes on a different ink color to temporarily change their relative durations. In the fourteenth century, when regular notes were black, the colored notes were written in red ink. Later, in the mid-fifteenth century, with regular use of hollow notes in white mensural notation, colored notes were filled-in in black. There are different types of coloration. Coloration types found in Renaissance music include:

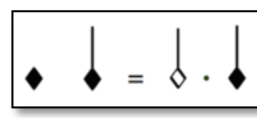
- *Hemiola coloration*. This type of coloration happens in perfect mensurations (Apel 1953, 130–42). In this coloration, three binary groups (the colored notes) take the place of two ternary ones (the uncolored notes). This is shown in Figure 2-23: the three binary groups are represented by either three imperfect breves (marked in orange) or a sequence consisting of an imperfect breve, two semibreves, and another imperfect breve (marked in green). Any of these groups of colored notes can take the place of the two perfect breves in tempus perfectum (marked in blue). Hemiola coloration results in a change in accent as shown in the bottom panel of Figure 2-23 (Apel 1953, 127).

Figure 2-23: Example of hemiola coloration in tempus perfectum. The top panel shows the example in mensural notation, while the bottom panel shows its interpretation in modern values (with the originally colored notes marked with brackets). Here, either three imperfect breves (orange) or the sequence of an imperfect breve, two semibreves, and another imperfect breve (green) can take the place of two regular perfect breves (blue); that is, three binary groups of colored notes may take the place of two regular (perfect) uncolored ones.

- *Triplet coloration*. Coloration can also be used in imperfect mensuration, where the three colored notes in the place of two regular, *imperfect* ones results in triplet rhythm as shown in the bottom panel of Figure 2-24 (Apel 1953, 127–8).

Figure 2-24: Example of triplet coloration in tempus imperfectum. While the top panel in mensural values looks similar to the top panel in Figure 2-23, the mensuration sign indicates that in this case the breves are imperfect by default (see blue in bottom panel). Therefore, the three in the place of two effect of the coloration results in completely new note durations (represented as triplets in modern notation, bottom panel).

- *Minor color*. This type of coloration happens in imperfect mensurations. Normally, it is applied to a group of a single semibreve followed by a minim in minor prolation (regardless of the tempus being perfect or imperfect); the interpreted duration of these colored notes is the same as a dotted minim followed by a semiminim. This is known as *minor color prolationis*. A much rarer case of minor color can be applied to a single breve followed by semibreve in imperfect tempus—interpreted as a dotted semibreve followed by a minim—known as *minor color temporis* (Apel 1953, 128–30).



By the late fifteenth century, there was a trend towards simplification of notation (Bent 2001). In the early sixteenth century, most music was in duple meter (imperfect mensuration), and the use of triple meter was less complex (with less alteration and an increase in the use of dots of perfection—dots of division following a perfect note—even in cases where previously they were not necessary) (Bent 2001; Chew and Rastall 2001; Kolb 2017; Tinctoris 1979; Busse Berger

1993; DeFord 2015). In the seventeenth century, features from instrumental music were introduced into vocal polyphony, such as the barline, slur, and beams (Chew and Rastall 2001). The seventeenth century also saw once again the adoption of the score format in vocal polyphonic music (Chew and Rastall 2001; Scholes, Nagley, and Grier 2011; Apel 1953, 271).

2.4.2 The Voice Alignment Issue of Mensural Notation and Related Tools

The separate-parts layout of mensural music obscures the visualization of vertical sonorities (the intervals between the voices) to our modern ears and eyes. It is not until the music is transcribed into a score, or musicians sing the parts together that the polyphonic texture of the piece can be appreciated (Owens 1997; Cumming 2013). In score layout, the reader has access to the same melodic information provided by the separate-parts layout, plus the information about vertical-intervals. However, the lining up of the mensural voices into a score requires the interpretation of the duration of the notes, which implies dealing with mensuration, note shapes, coloration, dots (of addition or division), and—in triple meter—dealing with the context-dependent nature of the notes' duration as well.

Tools that have dealt with the interpretation of mensural notation include Huang et al. (2015), Rizo et al. (2017), and Thomae, Cumming, and Fujinaga (2019). Huang et al. (2015) present a framework for OMR for mensural notation and transcription into modern values, returning a MIDI file encoding the music. However, the transcription work is restricted to imperfect mensurations, where the context-dependent issue of the notation is not present. On the other hand, Rizo et al. (2017) present a proposal for an interactive editing system to assist musicologists in the transcription of Hispanic mensural music from the sixteenth to the eighteenth centuries, together with some early prototypes of its different parts: the editor, the digital typographic font, and a mensural-to-modern finite state transducer.[65] The transducer is in charge of the interpretation of mensural notation. The transducer's weights are initialized so that the interpretation of the note's duration is consistent with the theory. However, as the user starts interacting with the music, the weights are adapted based on the user corrections. The idea behind the learning process is so that the mensural-to-modern notation transducer learns about common scribal errors (based on the user's previous corrections) and helps the user with suggestions for possible editorial corrections. Unfortunately, the paper does not report on the system's performance and the software does not seem to have been released.

Finally, Thomae, Cumming, and Fujinaga (2019) introduced a tool for automatic voice alignment (or scoring up) of music in mensural notation, called the *mensural scoring-up tool*.[66] This tool is a Python script that implements the principles of imperfection and alteration, dealing with the context-dependent nature of mensural music in triple meter. It also deals with the non-contextual cues of dots of addition—including a heuristic method to distinguish between dots of division and dots of addition—and hemiola coloration. The tool runs on the command line. It receives as input a set of Mensural MEI files (i.e., MEI files containing mensural notation values) encoding each of the parts and outputs a single Mensural MEI file encoding the same information plus the interpretation of the notes as perfect / imperfect / altered. If this Mensural MEI output file is given to Verovio (a rendering library for MEI),[67] the piece is shown as a score with all the voices lined up.

After testing the *automatic mensural scoring-up tool* on a small corpus of fourteenth- and fifteenth-century pieces, Thomae, Cumming, and Fujinaga (2019) found out that the main source

[65] A finite state transducer is a mathematical model that converts a string of characters into another.
[66] https://github.com/ELVIS-Project/scoring-up
[67] https://www.verovio.org/index.xhtml

of error for the automatic voice alignment was related to scribal errors, mostly the lack or misplacement of dots of division. To make all the voices line up properly, the user of the tool would need to correct these scribal errors. MEI allows for this. In addition to encoding mensural notation values, MEI allows the user to account for editorial corrections by encoding both the original and corrected readings. Therefore, to obtain a score with the voices correctly lined up, the user would need an interface to make the editorial corrections and run the scoring-up script once again in the corrected voices. The Measuring Polyphony Editor (MP Editor) is a software that allows for this, with automatic voice alignment and editorial correction functionalities (Desmond, Hankinson, et al. 2020).[68] The MP Editor is an online mensural notation editor developed under the supervision of Professor Karen Desmond as part of the Measuring Polyphony Project, with Juliette Regimbal as lead developer, myself as developer of the automatic voice alignment functionality, and great support from Laurent Pugin, Craig Sapp, and Andrew Hankinson for the related technologies of Verovio, Humdrum, and IIIF, respectively.[69] The Editor allows to enter pieces in mensural notation and encode them as scores (with editorial corrections, if relevant) in MEI format.

The MP Editor has two music-related interfaces: the input editor and the score editor.[70] The *input editor* is where the user enters the music by selecting the systems on the manuscript image and typing in the notes of the system. Each system is stored internally as **mens, the Humdrum format for mensural notation (Rizo, León, and Sapp 2018). The *score editor* shows the piece as a score (see Figure 2-25). This is done by automatically scoring up the piece—using a re-

[68] https://editor.measuringpolyphony.org/#/
[69] https://measuringpolyphony.org
[70] Videos on how to use the MP Editor, the specific features of the web app, and an example of a step-by-step transcription process can be found in its YouTube channel: https://www.youtube.com/channel/UCWB5WRs4LjKpYB_-J2Sb03Q/playlists.

implementation of my Python script in JavaScript—and rendering the resulting MEI file through Verovio. The user can also make editorial corrections in the score editor.

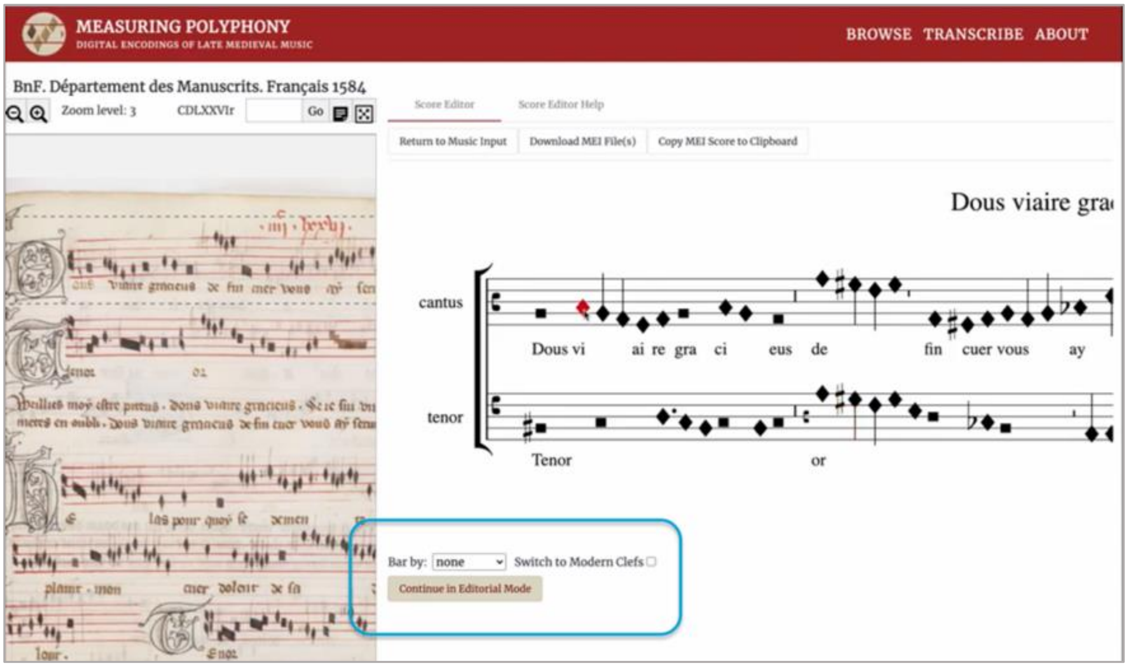


Figure 2-25: The score editor interface of the MP Editor. It renders the piece as a score and allows for barring by different note values (long, breve, and semibreve), using modern clefs, and encoding editorial corrections (see blue box).

In the MP score editor interface, the user can see the melodic and vertical-interval information of the piece by looking at the rendered score. Having access to this information allows the user to detect counterpoint errors more easily—these are errors in the vertical intervals or the ways in which they are approached according to the common practice of the time.[71] This is additionally facilitated by the "barring" and "change to modern clefs" options in the interface (see blue box in Figure 2-25). Counterpoint errors hint at errors in the score, either in the notes entered by the user, the interpretation of the notes' duration by the scoring-up algorithm, or an error in the manuscript (i.e., a scribal error). While the score layout facilitates the detection of these errors, the

[71] For more details about specific to Renaissance-style counterpoint, consult Schubert (2008).

use of computer-aided musicological tools that can retrieve and analyze information in symbolic representations of music (e.g., MEI) might further facilitate this task. This is the hypothesis I will test in this dissertation. The following section contains a survey of these computational music analysis tools.

2.4.3 Computational Music Analysis Tools

This section reviews computer-aided musicology tools. These tools can process music encoded in symbolic music files, which are machine-readable files encoding music documents (not audio). Some well-known symbolic formats are MusicXML,[72] an XML-based representation for CWMN that serves as an interchange format between different music notation, music analysis, music information retrieval, and musical performance applications (Good 2001); and MIDI,[73] widely used for three decades now and with abundant free MIDI scores online (Müller 2015, 13–16).[74] In academia, MEI and Humdrum are among the popular symbolic formats. As mentioned at the end of Section 2.3.1, MEI is an XML-based format that supports encoding early music notations (in addition to CWMN), musical editions, rich metadata, linking encoded music to facsimiles and recordings, and other features that satisfy its multi-disciplinary community (Roland 2002; Roland, Hankinson, and Pugin 2014; Hankinson, Roland, and Fujinaga 2011).[75] On the other hand, Humdrum is a text-based data format that has been around since the 1980s. It can be used to encode music in CWMN using the ***kern* format (Huron 1997), and mensural notation using ***mens* format (Rizo, León, and Sapp 2018).[76]

---

72 https://www.musicxml.com
73 https://www.midi.org
74 The original purpose of MIDI was not to work as a symbolic file. MIDI, which stands for Musical Instrument Digital Interface, was originally developed as an industry standard to allow digital electronic music instruments from different manufacturers to communicate and play together. However, it is widely used for sheet music representation, as proven by the abundant MIDI scores found online.
75 https://music-encoding.org
76 https://www.humdrum.org

Symbolic files that encode music notation in a document offer advantages over mere digital images. Since these files are machine-readable, the computer can extract a wide variety of information from them useful for music searching and analysis. There are many tools developed for empirical symbolic music research, including jSymbolic, music21, VIS Framework, and Humdrum. This section presents each of these tools, emphasizing their strong points and the symbolic formats they can process. A table comparing all four tools is presented at the end of this section.

2.4.3.1 ***jSymbolic***

jSymbolic is an open-source Java framework for feature extraction in (potentially a large collection of) symbolic music files (McKay, Cumming, and Fujinaga 2018).[77] It was created by Cory McKay at McGill University. A "feature" in jSymbolic is a piece of information that measures a single characteristic of a musical work; for example, the total range in semitones from the highest to the lowest note in all voices of a piece of music (McKay and Cuenca 2021, 17–18). The features extracted by jSymbolic relate to a wide range of musical characteristics, including pitch statistics, melodic intervals, vertical intervals, rhythm, instrumentation, texture, and dynamics.[78]

In addition to search-based tasks, the extracted features can be analyzed statistically or used as inputs for machine learning algorithms, leading to musicological findings (McKay, Cumming, and Fujinaga 2018). jSymbolic, along with machine learning, has been used for studying composer attribution (McKay, Cumming, and Fujinaga 2017; McKay et al. 2017); exploring the stylistic origins of genres (Cumming and McKay 2021); and distinguishing between regional styles, such as Iberian and Franco-Flemish (McKay 2018).

[77] http://jmir.sourceforge.net/jSymbolic.html
[78] http://jmir.sourceforge.net/manuals/jSymbolic_manual/home.html

jSymbolic can process MEI and MIDI files. It can be used via a graphical user interface (GUI), or through the command line to facilitate batch processing, through its (well-documented) Java API to facilitate programmatic use, or within the Rodan workflow manager presented in Section 2.3.3.2.2 (McKay, Cumming, and Fujinaga 2018).

2.4.3.2 *Music21*

Music21 is an open-source, cross-platform Python-based toolkit for analyzing, searching, and transforming music in symbolic forms (Cuthbert and Ariza 2010).[79] It was created by Michael Cuthbert at MIT and has been in use since 2008 (Cuthbert and cuthbertLab 2021). Music21 has a well-developed framework for importing scores from the most common symbolic music formats, such as MusicXML, **kern, MIDI, ABC, MuseData, and CMN MEI (Cuthbert, Ariza, and Friedland 2011; Music Encoding Initiative 2021).

With musci21, scholars and musicians can write simple and reusable scripts (Cuthbert and Ariza 2010). Many examples of the use of music21 can be found on its website.[80] These examples illustrate its use for answering musicological questions, analyzing large music datasets, plotting results, generating music examples, retrieving and processing music, etc. Music21 has been used to build other music-analysis or music-query applications, including the VIS Framework—introduced later in this section—and PatternFinder—a tool for searching polyphonic music (Garfinkle et al. 2017).

---

[79] A tutorial for music21 can be found at https://web.mit.edu/music21/doc/index.html. The GitHub repository can be found at https://github.com/cuthbertLab/music21.
[80] https://web.mit.edu/music21/

Music21 provides feature-extraction functionality by including some of the jSymbolic features and some of its own.[81] Its feature-extraction functionality, along with data mining toolkits, was used by Cuthbert, Ariza, and Friedland (2011) to distinguish between Monteverdi's and Bach's works and between German and Chinese folk music.

2.4.3.3 ***VIS Framework***

The Vertical Interval Successions (VIS) Framework for Music Analysis[82] is a modular Python library built on music21 and pandas.[83] It was developed by Christopher Antila at McGill University. Its primary function is to find contrapuntal patterns in symbolic music (Antila and Cumming 2014). These patterns are expressed in the forms of n-grams, defined in VIS as a sequence of *n* consecutive vertical intervals and the melodic motions—with respect to a single voice—that connect them (see Figure 2-26). The VIS Framework has been used to analyze different contrapuntal aspects. Antila and Cumming (2014) studied the stylistic change between the Ockeghem, Josquin, and Palestrina generations in the Renaissance by finding the most common 3-grams in each period and examining how their use changed across generations. Schubert and Cumming (2015) studied contrapuntal repetition in Lassus duos by making use of 3-grams as well. And Morgan (2017) studied the discrepancies between theory and practice by examining the "acceptable" 2-grams in Tinctoris' counterpoint treatise (*Liber de arte contrapunti*) against contemporaneous repertoire.

[81] For the music21 features implemented based on those found in jSymbolic, refer to https://web.mit.edu/music21/doc/moduleReference/moduleFeaturesJSymbolic.html. For the original music21 feature extractors, refer to https://web.mit.edu/music21/doc/moduleReference/moduleFeaturesNative.html.
[82] https://vis-framework.readthedocs.io/en/v3.0.5/
[83] Refer to https://web.mit.edu/music21/ and https://pandas.pydata.org/. Pandas is a Python library for data analysis. VIS uses Pandas's DataFrames to store and present data in the familiar table format of rows and columns.

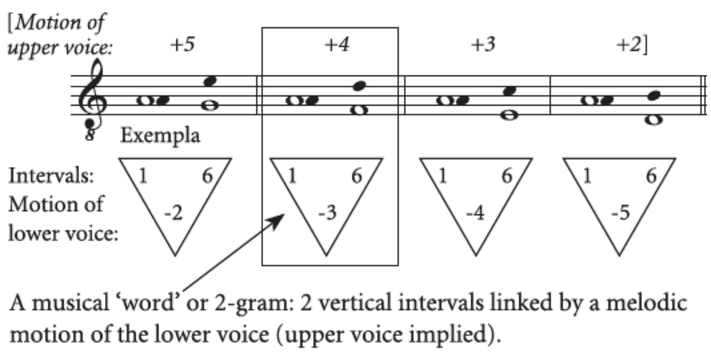


Figure 2-26: Example of 2-grams from Schubert and Cumming (2015). The 2-gram consists of two vertical intervals (the unison and the 6th in the selected part of the image) and the melodic interval of the lowest voice (downward 3rd).

VIS allows for a much broader set of queries. It includes *analyzers*, which are used to make music analysis decisions (Antila 2017). As indicated in the VIS documentation, there are two types of analyzers: indexers and experimenters. *Indexers* produce results that can be attached to a specific moment of a piece; in other words, they "index" the piece according to a certain type of event (e.g., the *NoteRestIndexer* index each of the notes and rests of the piece).[84] On the other hand, *experimenters* produce results that characterize the piece as a whole. The *FrequencyExperimenter* was used in the previous studies to determine the frequency of the n-grams—found with the *NGramIndexer*—used in a given corpus. In addition to the n-gram indexer, another indexer related to counterpoint is the dissonance indexer introduced by Alex Morgan, who later developed the Renaissance dissonance filter for the humlib library (also presented in this section).[85]

[84] For a complete list of all the indexers in the VIS-Framework (for the stable version, 3.0.5), consult https://vis-framework.readthedocs.io/en/v3.0.5/vis.analyzers.indexers.html#module-vis.analyzers.indexers.
[85] The VIS dissonance indexer can be found at https://vis-framework.readthedocs.io/en/v3.0.5/vis.analyzers.indexers.html#dissonance-module. The humlib dissonance filter, also developed by Morgan and Sapp, can be found at https://github.com/craigsapp/humlib/blob/master/src/tool-dissonant.cpp. For the VHV/humlib documentation of this filter, consult https://doc.verovio.humdrum.org/filter/dissonant/.

VIS can receive as input formats files that can be converted into music21 format, which includes MusicXML MEI, Humdrum, and MIDI.[86] The VIS Framework can be used through its Python API, the *Counterpoint Web App*, and the Rodan workflow manager presented in Section 2.3.3.2.2.[87] The Rodan wrapper for the VIS Framework, which allowed for using VIS within the Rodan workflow manager, was developed by Ryan Bannon at McGill University. It was used by Samuel Howes (also from McGill) for various music theory analyses including a study of medieval motets (Desmond, Hopkins, et al. 2020).

2.4.3.4 ***Humdrum***

Humdrum is a system for computational musicology research created by David Huron in the 1980s and maintained by Craig Sapp at Stanford's Center for Computer Assisted Research in the Humanities (CCARH).[88] As indicated on its website (Sapp 2021), Humdrum consists of the humdrum syntax (symbolic encoding) and the software for analyzing the encoded material. This software exists in two forms: (1) its original form, David Huron's *Humdrum Toolkit;* and (2) the updated and extended versions of it, such as the *Humdrum Extras*, the *humlib* library, and the *Verovio Humdrum Viewer*, all developed by Sapp. The *Humdrum Toolkit* is a set of Unix command-line tools (mostly written in Bash and AWK) to parse and analyze humdrum data. A guide to using the Humdrum Toolkit was provided by Huron (1999) and can now be accessed online through the Humdrum page.[89] The *Humdrum Extras* are a set of command-line tools (written in C++) that expand the capabilities of the humdrum toolkit.[90] *Humlib* is a C++ parsing

---

[86] VIS imports files through its "Importer" method, which converts the file into music21 format (https://vis-framework.readthedocs.io/en/v3.0.5/vis.models.html#vis.models.indexed_piece.Importer).
[87] For the counterpoint web app, refer to https://counterpoint.elvisproject.ca/. For the wrappers of the VIS Framework to work in Rodan, refer to https://github.com/ELVIS-Project/vis_rodan.
[88] https://www.humdrum.org/
[89] https://www.humdrum.org/guide/index.html
[90] http://extras.humdrum.org/

library for Humdrum data files.[91] This library is compiled into JavaScript for its use in the *Verovio Humdrum Viewer* (VHV), an online Humdrum file editor with integrated graphical notation display through Verovio (the rendering library designed for MEI).[92]

In addition to Humdrum data, the user can load MusicXML and MEI files into the VHV editor; these files are converted into Humdrum data to allow for the use of Humdrum processing tools, represented by filters (Sapp 2017a). According to the VHV documentation (Sapp 2017b), *filters* are embedded commands within the Humdrum data that modify the data. These processing tools come from the *humlib* library. Most available filters are related to data processing (e.g., filters to *extract* a part or parts from the encoded score or to *transpose* the score). However, more advanced filters are available too, such as the *dissonant* filter that analyzes the functions of dissonant notes in Renaissance-era music.[93] The dissonant filter, developed by Alex Morgan, is one of the analysis tools used in the Josquin Research Project (JRP) that make it possible to engage with Renaissance music.[94]

As noted by Condit-Schultz and Arthur (2019), though there is no framework that dominates the field of digital musicology, Humdrum has been cited as the direct inspiration for its biggest competitors, music21—as music-analysis software—and MusicXML—as a symbolic format (Cuthbert and Ariza 2010; Good 2001). Table 2-9 compares the four computational musicology tools presented in this section.

---

[91] https://humlib.humdrum.org/

[92] The Verovio Humdrum Viewer (VHV) can be found at https://verovio.humdrum.org/. Refer to https://doc.verovio.humdrum.org/ for documentation.

[93] The complete list of filters can be found at https://doc.verovio.humdrum.org/filter/index.html.

[94] https://josquin.stanford.edu/

Table 2-9: Comparison of the four computational music analysis tools presented before. The 'accepted input formats' column only compares MusicXML, MIDI, MEI, and Humdrum, but some of these tools (e.g., Music21) can process other formats as well.

<table>
<tr><th colspan="2" rowspan="2">Tools</th><th rowspan="2">Creator / Lead Developer</th><th rowspan="2">Programming Language</th><th rowspan="2">Main Use</th><th colspan="4">Accepted Input Formats</th></tr>
<tr><th>MusixXML</th><th>MIDI</th><th>(CMN) MEI</th><th>Humdrum (**kern)</th></tr>
<tr><td colspan="2">Music21</td><td>Michael Cuthbert, Christopher Ariza, Benjamin Hogue, and Josiah Wolf Oberholtzer</td><td>Python</td><td>music writing, processing, searching, and analyzing (some feature extraction)</td><td>yes</td><td>yes</td><td>yes</td><td>yes</td></tr>
<tr><td colspan="2">VIS Framework</td><td>Christopher Antila, Reiner Krämer, and Alex Morgan. Rodan wrappers by Ryan Bannon.</td><td>Python</td><td>counterpoint and indexed data</td><td>yes</td><td>yes</td><td>yes</td><td>yes</td></tr>
<tr><td rowspan="3">Humdrum</td><td>Humdrum Toolkit</td><td>David Huron, in the 1980s</td><td>Bash and AWK.
It is program-language agnostic since it is a set of command-line tools.</td><td rowspan="3">music writing, processing, searching, and analyzing</td><td>no</td><td>no</td><td>no</td><td>yes</td></tr>
<tr><td>humlib</td><td rowspan="2">Craig Sapp</td><td>C++ (also compiled to command-line tools)</td><td rowspan="2">if converted to Humdrum first</td><td>no</td><td rowspan="2">if converted to Humdrum first
Mensural MEI too</td><td rowspan="2">**mens too</td></tr>
<tr><td>VHV humlib</td><td>JavaScript (compiled humlib)</td><td>no</td></tr>
<tr><td colspan="2">jSymbolic</td><td>Cory McKay</td><td>Java</td><td>feature extraction</td><td>no</td><td>yes</td><td>yes</td><td>no</td></tr>
</table>

Chapter 3 Methodology – Part 1: The Digitization and Encoding Workflow's Tools

I used the following three-stage methodology/workflow to retrieve digital images and symbolic scores (i.e., scores encoded in a symbolic format) for each of the pieces of GuatC 1 (see Figure 3-1). This method is designed to retrieve these symbolic scores in a semi-automatic way through OMR-assistance and automatic voice alignment. The user is involved in the correction of the output of these two stages: correcting the results of the OMR (the recognized symbols and its pitches) at the end of the second stage; and correcting the results of the voice alignment (which normally implies the correction of scribal errors in the form of editorial corrections) at the end of the third stage.

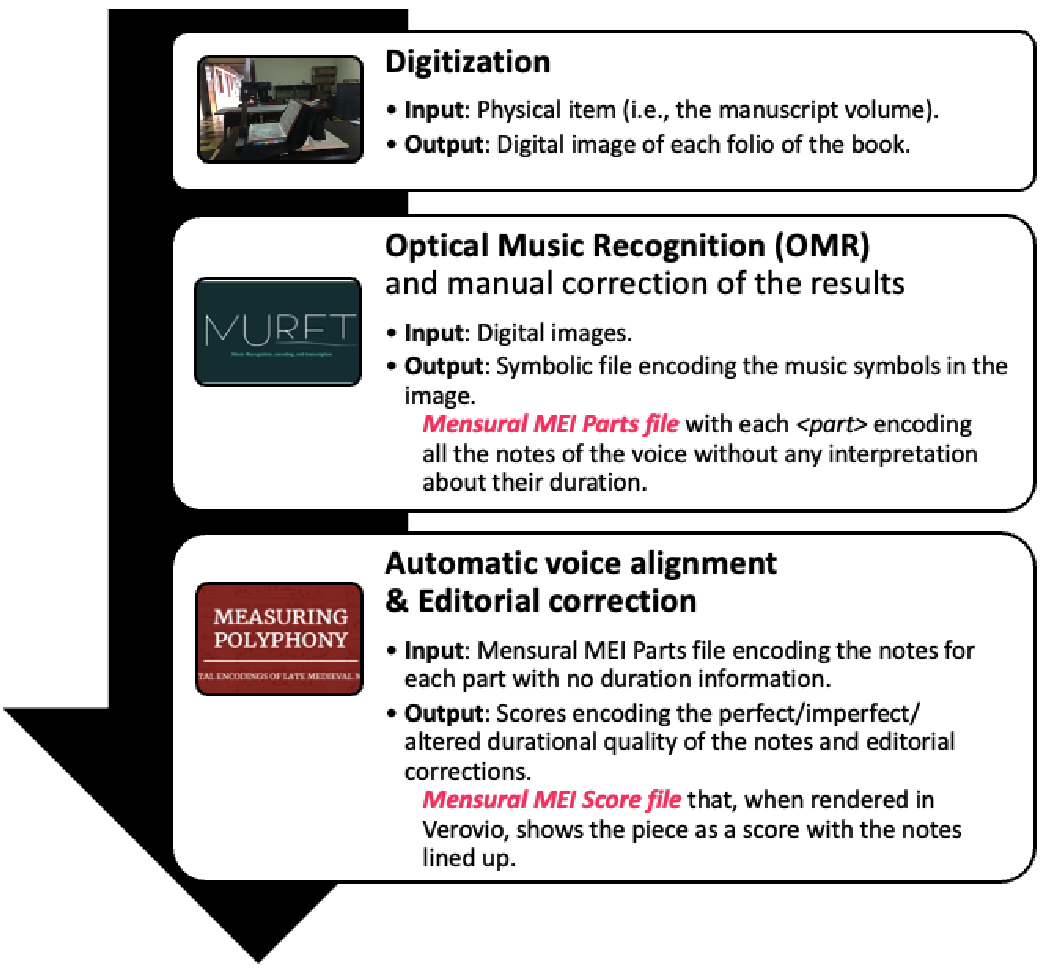


Figure 3-1: Music document digitization and encoding workflow followed for GuatC 1.

The first stage in the workflow, the digitization stage, consists of photographing the folios of the manuscript. The images obtained are then passed through an OMR process to retrieve the music symbols in the image and encode the music of a composition in a symbolic file. However, the retrieved file encoding the mensural symbols and their pitches does not solve the vertical alignment issue presented in Section 2.4.2. To have a file encoding a score with the voices properly lined up, one needs an additional stage. This third stage in the workflow performs the automatic alignment of the voices and allows for entering editorial corrections.

In addition to the stages of the encoding workflow, Figure 3-1 shows the tools used to perform each of these stages. Each of these tools are presented in Sections 3.1 to 3.3. For the digitization stage, after a survey of digitization equipment and services in Guatemala, I chose to design a do-it-yourself (DIY) book scanner. Section 3.1 presents the details of this survey, the reason behind the decision to use a DIY book scanner, and the setup of the latter; it also presents the steps prior to digitization such as requesting permission for the project to the holding institution (the Archivo Histórico Arquidiocesano de Guatemala, AHAG) and looking for digitization and conservation staff. For the OMR stage, I used the Music Recognition Encoding and Transcription (MuRET) framework, which allows to perform each of the individual steps involved in the OMR process (presented in Section 2.3.1) and correct the results of each of these steps. Section 3.2 presents the reasons behind this decision and how to use MuRET to perform OMR over a collection of images. Finally, for the automatic voice alignment and editorial correction stage, I used the Measuring Polyphony Editor (MP Editor). Section 3.3 presents how to use this online mensural-notation editor to enter a piece and obtain a correct score encoded in MEI.

While Sections 3.2 and 3.3 focus on the current state of MuRET and the MP Editor, Sections 3.4 and 3.5 present the modifications to these tools as part of this dissertation. Section 3.4

contains the work done to allow for the interoperability between MuRET and the MP Editor, allowing to retrieve symbolic scores through OMR assistance. The MP Editor was originally designed for manual input; in other words, the user enters the piece by typing its notes on a computer keyboard (and then the MP Editor performs the automatic voice alignment of the entered notes and allows the user to provide editorial corrections). In this dissertation, with the help from the developers of both tools, I am allowing the MP Editor to use the output of MuRET as input (see the last two stages in Figure 3-1); this avoids the manual labor of typing in all the notes of a piece and instead gather these symbols through OMR. Finally, Section 3.5 presents the integration of humlib's dissonance filter into the MP Editor, with the help of humlib's developer. This allows to study one of the key aspects of this dissertation: determine whether the use of a counterpoint error detector makes the correction of the scored-up OMR file in the MP Editor more efficient.

3.1 Digitization Details (First Visit)

This section focuses on the first stage of the workflow, the digitization. This stage includes the tasks of requesting permission to the archive for the execution of the project (Section 3.1.1), selecting one choirbook of the collection for a pilot project, looking for digitization equipment options (Section 3.1.2), and finding the personnel to conduct the digitization and conservation assessment and treatment of the selected book. All these actions were conducted prior, during, or based on my first visit to Guatemala in November 2018. For the details of what happened during the digitization process itself, please consult Section 5.1 on the Experiment chapter that provides information on the second visit to the AHAG in January 2019, when the digitization process took place.

3.1.1 Permission Request, Conservation Report, and Choirbook Selection

I contacted Guatemalan musicologists Omar Morales Abril on 19 July 2018 asking for his advice on how to approach the archive about this project. Morales Abril has conducted projects that involve the study of music sources in the AHAG, the archive holding the GuatC choirbook collection. He referred me to the AHAG's archivist Alejandro Conde. I called the AHAG on 18 August to speak with Conde, who already knew details about the project thanks to Morales Abril. He had two suggestions: (1) that I should go to the AHAG in person to get to know the archive and its collections and (2) he proposed to ask permission to the AHAG authorities to make a pilot project with only one of the choirbooks and, depending on the results of the pilot, in the future I could approach them asking permission to work with the full collection. Following Conde's advice, I wrote a letter to the chancellor of the Ecclesiastical Curia of Santiago de Guatemala and director of the AHAG, Father Eddy René Calvillo, on 23 October 2018. In this letter I asked for permission to conduct a pilot project with one of the GuatC choirbooks. The request was approved eight days later, on 31 October. The letter of permission with the chancellor's signature can be found on Appendix B.

With the approval of the chancellor of the Curia and under Conde's advice, I planned my trip to Guatemala to consult the GuatC collection in person (for the first time in my life). Before this visit, I recompiled all the information I could about the GuatC repertoire. Of course, there was minimal information about choirbooks 5 and 6 since they were discovered relatively recently, and the DjVu images of the microfilm of the folios of GuatC 1–3 also raised questions about missing or cropped folios. I traveled to Guatemala on 03 November and visited the archive numerous times during the week of November 5–9. After this visit, I realized that most music in the collection was in duple meter and got a general idea of the level of degradation of the different books.

In addition to visiting the archive, I used my November 03–11 trip to Guatemala to look at different digitization and conservation options. Prior to the visit, and to reduce the options to look at in one week, I called and sent emails to various digitization services in Guatemala and looked among my contacts for information about paper conservation options in the country. Section 3.1.2 presents the result of this survey for digitization options. Regarding the options for conservation, I only found one institution that performed paper conservation: the *Centro de Rescate, Estudios y Análisis Científico del Arte* (*CREA*).[95] CREA is a non-profit organization focused on the conservation and restoration of Guatemalan religious heritage.

I visited CREA's installations on 6 November and presented the project to the lead conservator Lucy González Muñoz. The next day, 7 November, three conservators from CREA visited the archive and evaluated the conservation state of the six GuatC manuscripts. On 26 November, I received CREA's full conservation report about the GuatC collection and, based on this report and on the amount of triple meter in the music, I chose to digitize choirbook GuatC 1. CREA's conservation report can be found on Appendix B.

3.1.2 **Survey for Digitization Equipment and Services**

Having obtained the permission to digitize and encode the GuatC 1 book, I started looking for suitable digitization options. Since the AHAG does not have a digitization department, the digitization task needed to be outsourced. As indicated in Section 2.2.2.1, given the nature of the choirbook as a large manuscript volume of a special collection, it should be digitized from the top with a book scanner, avoiding the use of platens to flatten the pages. I surveyed on available book scanners and digitization services in Guatemala City, making calls and sending emails to the National Library, the National Bank of Guatemala, University Libraries, and recommended

[95] https://fundacionrozasbotran.org/cultura-y-arte/crea/

digitization services by the former institutions. After these calls, I made a list of places to consult during my visit to Guatemala in November 2018. The National Library has not carried out any digitization work to preserve its collections because they do not have the equipment or the budget for such an enterprise, as I was told when I called them. And, although the National Bank of Guatemala has digital records of many important documents, they do not have a large book scanner on site; however, they provided suggestions on various digitization services.

Given the circumstances outlined above, I focused my survey on University Libraries, archives, and digitization services. Table 3-1 shows the summary of the digitization options found and the final decision regarding how suitable these were for the GuatC books considering their characteristics: large dimensions, bound items, and part of a special collection.

Table 3-1: Book scanner survey.

No.	Name of Institution	Type of Institution and Description	Digitization Technology	Comments
1	*Archivo Histórico Arquidiocesano de Guatemala* (*AHAG*), Guatemala City, Guatemala. Visit: 5–9 November 2018.	The archive holding the GuatC collection.	Portable planetary scanner for A3 landscape documents (17 x 11.8 in^2).  *Fujitsu ScanSnap SV600* https://www.fujitsu.com/global/products/computing/peripheral/scanners/soho/sv600/	The scanning head does not cover the area of a single folio from GuatC 1 (18 x 12 in^2).
2	*Centro de Investigaciones Regionales de Mesoamérica* (*CIRMA*), La Antigua Guatemala, Sacatepéquez, Guatemala. http://cirma.org.gt/ Visit: 6 November 2018.	The archive holding the microfilms of some of the music manuscripts in the AHAG, including GuatC 1–3.	Portable planetary scanner for A3 landscape documents. Same scanner as in AHAG.	The scanning head does not cover the area of a single folio from GuatC 1 (18 x 12 in^2).
3	*Centro de Escaneo, Biblioteca Ludwig von Mises, Universidad Francisco Marroquín* (*UFM*), Guatemala City. https://biblioteca.ufm.edu/rescate-digital Visit: 9 November 2018.	University library with a digitization center. The library joined the Internet Archive project in 2008, contributing to it through the digitization of books related (mostly) to the History of Guatemala and Central America.	Book scanner with V-shaped cradle, glass platen, two cameras, and black cover to filter outside lighting. Cameras CANON EOS 5D MARK III (maximum resolution of 5760 x 3840 pixels). They work in collaboration with Internet Archive, which implies sharing all images taken on the web. (https://archive.org/details/sigloxx)	Using this scanner would have implied transportation of the manuscript outside the archive. The scanner is not big enough and it has a non-removable platen. It only takes JPEG images (which imply compression), and they are automatically uploaded to the web upon capture. Need to get an agreement with both parts (archive and university).

4	*Infile*	Company digitizing books for the *Biblioteca Central de la Universidad San Carlos de Guatemala* (*USAC*), a university library.	Portable planetary scanner for A3 landscape documents. Same scanner as in AHAG.	The scanning head does not cover the area of a single folio from GuatC 1 (18 x 12 in^2).
5	*RICOH*, Guatemala City.	Electronics company that sells and rents scanners. This company was recommended by Amanda Judith López, head of the *Archivo General of the Universidad San Carlos de Guatemala* (*AGUSAC*), a university archive. AGUSAC was to buy a book scanner from this company in March/April 2019.	-	According to the company itself, not an option (too expensive for a small project).
6	*Productive Business Solutions* (*PBS*), Guatemala City. https://www.grouppbs.com/contact-us/ servicioalcliente.gt@grouppbs.com	Technology company providing products and services to different countries. PBS has an office in Guatemala. This company was recommended by the administrative Assistant Secretary at the National Bank of Guatemala. http://www.banguat.gob.gt/	Portable book scanner for A3 (landscape) documents. *CZUR ET16 Plus book* https://www.czur.com/product/et16/	Does not cover the area of a single folio from GuatC 1 (18 x 12 in^2).
7	*DataScan*, Guatemala City. www.datascan-site.com	Scanning-service company, also recommended by the administrative Assistant Secretary at the National Bank of Guatemala. http://www.banguat.gob.gt/	*ATIZ BookDrive Pro* Two pages A2 (16.5 x 23.4 in^2) or landscape A1 (23.4 x 33.1 in^2). https://www.atiz.com/atiz_connect/bookdrive_pro/brochures/bookdrive-pro.pdf	Perfect dimensions for digitizing the GuatC 1 choirbook. However, it was already borrowed by another institution and the service providers never confirmed if it was going to be available for the digitization period (January 2019). In addition, while the cameras are

				interchangeable, the cameras included with the scanner have a resolution of 5,184 x 3,456 pixels, which would generate images below 300 ppi given the dimensions of the GuatC 1 (18 x 12 in^2).
8	*Instituto de Nutrición de Centro América y Panamá* (*INCAP*), Guatemala City. Visit: 9 November 2018.	The INCAP research institute has a digitization service. http://www.incap.int/biblio/index.php/es/servicios-de-digitalizacion-2	Scanner COPIBOOK, designed by i2S, to digitize bounded items of size up to A2 (16.5 x 23.4 in^2).  https://www.i2s.fr/en/heritage-digitization/small-book-scanners-a3-a2/copibook-open-system-a2 Installation at INCAP, with additional lights added at each side: 	Unsatisfactory results when the personnel did trials, the images did not look well defined. In addition, the platen put a lot of pressure on one of the sides of the book, not ideal for digitizing old material.

The scanning area of all options consulted was not big enough to digitize the choirbooks except for the ATIZ BookDrive Pro (entry 7 in Table 3-1). Even though the ATIZ BookDrive Pro was big enough to fit the manuscript in the cradle and the cameras could be substituted by others of higher resolution, the option was abandoned due to the use of a platen and the lack of confirmation regarding its availability for the required period. McGill University Library's digitization and new media administrator, Greg Houston, was consulted about the use of this scanner for the project. Although the McGill Library uses this scanner to digitize many of its books, Houston does not recommend it when dealing with old documents.[96] For these documents, the McGill library personnel uses a similar scanner that does not include the glass platen for flattening the pages.

I also investigated the option of borrowing cameras to either use them in the place of the current lower-resolution cameras of the book scanners in Table 3-1 or to use them with a DIY book scanner. The camera options consulted are shown in Table 3-2.

[96] Tour to Rare Books and Special Collections & Digitization Lab of the McGill University Library on 16 November 2018 as part of the "Preservation Management" course (GLIS 642).

Table 3-2: Survey of Digitization Cameras.

Institution	**Description / Relevance**	**Camera Specifications**	**Status**
Centro de Investigaciones Regionales de Mesoamérica (*CIRMA*), La Antigua Guatemala, Sacatepéquez, Guatemala. http://cirma.org.gt/	The archive holding the microfilms of some of the music manuscripts in the AHAG, including GuatC 1–3.	*Canon EOS 5DSR* Resolution: 8688 x 5792 (50.6 effective megapixels). Resulting in an approximate value of 480 ppi for the GuatC 1 folios (18 x 12 in^2).	Not available for borrowing.
Centre for Interdisciplinary Research in Music Media & Technology (CIRMMT), Montreal, Quebec, Canada. https://www.cirmmt.org/	I am affiliated to this research institution.	*Nikon D3200* Resolution: 24.2 Megapixels.	Available, but usually reserved for campus use.
McGill University Photography Students' Society (MUPSS). http://www.mupss.ca/	Student association at my university.	All cameras resulted in a resolution lower than 300 ppi given the size of the GuatC1 folios.	Nothing suitable.
Marvin Duchow Music Library, McGill University, Montreal, Quebec, Canada. https://www.mcgill.ca/library/branches/music/av	Library at my university.	*Nikon D750* (2 cameras) Resolution: 24.3 Megapixels.	Booking system for students.

Although the camera from Centro de Investigaciones Regionales de Mesoamérica (CIRMA) has a resolution that even exceeds the one given by the Federal Agencies Digital Guidelines Initiative (FADGI) for a four-star project (400 ppi, see Table 2-5), the camera was not available for borrowing. Only the cameras from CIRMMT and the Marvin Duchow Music Library were available and of a resolution high enough for a three-star FADGI project (300 ppi as indicated in Table 2-5).

I also consulted the last three institutions in Table 3-2 for lighting options. CIRMMT was the only one that had light equipment; unfortunately, these lights are fragile and would not endure travel. The Distributed Digital Music Archives & Libraries (DDMAL) Lab at McGill University,[97] to which I belong to as a research assistant, bought a pair of LED natural-light panels for the execution of this project—the lights have been donated to the Marvin Duchow Library so that students can borrow them for future projects. The lights were approved by an expert, BAnQ's photographer Michel Legendre, before their purchase.

With no viable option for book scanner and with options for borrowed cameras and lights, I decided on making a do-it-yourself book scanner, the details of which are presented in Section 3.1.3. I needed additional equipment, which the DDMAL Lab also lent me. Table 3-3 shows the equipment bought by the DDMAL Lab for this project, including the lights.

[97] The Distributed Digital Music Archives and Libraries (DDMAL) Lab is directed by Professor Ichiro Fujinaga. The DDMAL Lab focuses on the development of technology to facilitate the construction of worldwide distributed digital musical archives' and libraries' resources. https://ddmal.music.mcgill.ca

Table 3-3: Equipment bought by the DDMAL Lab for the execution of the digitization phase of the project.

Item	**Quantity**	**Store**	**Description**	**Price (CAD)**	**Total including shipping (CAD)**
External hard drives (1TB of storage)	2	amazon.ca	Western Digital 1TB My Passport Portable External Hard Drive-USB 3.0-WDBYNN0010BBL-WESN, Blue	$64.99	$276.30
		amazon.ca	Seagate Backup Plus Slim 1TB Portable External Hard Drive USB 3.0, Red (STDR1000103)	$64.99	
AC Power adapter kit for camera	1	amazon.ca	Glorich EH-5 Plus EP-5B Replacement AC Power Adapter Kit for Nikon 1 V1 D500 D600 D800 D7000 D7200 DSLR Cameras (With Built-in Smart Decoding Chip). The product description included Nikon D750	$56.99	
Memory cards (32GB)	2	amazon.ca	Sandisk Ultra 32GB Class 10 SDHC UHS-I Memory Card Up to 80MB, Grey/Black (SDSDUNC-032G-GN6IN)	$20.33	
Wired remote control for the camera	1	amazon.ca	Foto&Tech Wired Remote Shutter Release Control Compatible w/ NIKON MC-DC2 for Nikon Z6 Z7 D7500 D750 D5600 D5500 D7200 D7100 D5200 D5100 D5000 D3200 D3100 D90 COOLPIX P1000 (Compatible w/ Nikon MCDC2)	$10.99	
Infrared remote control for camera	1	amazon.ca	Infrared Remote Control ML-L3 for Nikon	$5.49	
Daylight LED panels	2	B&H	Genaray SpectroLED Essential 500 Daylight LED Light	$297.90	$411.74
Stands for the lights	2	B&H	Impact Light Stand, Black (6’)	$39.98	
				TOTAL	**688.04**

The memory cards were used for the camera to save the set of captured photos, the master files were then stored in hard drives (two copies), the power adapter kit was used to keep the camera charged since the battery would run out of power due to the long hours of digitization (another option is to have two batteries at hand and a charger). The remote controls for the camera (gray rows in Table 3-3) were not used in the end, since I shot remotely from my laptop to be able to view the pictures taken on the laptop's screen.

In addition to the equipment in Table 3-3, the DDMAL Lab lent me a grayscale and color reference targets, each including a dimensional scale. The price online for these reference targets is approximately of $50 USD ($65 CAD) in total. I also bought a pair of flexible snake weights as the cradle configuration that I used at the end (this is, the open-at-one-side cradle configuration that will be presented in more detail in Section 3.1.3.1) required something to hold the pages in place, so that they do not fall over the page to be photographed.[98] The pack of two snake weights costs $20.50 USD (approximately $27 CAD). Images of the snake weights holding the pages of the GuatC 1 in the open-at-one-side cradle configuration are shown on the Experiment Chapter (Section 5.1).

3.1.3 The DIY Book Scanner Setup

Since I did not find any suitable digitization option in Guatemala after my first visit in November 2018, I decided to make my own do-it-yourself (DIY) book scanner to create high-quality color images of the manuscript. The three main parts of the book scanner were either built or borrowed:

[98] Some links to get archival supplies include: http://www.carrmclean.ca/, https://www.preservationequipment.com/, and https://www.gaylord.com. Here one can find book serpents (or snake weights) and weights bags among other things. I bought a pack of two flex snake weights from https://www.gaylord.com/Preservation/Book-%26-Pamphlet-Preservation/Repair-Tools-%26-Supplies/Other-Tools-%26-Supplies/Flex-Snake-Weight-%282-Pack%29/p/58086.

- **Book cradle.** The book cradle was built by German Thomae (my father) in Guatemala. See Section 3.1.3.1 for more details.
- **Camera.** A Nikon D750 camera was borrowed from the Marvin Duchow Music Library (McGill University). The camera's image sensor has an effective resolution of 24.3 Megapixels, which, given the dimensions of the manuscript pages (18 x 12 square inches), allows for an image resolution of 300 ppi (with no interpolation), an adequate archival resolution value according to the FADGI specifications.[99] The camera included an AF-S Micro NIKKOR 24-85 mm f/3.5-4.5G ED VR lens. It was lent together with a tripod.
- **Lights.** Two natural-light LED panels were borrowed from the DDMAL Lab (McGill University).

The book cradle and camera setup is shown in Figure 3-2. The LED panels for lighting are positioned one at each side of the book, as shown in Figure 3-3. In this setup, I am adhering to the following digitization guidelines, which can be found within the guidelines' compilation in Section 2.2.2:

- cradle opened at 110º
- avoid the use of a platen to flat the pages
- therefore, use the open-at-one-side cradle configuration to keep the digitized page as flat as possible without a platen
- keep the camera perpendicular to the page

[99] As indicated in Section 2.2.2, the ideal resolution would be 400ppi, to have a four-star project (FADGI's maximum score). A camera of 34.6 Megapixels would be necessary to achieve such a high-resolution value for these manuscripts. A 300 ppi resolution is consistent with a three-star FADGI project.

- use of a copystand to move the camera up/down to keep the distance between the page and the camera constant[100]
- use of black felt for neutral background and a stack of black cardboards to position the color patch at the same height as the digitized page
- keep the lights at an angle (approximately 45°) to avoid reflections on the camera lens and have the same amount of lighting at each side of the page to be digitized

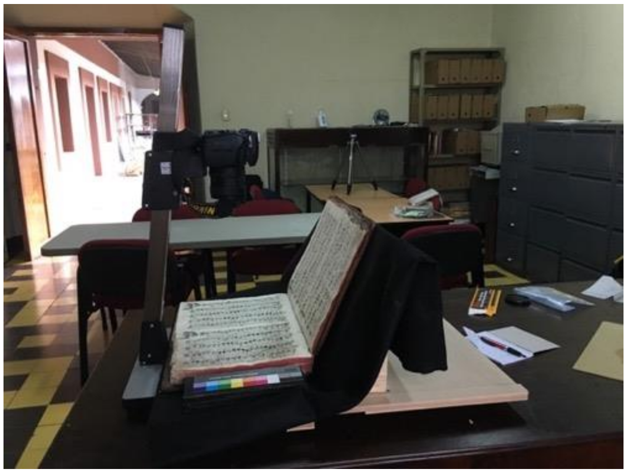


Figure 3-2: Book cradle and camera setup of the DIY book scanner.

Figure 3-3: DIY book scanner light setup, one on each side for even distribution of the light over the page.

[100] The copystand was borrowed from the Music Technology area at McGill University. Darryl Cameron, Chief of Electronics Technician, commented me that he had one in case it was useful.

3.1.3.1 *The Book Cradle*

The constructed book cradle, which will be shown later in the "The Book Cradle Specifications" subsection, allowed for two configurations: a facing-up V-shaped cradle and an open-at-one-side configuration. These two configurations of the cradle were a result of an at-home test.

The Home Test

To be better prepared for the actual digitization process and foresee any upcoming issues, I experimented at home with different configurations for the book cradle and evaluated which might work best with the equipment I had available. During this test, I tried to imitate the on-stie conditions as much as possible. I started the test with the facing-up V-shaped cradle configuration, which I made from a cardboard box by cutting it diagonally into two pieces, which I used as two big wedges for the book to rest open (see Figure 3-4, left). The facing-up V-shaped configuration has the advantage that one can use two cameras to photograph two pages at once. However, this means finding two cameras of the same resolution and using the same parameters to guarantee that the images of recto and verso pages have the same quality. Moreover, when testing this configuration with a book of my own library, I found out that it was hard to keep the camera perpendicular to and at a constant distance from the photographed page when using a tripod (see Figure 3-4, left); usually the scanners with a V-shaped cradle have a scanner rig with adjustable mounts or rails to place the camera. I also found out that, without a glass or plastic platen—which is not recommended for special collections—it is difficult to keep the pages flat in this V-shape configuration. Because of this, I prepared an alternative configuration for the cradle, the open-at-one-side configuration (see Figure 3-4, right). Although this configuration only allows for photographing one page at once, it proved to work better using a copystand that allowed to hold the camera perpendicular to and at a constant distance from the page. The photographed page still curved in this configuration, but it curved less than in the facing-up V-shaped arrangement. I used a pair of book serpents, one for holding the pages on the wedge to stop them from falling over the photographed page, and another for pressing on the photographed page to reduce its curvature—still, the page was not as flat as it would be if using a platen.

Figure 3-4: At home tests with a cardboard box cut diagonally in half to obtain two big wedges for the book to rest open. I used the two wedges for testing the facing-up V-shaped configuration (left) with the camera placed on the borrowed tripod. I took pictures with and without a platen. I only used one wedge to test the open-at-one-side configuration with the borrowed copystand, with a flexible snake weight to keep the pages resting in the wedge from falling on top of the photographed page.

After the results of this at home test, and although I suspected the best cradle configuration was the open-at-one-side, I decided that the book cradle to be used on site should allow for both configurations for quickly testing which one worked best with the original manuscript source. The next section provides the details about the built book cradle, diagrams of its configurations, and specifications about the cradle components and dimensions for others interested in similar DIY digitization projects.

The Book Cradle Specifications

German Thomae (my father) constructed the cradle to allow for both configurations: the facing-up V-shaped configuration (Figure 3-5) and the open-at-one-side configuration (Figure 3-6); so on the next visit to the archive (the digitization visit detailed on Section 5.1), I could test both configurations quickly, decide which one works best with the original manuscript source, and start the digitization process immediately.

I used the open-at-one-side configuration with the original source as it made it easier to: (1) set the camera perpendicular to the page being digitized by using a copystand, (2) maintain a constant distance between the camera and the page while moving through the pages of the book by moving the camera up and down the copystand arm, (3) keep the page of the manuscript source flat with no platen, and (4) provide an even distribution of the light on the page when using two sources of light, one at each side.

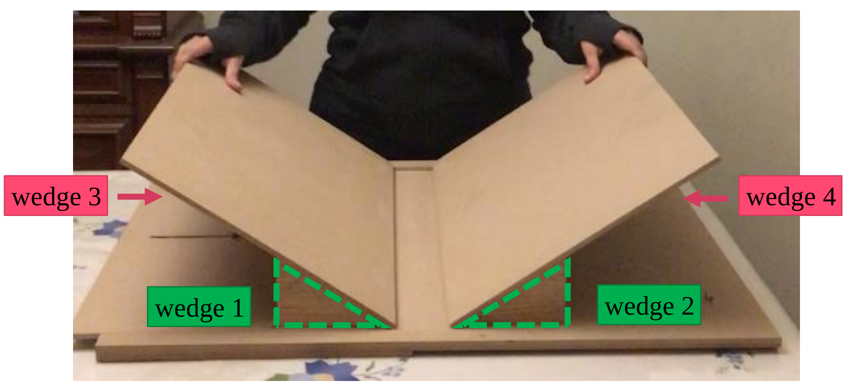


Figure 3-5: V-shaped book cradle facing-up configuration, which allows for photographing two folios at once. This configuration includes two doors, one of them is movable to fit the book spine. Each door is held by a pair of wedges at a 35º angle, allowing to open the book at 110º (angle between the doors).

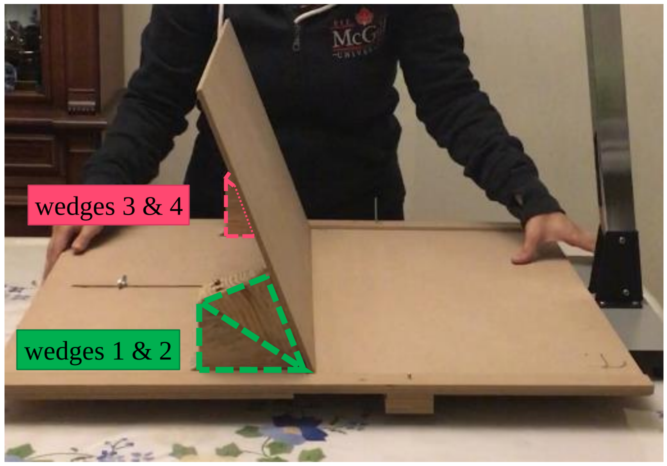


Figure 3-6: Book cradle open-at-one-side configuration, which allows for photographing one folio at a time using a copystand (see Figure 3-2). In this configuration, the right door has been removed, leaving only the left (movable) door. This door is held at a 70º angle by a pair of two 35º wedges stacked together, holding the book opened at 110º.

The cradle includes two doors (see Figure 3-5), one of them is movable (left door) and the other is removable (right door). In the facing upward configuration (Figure 3-5), the movable door allows to adjust the distance between both doors to fit the book spine in between. The doors are held at an angle by using four wooden wedges, each of them forms a 35º angle. In Figure 3-5, each door lies over two of these wedges (one at each side), which keeps the door at a 35º angle. In Figure 3-6, the one door in the cradle lies over the four wedges, having two of them stacked at each side, forming a 70º angle. Both configurations of the cradle (Figure 3-5 and Figure 3-6) allow for the book to be opened at a 110º angle, following the digitization guidelines outlined in Section 2.2.2.2.

The pieces used to build the book cradle and their dimensions are shown in Table 3-4, and their placement is shown in Figure 3-7–Figure 3-13. The cradle was made from one Medium-density fibreboard (MDF) panel of 4 x 8 square feet and ¼ inch width.

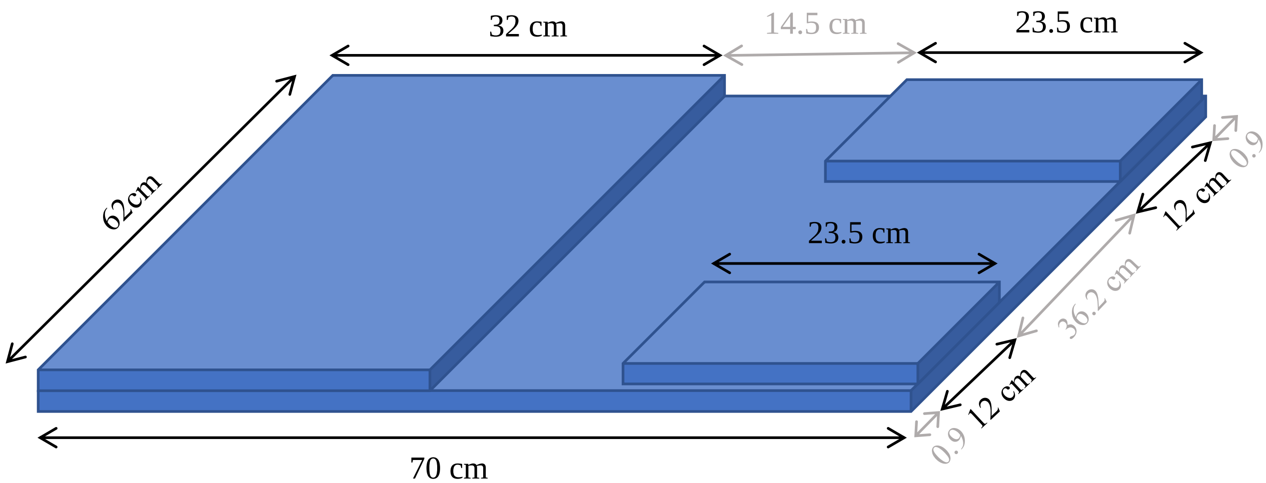


Figure 3-7: Bottom of the cradle (flipped cradle).

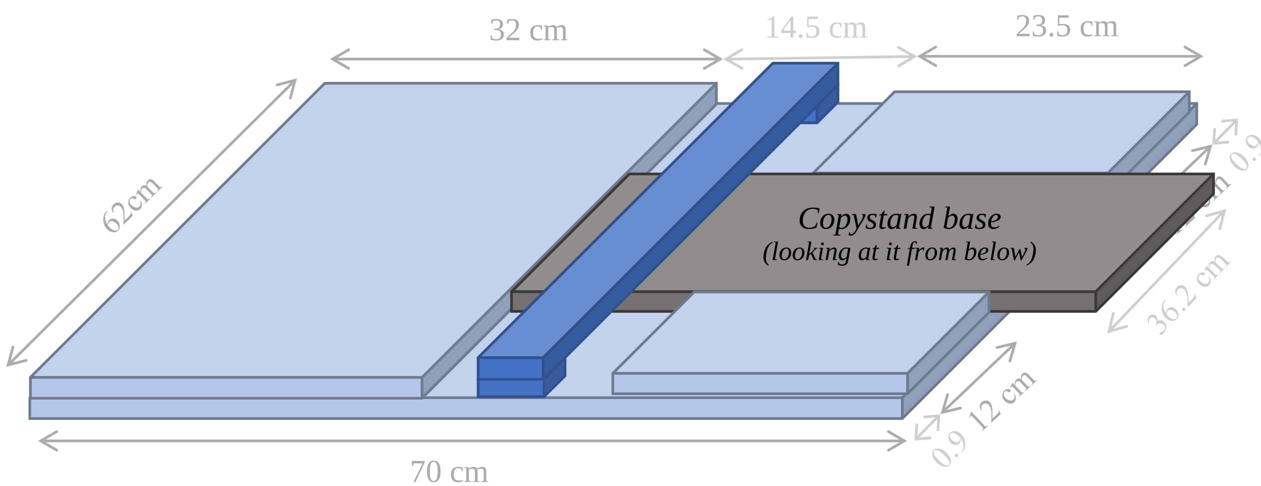


Figure 3-8: Bottom of the cradle (as in Figure 3-7) showing the joint bar used to fix the copystand base to the book cradle.

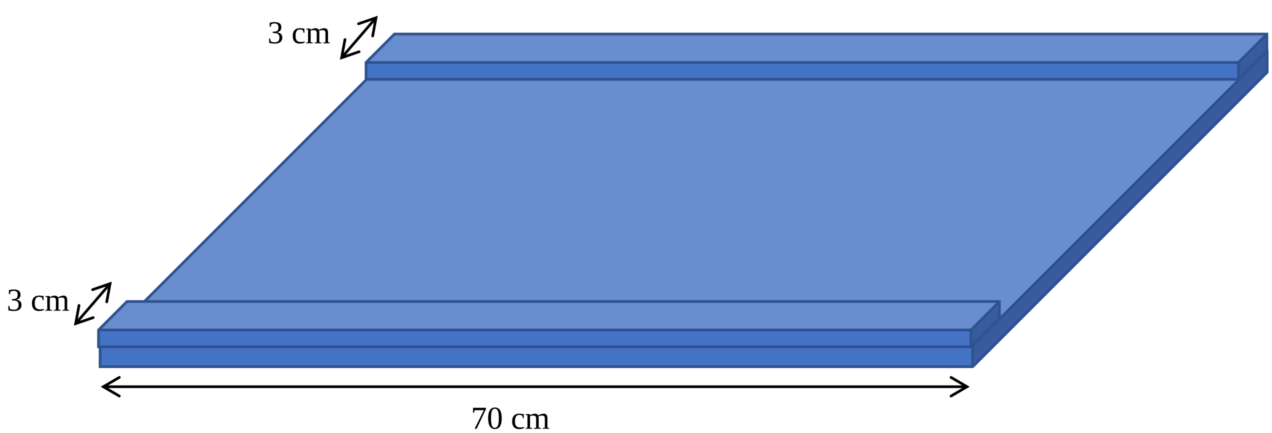


Figure 3-9: Top of the cradle.

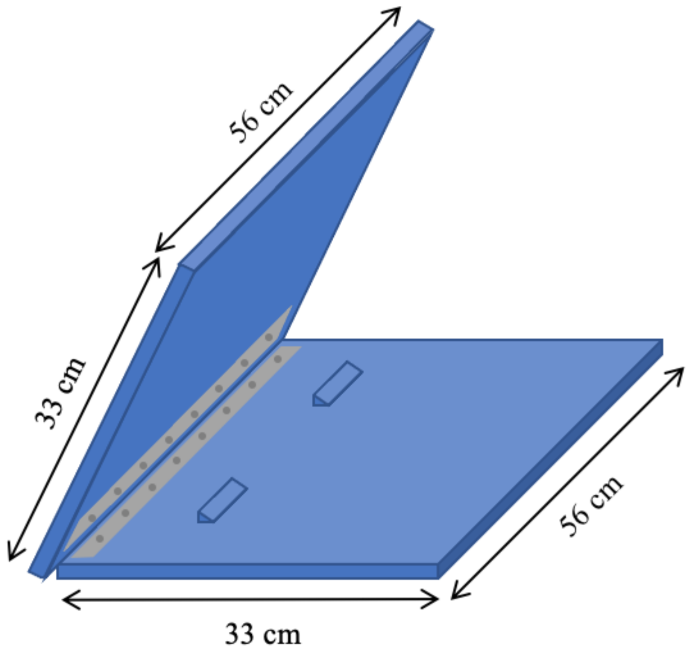


Figure 3-10: Right-side door. This is a removable door to be slid into the right side of the empty space in the cradle shown in Figure 3-9. The small triangular pieces work as stops for the wedges that keep the door open.

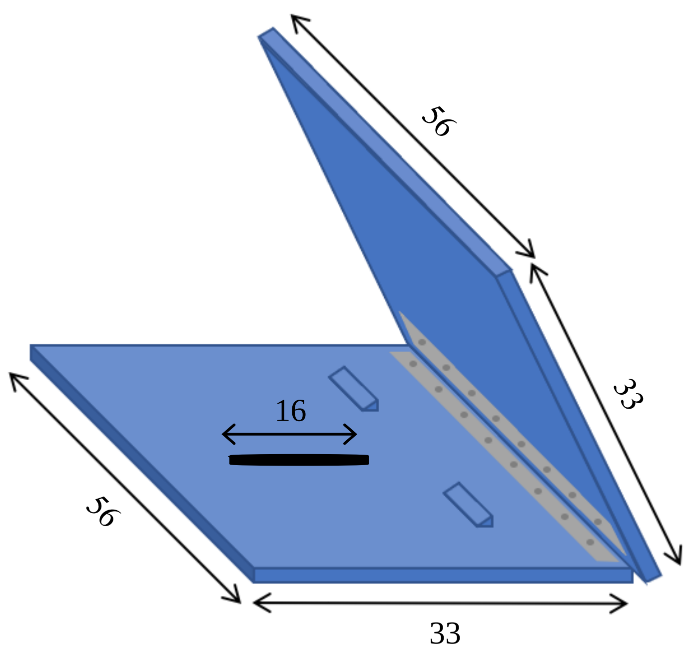


Figure 3-11: Left-side door. This is a movable door to be inserted at the left side of the empty space of the cradle shown in Figure 3-9. The large slide hole in the middle of its lower panel allows the door to stay in different positions by using a wing nut (which can be seen in Figure 3-5 and, more clearly, in Figure 3-6).

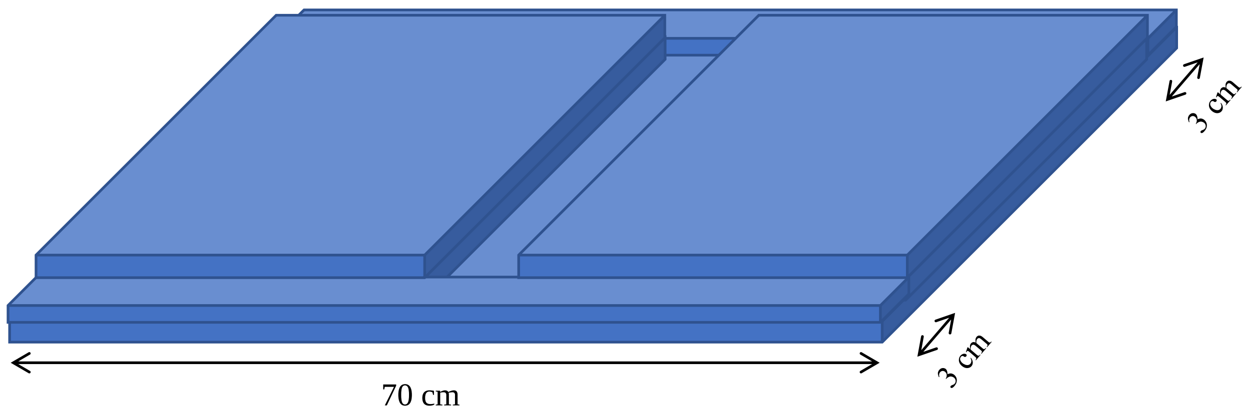


Figure 3-12: Cradle with doors closed.

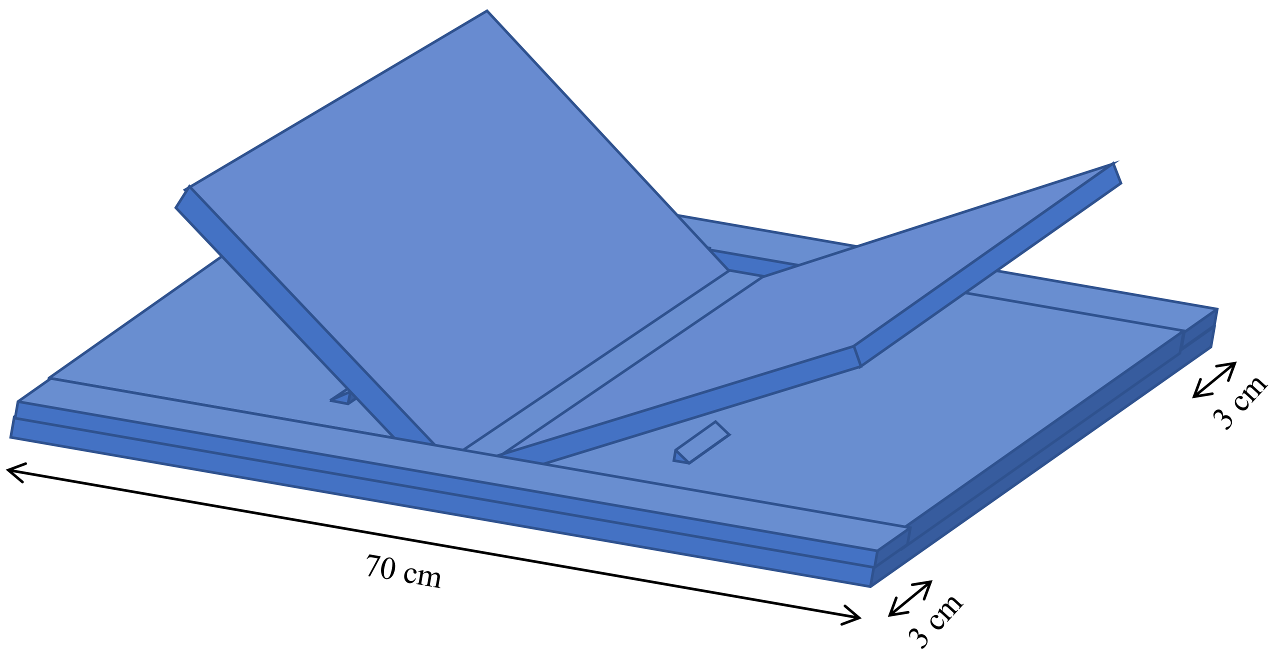


Figure 3-13: Cradle with both doors open. Wedges are used to keep the doors open (as shown in Figure 3-5 and Figure 3-6).

Table 3-4: Cradle components and their dimensions.

Type of Piece	Size (cm)	Number of Pieces	Description
Base	width: 62 length: 70	1	The **base** where the doors rest (largest piece in Figure 3-7 and Figure 3-9).
Bottom piece A	width: 62 length: 32	1	Three pieces to fit the base of the copystand in the bottom side of the cradle (Figure 3-7).
Bottom piece B	width: 12 length: 23.5	2	
Joint Bar piece A	width: 62 length: 5	1	These three pieces provided a joint bar that allowed to fix the copystand base in the place delimited by the three bottom pieces (see Figure 3-8).
Joint Bar piece B	width: 3.5 length: 5	2	
Lateral guides	width: 3 length: 70	2	Two **lateral guides**, pasted on top of the base, to help sliding the doors in (Figure 3-9).
Door panels	width: 56 length: 33	4	These four pieces are connected in pairs with a continuous hinge to make the two doors (see Figure 3-10 and Figure 3-11). The doors are to be inserted in the upper part of the cradle by sliding them in (see Figure 3-12 and Figure 3-13).
Hinge	length: 54	2	Continuous hinge to be used in each door (see Figure 3-10 and Figure 3-11).
Wedges	height: 7.8 length: 11.3 (minus the tip of 1.9 cm) width: 4.5	4	Each wedge has a 35º angle, as can be seen given their height and length. The points of the wedges were cut, and they were used to fix the wedges in place (see four small triangular pieces in the doors, Figure 3-10 and Figure 3-11).

3.2 The MuRET Framework

For the second stage of the workflow, I used the Music Recognition Encoding and Transcription (MuRET) framework to perform optical music recognition (OMR) on the GuatC 1 choirbook. The reason for choosing MuRET is that, among all the OMR tools presented in Section 2.3, it has the most support for handwritten mensural notation. *Aruspix* only works with printed mensural notation, which is not the case of the Guatemalan choirbooks that are all manuscript sources, and *OMMR4All* only supports neumatic notation. SIMSSA’s OMR process run in *Rodan* (see Section 2.3.3.2.2) does allow some processing of handwritten mensural notation since most of the jobs involved are applicable to different types of sources and notations. However, the last two jobs (i.e., the *Music Encoding* job and the *Neon* job), which cover the last step of the OMR process by allowing encoding the music in a symbolic format (MEI) and correcting the encoded symbols through a graphical interface, are designed specifically for neumes. Therefore, to have full support for OMR of mensural sources, these jobs would need to be rewritten to work with mensural notation. This is relatively easy for the Music Encoding job, but harder for the Neon webapp. On the other hand, MuRET has support for performing and correcting the results of all OMR stages for handwritten mensural notation (as well as printed mensural notation and Common Western Music Notation). Moreover, MuRET includes now machine learning algorithms for the document analysis, symbol recognition, and holistic staff recognition (see Section 2.3.3.3) that were trained on handwritten mensural notation from Spanish sources and can be used as they are or as a starting point to train other models. The following sections present the MuRET webapp and the steps involved in the OMR process, focusing on the various interfaces that the user has to navigate to perform OMR on a collection. It starts with the collection and subcollection interfaces to access the images (Section 3.2.1); this is followed by the three interfaces related to the OMR process of a given image: the document analysis (Section 3.2.2), parts assignment (Section 3.2.3), and transcription interfaces (Section 3.2.4); and concludes with the summary interface (Section 3.2.5) that shows which OMR steps have been completed and any comments left by the user for a given image.

3.2.1 **Collection and Subcollection Interfaces**

The "collection interface" shows all the items in the user's collection; in this case, all the items in the GuatC 1 choirbook (see Figure 3-14). Each item is called a *subcollection* since it consists of a collection of images. In this case, each piece of the GuatC 1 choirbook is a subcollection. By clicking on the subcollection link, the user can access the images (see Figure 3-15 and Figure 3-16). For any image, there are three processes available: *document analysis*, assignment of the *parts*, and *transcription*; each of these processes has its own interface. There is also an additional interface that opens when clicking on the link of the *summary* column (Figure 3-15). The following sections present each of the four interfaces and their purpose.

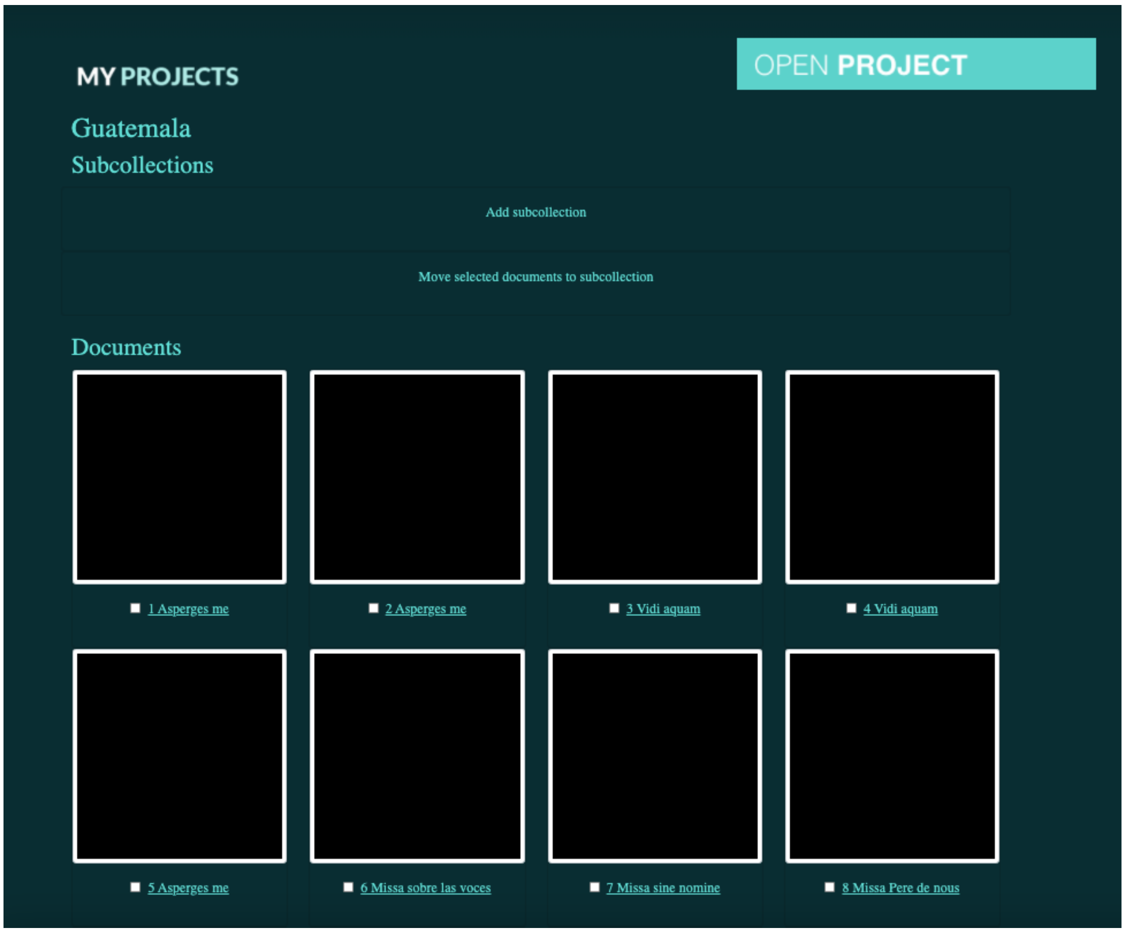


Figure 3-14: MuRET´s collection interface. This interface shows all the subcollections within a collection. In this case, it is showing the 'Guatemala' collection, which has various subcollections of images, each representing one composition in the manuscript.

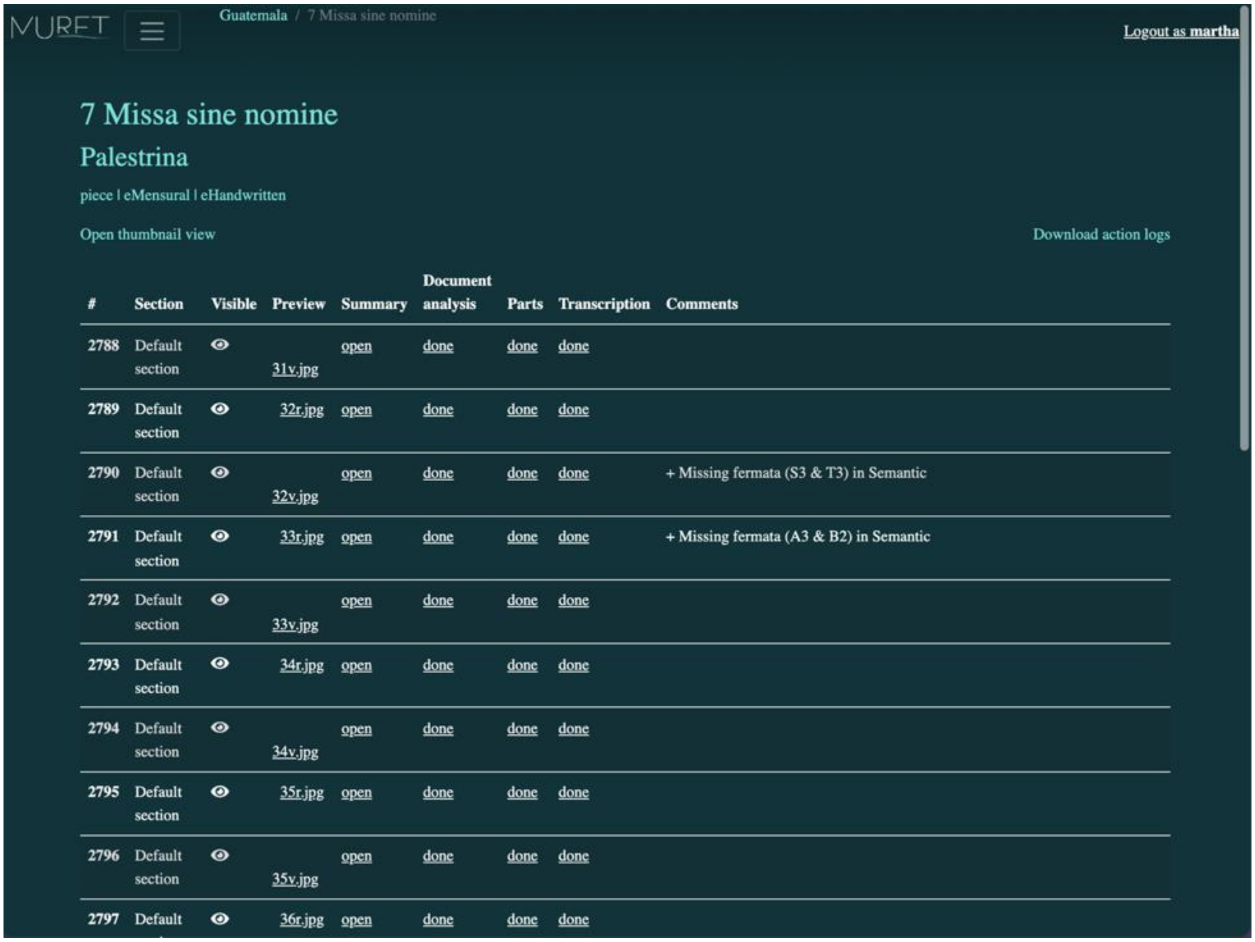


*Figure 3-15: Subcollection interface, list view. This shows the subcollection '7 Missa sine nomine' by Palestrina, listing all the images belonging to this piece. The columns **Document analysis**, **Parts**, and **Transcription** show the different processes involved in the OMR of the images. Each of these columns can have a status of **open**, **in progress**, or **done**. The notes in the **Comments** column can be entered when selecting an entry in the **Summary** column.*

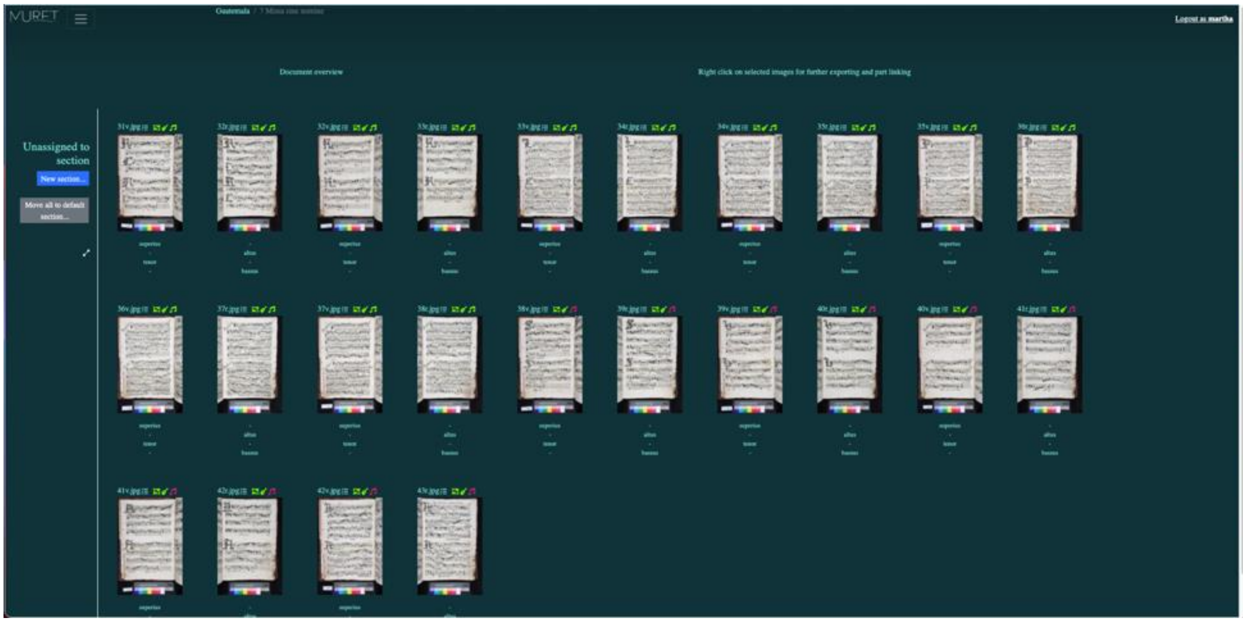


Figure 3-16: Subcollection interface, thumbnail view. Same subcollection as in Figure 3-15 ('7 Missa sine nomine' by Palestrina) but with all its images displayed as thumbnails. The status of the ***Document analysis****,* ***Parts****, and* ***Transcription*** *processes is represented by the square, guitar, and beamed-notes icons above each image, respectively, with gray meaning* ***open*** *(i.e., not started), red meaning* ***in progress****, and green meaning* ***done****. The checklist icon next to the filename allows for seeing the image and entering comments.*

3.2.2 Document Analysis Interface

This interface provides support for the first OMR stage: document segmentation. Here the user can manually draw bounding boxes indicating the different regions of the manuscript (e.g., staff regions, lyrics, and text), or choose to use the Selectional Auto-Encoder (SAE) model trained for the automatic detection of staff regions and correct the bounding boxes generated by it if needed (see Figure 3-17).

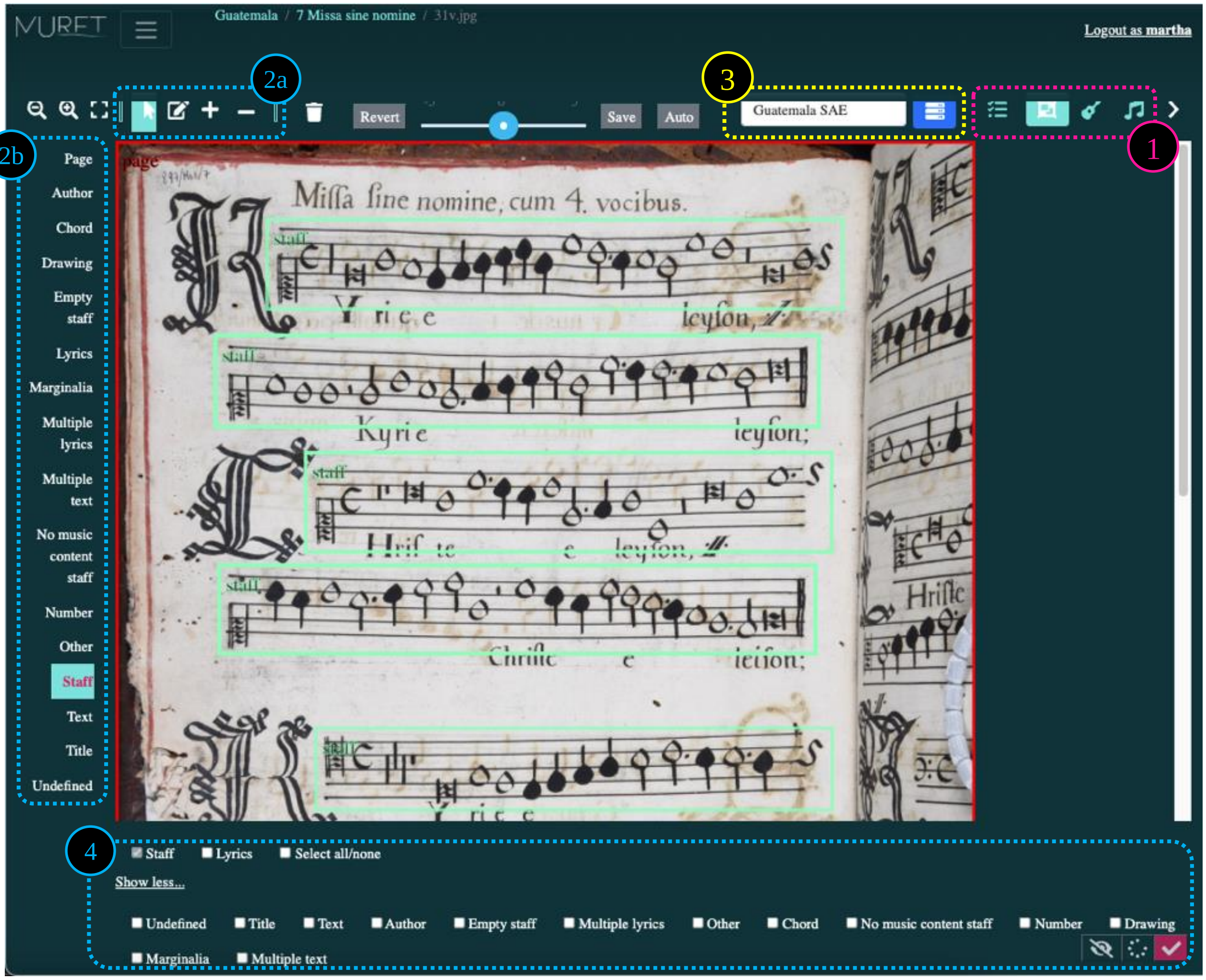


Figure 3-17: Document analysis interface, as indicated by the highlighted button in (1). In this interface, the user can draw bounding boxes (add, edit, or remove them using the tools in 2a) to identify the different types of regions present in a document (using the categories given in 2b). The user can also select the Selectional Auto-Encoder (SAE) model to automatically detect the staff regions (3). The bottom panel (4) provides checkboxes to select the type of region to display (in this case, only the staff checkbox is checked and, therefore, only the staff regions are displayed).

3.2.3 **Parts Interface**

In this interface, the user can select the staff regions created in the Document Analysis interface and assign them to a part (i.e., voice) as seen in Figure 3-18.

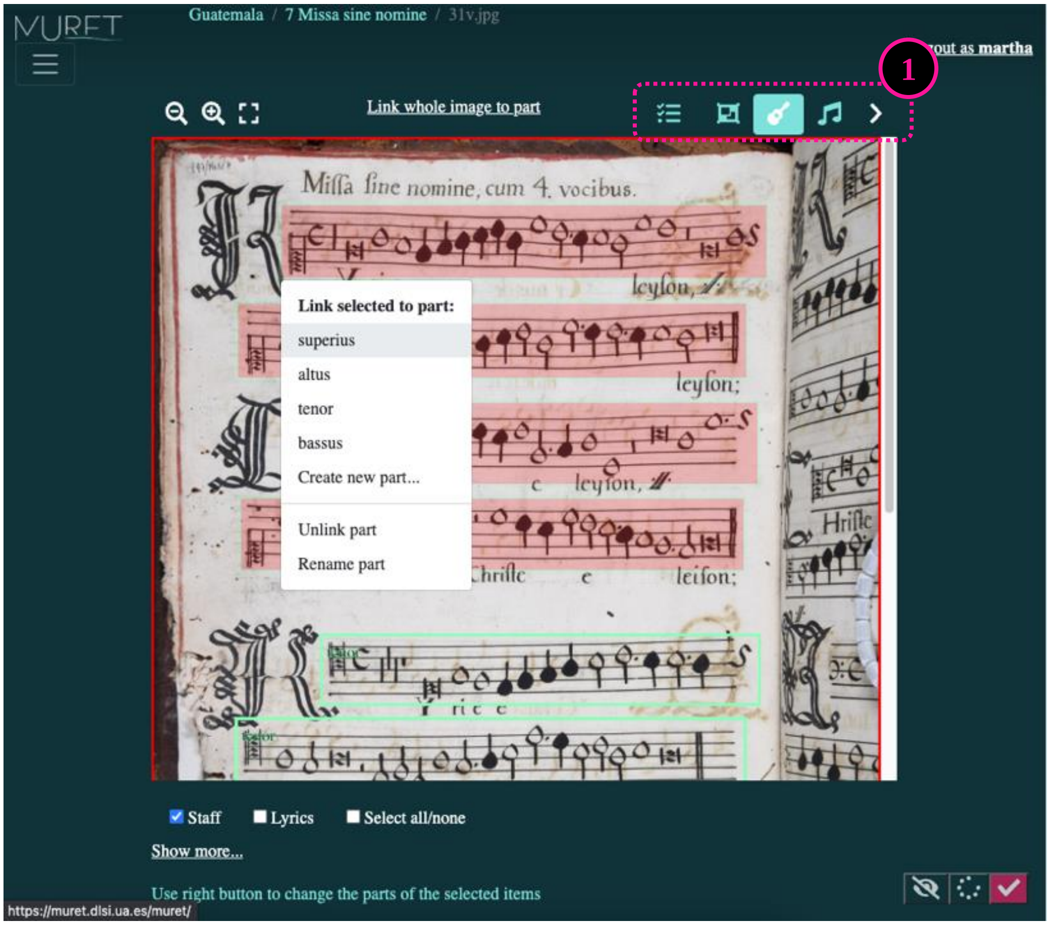


Figure 3-18: Parts interface, as indicated by the highlighted button in (1). In this interface, the user can select one or multiple staves (or the whole page) and assign those staff regions to a particular part (i.e., voice).

3.2.4 **Transcription Interface**

This interface provides support for the second and third OMR stages: symbol recognition and semantic encoding. Regarding the symbol recognition stage, the user can select any of the staff regions created in the Document Analysis interface, as well as the holistic staff-level recognition model to classify the symbols in that staff (see Figure 3-19). The selected staff is shown as a region of the image, and the recognized symbols are shown in the "agnostic" transcription below.[101] The coordinates of the bounding boxes of the recognized symbols shown in the selected staff-region

[101] The agnostic representation of the symbols is a representation based solely on glyph class and position in the staff, with no musical/semantic meaning (see page 102 for details).

image panel (see panel 3 in Figure 3-19) are approximate as the holistic staff-level symbol recognition model is designed to retrieve the sequence of symbols in the staff, not their coordinates; however, if the user wants the coordinates of the symbol's bounding box to be more accurate, they can drag the corners of the bounding box with the computer mouse to correct it. The user can also correct any of the recognized symbols in the agnostic transcription (see panel 4) by clicking on the symbol and changing either the category of the symbol (by using panel 4a) or its position in the staff (by using the arrows in panel 4b). If any symbol is missing in the transcription, the user can employ the symbol-level classifier to draw a bounding box on that symbol and have it automatically recognized (see red box); they can also manually correct the result if needed. When there is no holistic staff-level recognition model suitable for the manuscript source, one can use the symbol-level classifier to create ground truth for training a new (more suitable) holistic staff-level recognition model. Currently, these new models are trained offline by MuRET's developers (with the training data exported from MuRET) and then they upload the models into MuRET for the users.[102]

[102] Juan Carlos Martínez Sevilla has recently (in 2022) provided access to the command-line script and a Google Colaboratory Python notebook for training the Document Analysis SAE, the CRNN for holistc staff-level symbol recognition, and the symbol-level classifier models: https://github.com/JuanCarlosMartinezSevilla/MuRET-UserTool. Downloading training data from MuRET and uploading classification models to it is currently in the works. Soon, the user will be able to download the training data from MuRET, run Martínez Sevilla's scripts to generate the classification models, and upload these models back to MuRET to use them for document analysis and symbol recognition.

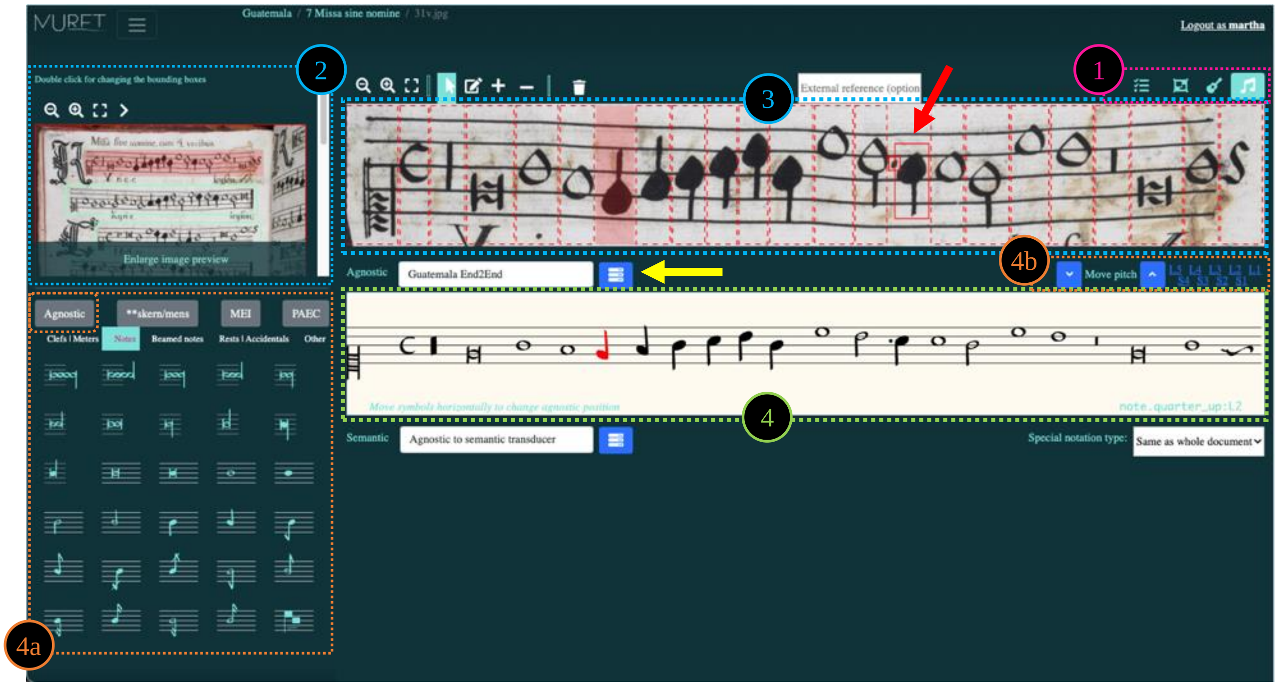


Figure 3-19: Transcription interface, as indicated by the highlighted button in (1). The panels shown in this interface are: (2) image display to allow for staff selection, (3) selected staff-region image panel, (4) agnostic transcription panel, and (4a) symbol category options and (4b) staff-position editing options, both of which allow for correcting the results of the holistic staff-level and symbol-level classifiers. The yellow arrow shows the place where the user selects the type of holistic (end-to-end) classifier to use in the selected staff region to obtain the agnostic transcription. On the other hand, the red arrow shows an example of the use of the symbol-level classifier.

Once the agnostic transcription of the symbols in the staff is correct, the user can ask MuRET for its semantic representation encoded in **mens (see Figure 3-20). This will provide a second transcription panel rendering the encoded symbols. Any errors can be corrected by switching the left-lower panel from ‘Agnostic’ to ‘**skern/mens,’ which will present the **mens encoding as a table. Each token in the table is linked to the corresponding symbol in the semantic transcription panel, the agnostic transcription panel, and the selected staff-region image panel (see highlighted symbols and regions in Figure 3-20). The user can correct the semantic encoding of the recognized symbols by editing the **mens code provided in the table. Currently, the **mens encoding requires very little correction; regular note’s *pitch* and *note shape* are correct since this information comes from the *position in the staff* and *symbol category* data contained in the note's agnostic encoding. However, at the moment, it is common for the user to correct the semantic encoding of ligatures. This is because MuRET has only one agnostic symbol to represent all types of ligatures (regardless of the different number of notes in a ligature, and their individual note values—i.e., note shapes—and pitches), a shortcoming that its developers are addressing.

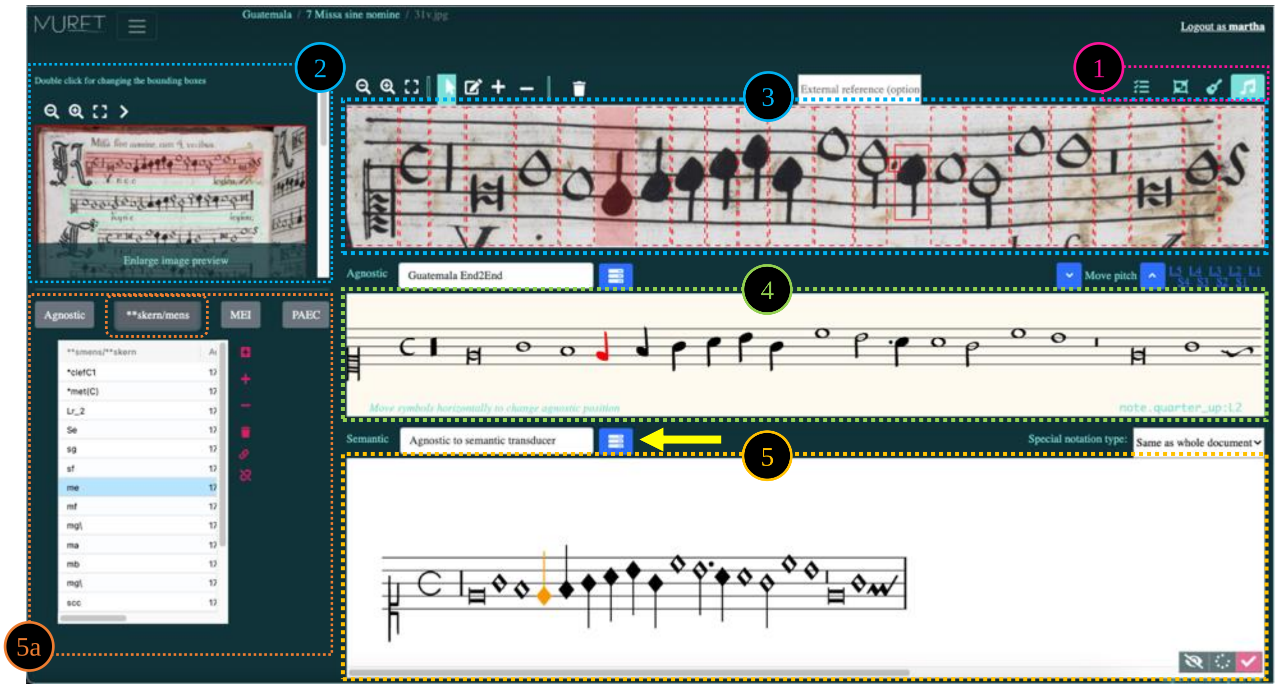


*Figure 3-20: Transcription interface, as indicated by the highlighted button in (1). The panels shown in this interface are: (2) image display to allow for staff selection, (3) selected staff-region image panel, (4) agnostic transcription panel, (5) semantic transcription panel generated from the agnostic one when clicking on the button shown by the yellow arrow, (5a) **mens table that allows for the correction of the semantic transcription. Every entry in the **mens table is linked to the corresponding symbol in the panels 3, 4, and 5 (see the selected entry of the table 'me' and the highlighted 'E minim' in the transcription panels and the selected staff-region image panel).*

The **mens table can also be used to encode the perfect / imperfect / altered quality of mensural notes, which provide the actual duration of the notes and are necessary for voice alignment. Currently, MuRET does not include an automatic way to determine these quality values and the user needs to enter them manually using the **mens syntax. There are plans to include the quality values in the agnostic-to-semantic translation process in the future. However, for the purposes of this dissertation, the encoding of quality values in MuRET—either manually or (eventually) automatically—is not an essential feature because this is handled by the next tool of the workflow, the Measuring Polyphony Editor (MP Editor, presented in Section 3.3). In addition to the encoding of the notes' perfect / imperfect / altered quality, the MP Editor allows for editorial corrections that are very much needed to have a score with its voices correctly lined up.

3.2.5 **Summary Interface**

In the Summary interface, one can enter comments for the particular image, as well as view the regions and status of the document analysis, parts, and transcription processes (see Figure 3-21).

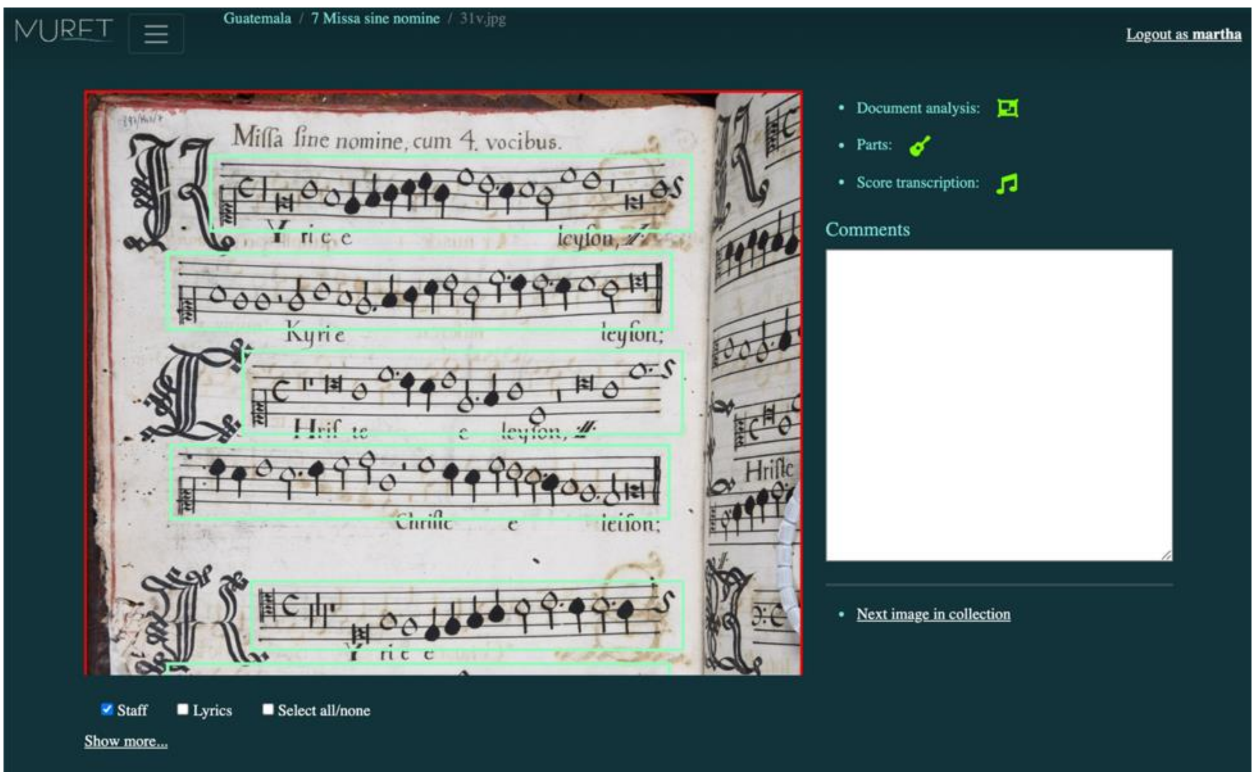


Figure 3-21: Summary interface to enter comments on an image. It also provides access to the document analysis, parts, and transcription processes.

Finally, to perform the last step of the OMR process, music encoding, the user needs to return to the subcollection interface in thumbnail view, select the files (i.e., images) that they want to save in a symbolic format, right click, and select the “Engrave and export (MEI)” option (see Figure 3-22). This will provide the option to “Export parts and facsimile for MP Editor,” which exports a Parts-based MEI file that complies with the MP Editor requirements. Details about this MP-Editor-compliant MEI Parts file are presented at the end of Section 3.3 about “The MP Editor,” and the process to make the output of MuRET compliant with the MP Editor is presented in Section 3.4. The exported MEI Parts file is a file that encodes the music as a collection of parts (i.e., voices) rather than as a score. Of course, an MEI Score file can be exported too, but this would not have the correct note alignment since the quality of perfect / imperfect / altered of the notes is not provided in the semantic encoding—though it can be provided, but it involves manually encoding this information rather than automatically encoding them as mentioned before.

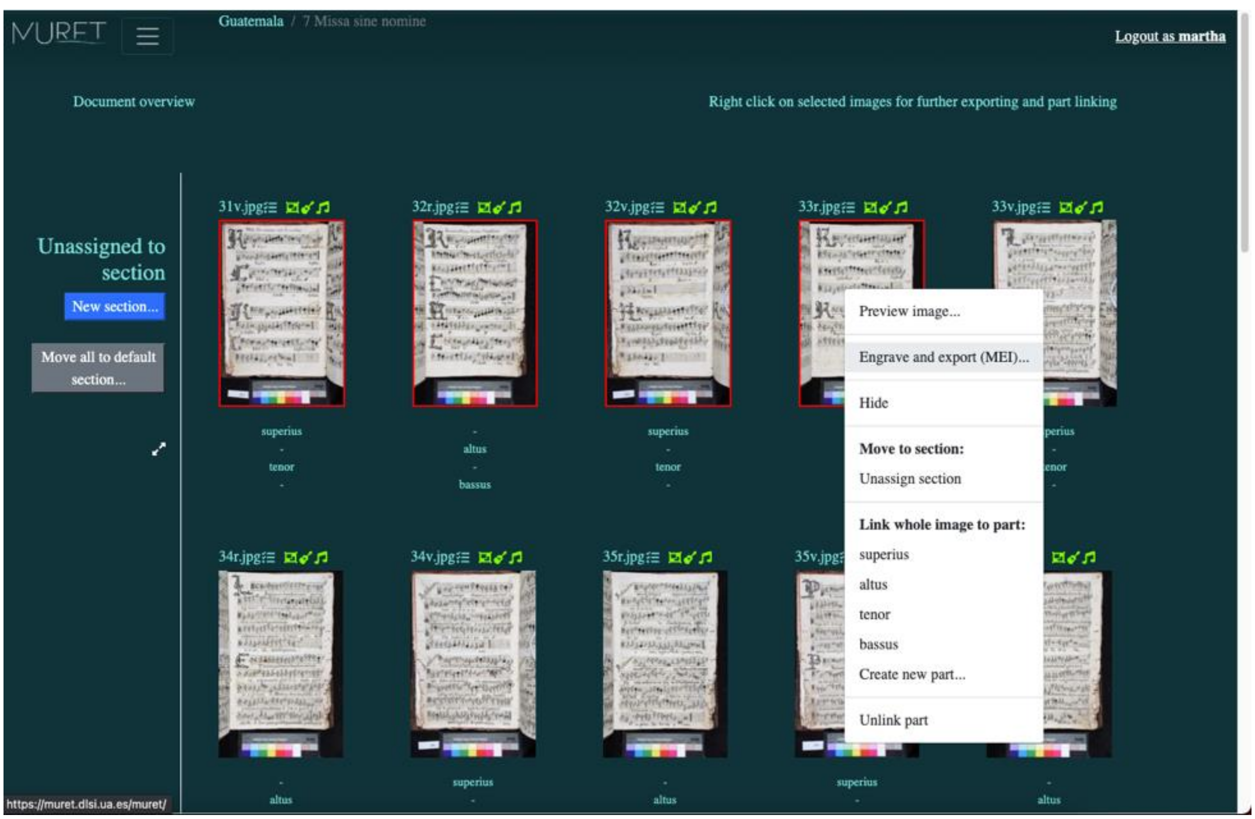


Figure 3-22: Selection of images (folios) to export into an MEI file in the subcollection interface in thumbnail view.

3.3 The MP Editor

The next (and final) tool used in the workflow is the Measuring Polyphony Editor (MP Editor). MuRET was used to retrieve an MEI file encoding the notes for each part (i.e., voice) of the piece. However, the parts are still not lined up in a score, which would provide the full rhythmic information about the notes and would also allow for visualizing vertical sonorities that the separate-parts arrangement does not display. To retrieve a symbolic score correctly lined up from the OMR results, I am using the MP Editor given its automatic voice alignment and manual editorial correction functionalities. This section presents the steps followed in the MP Editor to enter a mensural piece and retrieve a properly lined-up score encoding the piece in (Mensural) MEI, with the correct durational values of the mensural notes and including editorial corrections for scribal errors present in the original source.

3.3.1 Upload Page

The first page displayed in the MP Editor is the upload page (see Figure 3-23). Here, the user provides either the link to a IIIF manifest pointing to the digital images of the musical source they want to work with (top part in Figure 3-23), or a Parts-based Mensural MEI file that follows the encoding conventions of the MP Editor (bottom part in Figure 3-23). The uploaded file is usually a file that the user has generated before with the MP Editor and that they want to upload again to keep working on it.

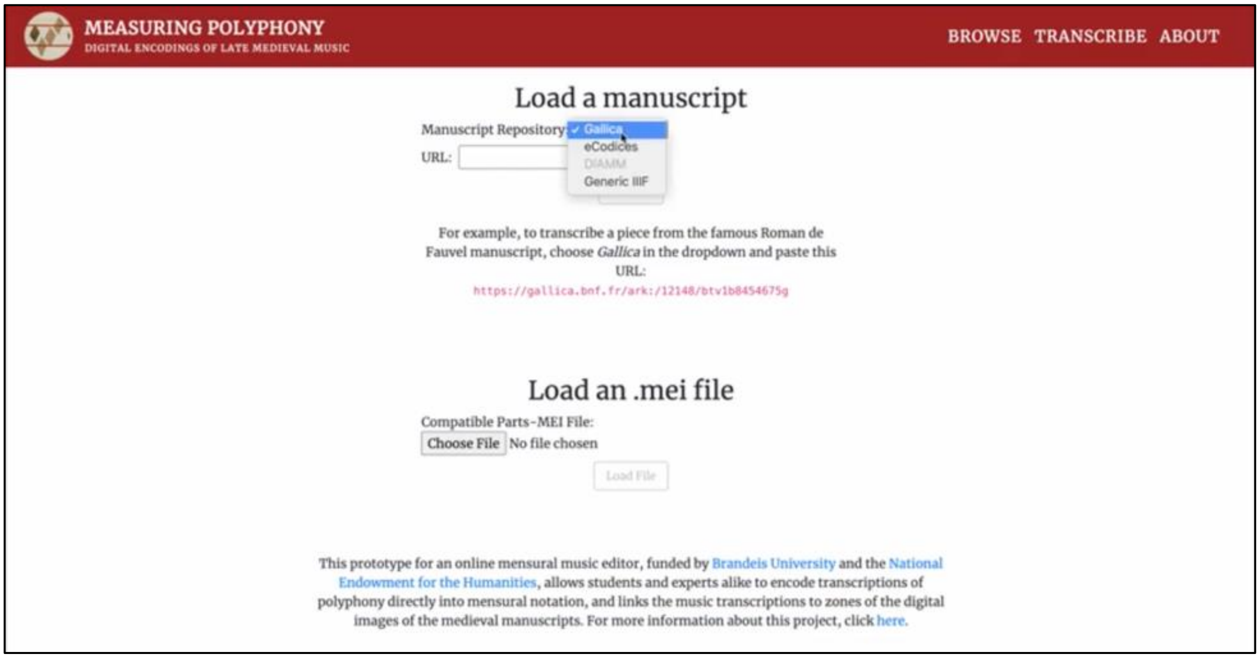


Figure 3-23: MP Editor's upload page. The user can select whether to upload a manuscript, using a link to the IIIF manifest, or a Parts-based MEI file, which encodes the individual parts (i.e., voices) of the piece.

3.3.2 **Metadata-Entry Page**

The second page of the MP Editor is the metadata-entry page. Here the user can enter basic metadata information about the piece they are about to transcribe, including the title, composer, notation style, manuscript siglum, and genre. For notation, there are three categories available in the dropdown menu: black mensural *Ars antiqua* notation, black mensural *Ars nova* notation, and white mensural notation. The rules of automatic voice alignment depend on the notation chosen by the user. The Guatemalan sources are mostly written in white mensural notation. The genre field also has a few options available in the dropdown menu. At the moment, these options are motet, song, conductus, mass movement, and plainchant. This interface also allows for adding different types of contributors: encoders, proofreaders, and editors. Once the metadata-entry phase is done, the user can click on the 'Continue to Music Input' button to move into the next page, the *input editor* page.

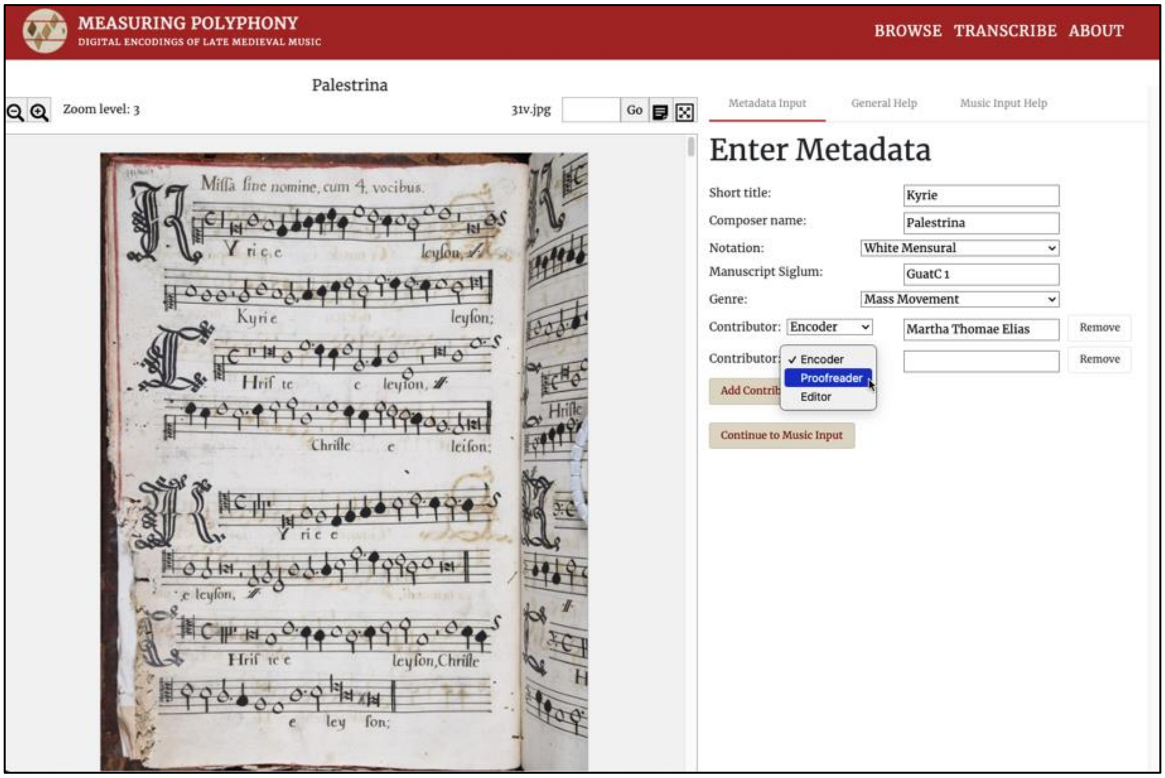


Figure 3-24: MP Editor's metadata-entry page.

3.3.3 **Input Editor**

In the 'input editor' page, the user enters the notes of each system for all voices by selecting the area corresponding to the system in the image and then typing the notes using a computer keyboard (see Figure 3-25). The user must assign a voice type and mensuration to the system they are entering (top-right panel in Figure 3-25).[103] The information for each system is internally stored in the MP Editor using **mens, the Humdrum format for mensural notation. One can download the **mens encoding of the system using the 'Download System Humdrum' button.

[103] The mensuration of the voice only needs to be entered once. Then every system assigned to that voice will be assigned that mensuration.

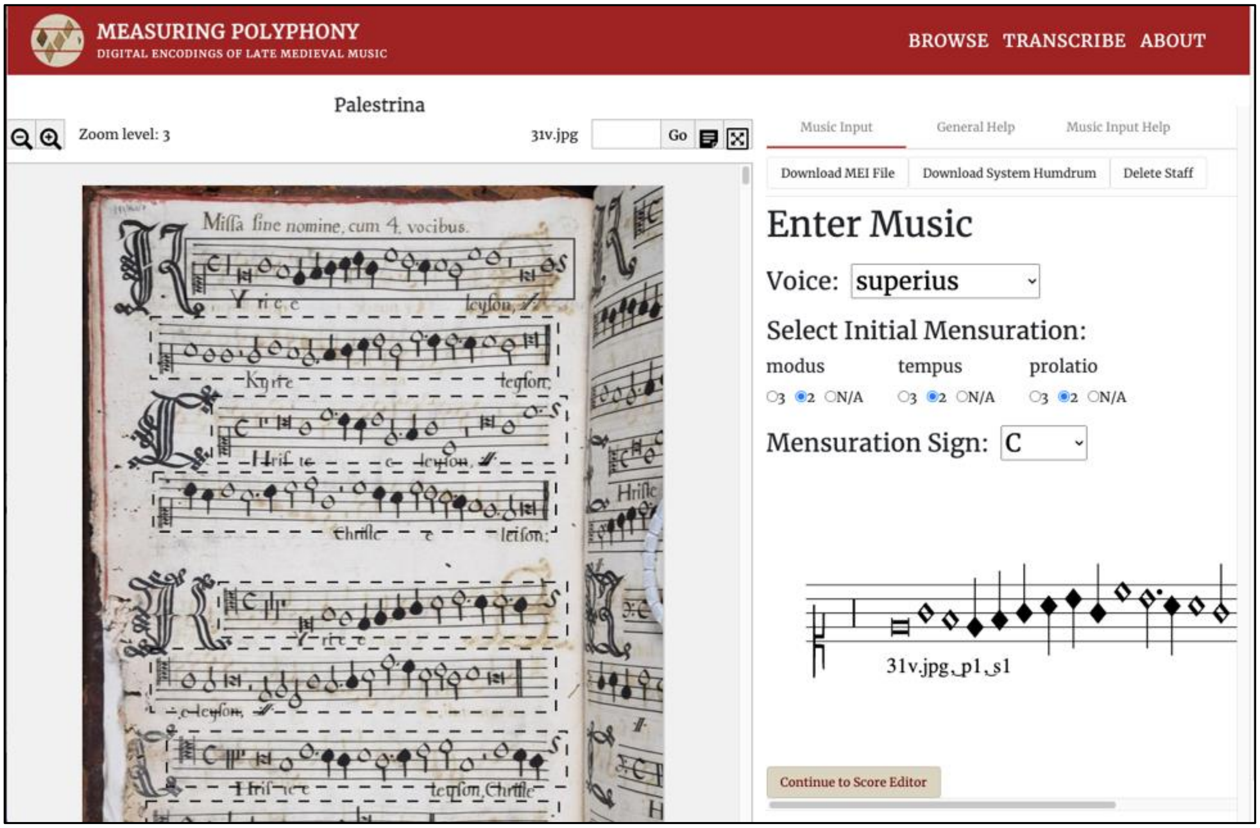


Figure 3-25: MP Editor's input editor page. In this interface, the user enters the music of each system of a voice by selecting the area of the system in an image, selecting its voice from a dropdown menu, providing the mensuration for the voice using a set of radio buttons, and typing the sequence of notes in the system using a computer keyboard.

At this point in the MP Editor, the entered piece is fully encoded in a Parts-based Mensural MEI file; this is, an MEI file that encodes mensural values using the <parts> element that contains the collection of parts, each encoded in a child <part> element. To automatically line up the voices and produce the Score-based Mensural MEI file (that uses a <score> element rather than <parts>), the user must click on the 'Continue to Score Editor' button, which transfers them into the *score editor* interface. Details about these two types of MEI files, parts- and score-based, as exported by the MP Editor, will be presented at the end of the next section.

3.3.4 **Score Editor**

The user launches the automatic scoring up of the voices by moving into the 'score editor' interface of the MP Editor. This automatic scoring-up functionality is a re-implementation of my scoring-up Python script (Thomae 2017) into JavaScript plus additional functionalities to deal with older repertoire, which is the focus of the Measuring Polyphony Project. The output of this automatic voice alignment process is a Score-based Mensural MEI file, which uses the element <score> and provides the perfect, imperfect, and altered value of the notes in the @dur.quality attribute, with which the notes can be lined up. This file is rendered on screen, showing the piece lined up as a score (see Figure 3-26).

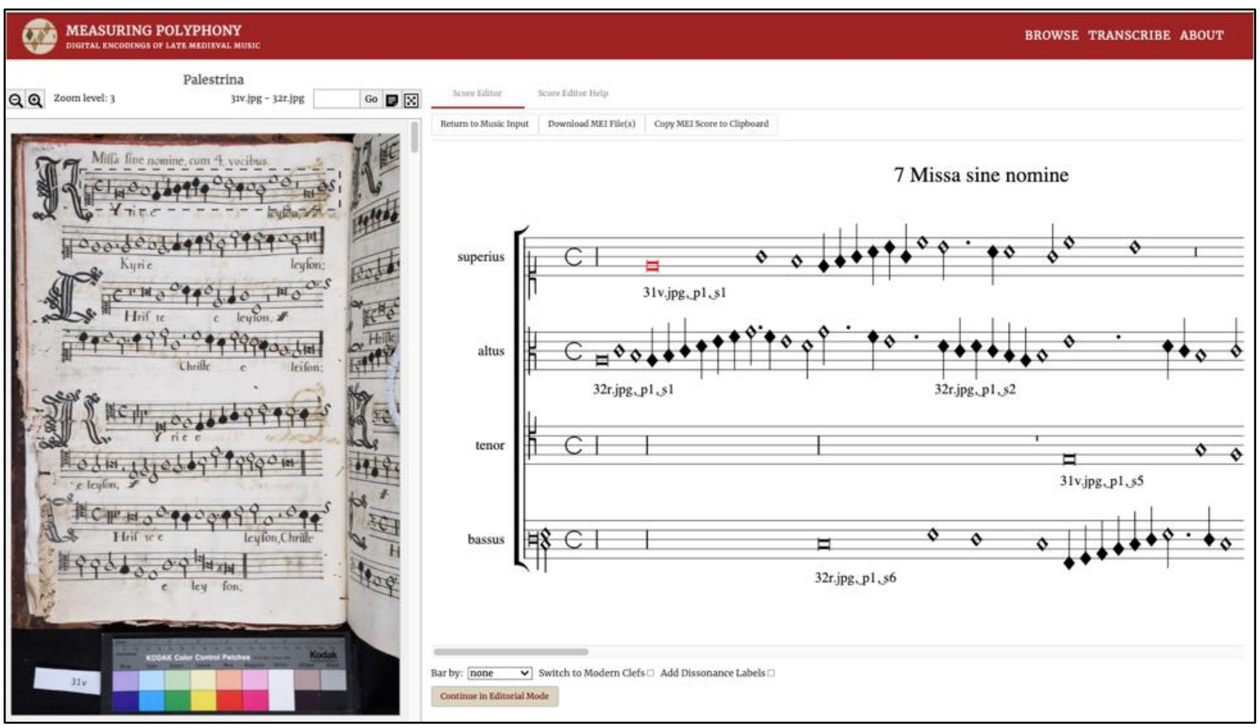


Figure 3-26: MP Editor's score editor page. In this interface, the piece is rendered as a score, based on the Score Mensural MEI file encoding it. The user can further correct the piece and the voices will be automatically re-aligned accordingly.

In the score editor interface, the user can make further corrections to their input (the notes typed in). They can also click on the 'Continue in Editorial Mode' button to correct scribal errors that will be stored in the Score-based MEI file as editorial corrections, encoding both the original and corrected readings (using the <sic> and <corr> MEI elements). With any correction, editorial or not, the piece will be automatically lined up accordingly.

To facilitate the correction process, the user can use the options on the bottom of the score editor to bar the piece by different note values (including long, breve, and semibreve) and to switch into modern clefs, as shown in Figure 3-27.[104]

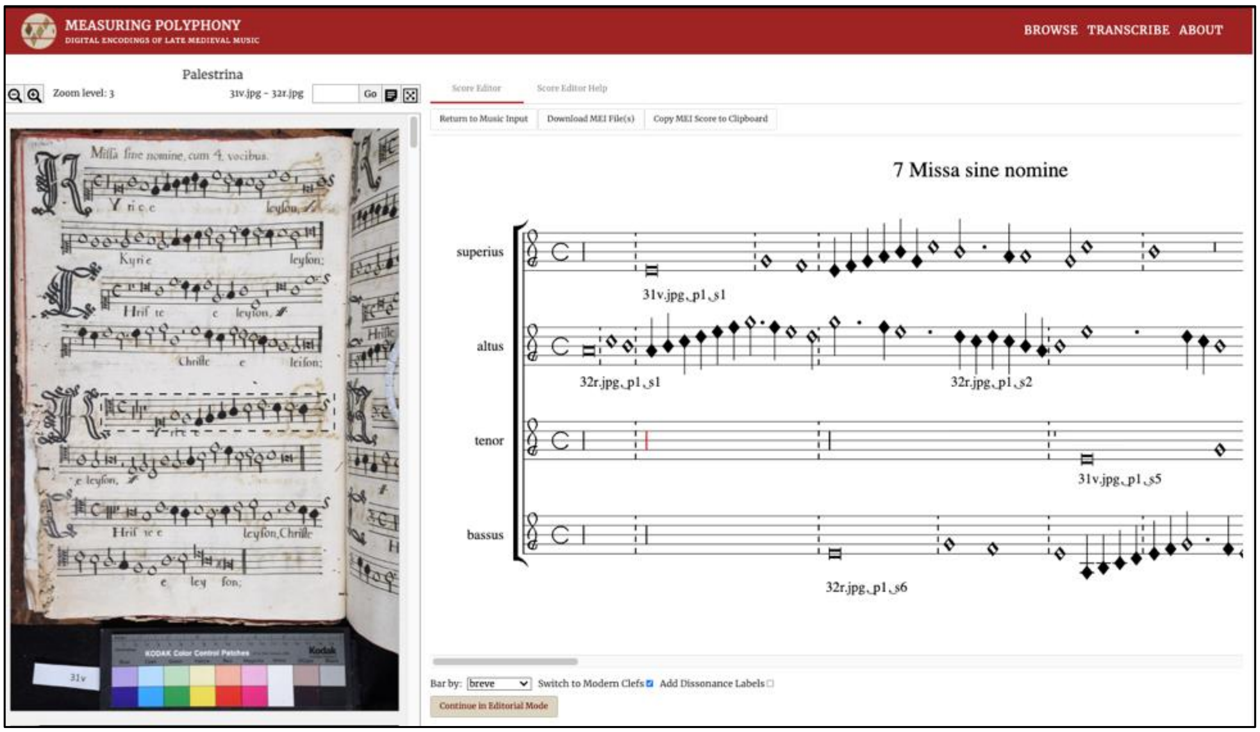


Figure 3-27: Same piece as in Figure 3-26, but barred by the breve and with modern clefs using the options at the bottom of the screen.

Finally, the user can download the MEI files encoding the piece as a collection of parts or as a score. When clicking on the 'Download MEI File(s)' button, the user is given the option to save the Parts-based MEI file, the Score-based MEI file, or both. The Parts-based MEI file can be re-uploaded again into the MP Editor using the upload page (shown in Figure 3-23) to continue

[104] The lower voices are in sub-octave treble clef. However, the numeral eight below the G clef is not shown.

working on the piece. The Score-based file can be uploaded into Verovio to visualize the rendered score and play back the encoded music. The figures below show the general structure of a Parts-based (Figure 3-28) and a Score-based (Figure 3-29) MEI file as exported by the MP Editor.[105]

[105] The <scoreDef> and <section> elements are both valid children of <score> and <part>, and they include information about the metadata of the staves/voices (e.g., their notation style and mensuration) and their music content, respectively.

```xml
<mei xmlns="http://www.music-encoding.org/ns/mei" meiversion="5.0.0-dev">
  <meiHead/>
  <music>
    <body>
      <mdiv>
        <parts>
          <part>
            <scoreDef>
              <staffGrp>
                <staffDef n="1" lines="5" notationtype="mensural.white" label="superius"/>
              </staffGrp>
            </scoreDef>
            <section>
              <staff n="1">
                <layer n="1"> <!-- music content of 1st voice --> </layer>
              </staff>
            </section>
          </part>
          <part>
            <scoreDef>
              <staffGrp>
                <staffDef n="1" lines="5" notationtype="mensural.white" label="altus"/>
              </staffGrp>
            </scoreDef>
            <section>
              <staff n="1">
                <layer n="1"> <!-- music content of 2nd voice --> </layer>
              </staff>
            </section>
          </part>
          <part>
            <scoreDef>
              <staffGrp>
                <staffDef n="1" lines="5" notationtype="mensural.white" label="tenor"/>
              </staffGrp>
            </scoreDef>
            <section>
              <staff n="1">
                <layer n="1"> <!-- music content of 3rd voice --> </layer>
              </staff>
            </section>
          </part>
        </parts>
      </mdiv>
    </body>
```

Figure 3-28: Structure of Parts-based MEI file as exported by the MP Editor. There is a global <parts> element that contains the set of <part> elements, each of them encoding the music content for a specific part. In this example, there are three voices, each represented by a <part> element (see blue, green, and yellow dotted boxes). The voices are labelled as superius, altus, and tenor. Note that in a <part> the <scoreDef> contains a single <staffDef> element and the <section> contains also a single <staff> element.[106]

[106] See note 105 about <scoreDef> and <section>.

```xml
<mei xmlns="http://www.music-encoding.org/ns/mei" meiversion="5.0.0-dev">
  <meiHead/>
  <music>
    <body>
      <mdiv>
        <score>
          <scoreDef>
            <staffGrp>
              <staffDef n="1" lines="5" notationtype="mensural.white">
                <label>superius</label>
              </staffDef>
              <staffDef n="2" lines="5" notationtype="mensural.white">
                <label>altus</label>
              </staffDef>
              <staffDef n="3" lines="5" notationtype="mensural.white">
                <label>tenor</label>
              </staffDef>
            </staffGrp>
          </scoreDef>
          <section>
            <staff n="1">
              <layer n="1"> <!-- music content of 1st voice --> </layer>
            </staff>
            <staff n="2">
              <layer n="1"> <!-- music content of 2nd voice --> </layer>
            </staff>
            <staff n="3">
              <layer n="1"> <!-- music content of 3rd voice --> </layer>
            </staff>
          </section>
        </score>
      </mdiv>
    </body>
  </music>
</mei>
```

Figure 3-29: Structure of Score-based MEI file as exported by the MP Editor. There is a score element that contains the information for the whole score. This score file can be rendered by Verovio that lines up the notes as a score, based on the mensuration and any perfecta, imperfecta, or altera value encoded in the notes. In this example, there are also three voices indicated by the blue, green, and yellow dotted boxes (notice the three <staffDef> within <scoreDef> and the three <staff> within <section>).[107]

[107] See note 105 about <scoreDef> and <section>.

3.4 Interoperability Decisions between MuRET and MP Editor

To allow for the interoperability between MuRET and the MP Editor, needed for the last part of the workflow to obtain correct symbolic scores though OMR assistance (as presented in Figure 3-1), certain decisions were taken to make the output of MuRET compatible with the MP Editor's. These decisions involved: (1) Including a IIIF link to the manifest in the <source> element, (2) including a reference to each image as a IIIF URIs in the <graphic> element, (3) export a Parts-based MEI file, and (4) follow certain conventions regarding the encoding of elements indicating page beginnings (pb), system beginnings (sb), mensuration signs (mensur), and clefs (clef). Details of each of these conventions are provided below. While I led the efforts for the interoperability between MuRET and the MP Editor, these conventions were implemented by David Rizo and Juliette Regimbal, the respective developers of these tools.

To process a music source, the MP Editor needs the link to the IIIF manifest pointing to the images of that source. This is the link the user provides in the MP Editor's upload page to "load a manuscript" (Figure 3-23). The MEI files exported by the MP Editor have a reference to this IIIF link within the metadata they encode (in the <source> element as shown in Figure 3-30). Therefore, when uploading an MEI file compatible with the MP Editor using the "load an .mei file" function on the upload page (Figure 3-23), the file is expected to include the IIIF link in the <source> metadata element following the conventions shown in Figure 3-30.

```xml
<mei>
  <meiHead>
    <fileDesc>
      …
      <sourceDesc>
        <source target="http://[IIIF-server]/[work-name]/manifest.json" targettype="IIIF"></source>
      </sourceDesc>
    </fileDesc>
  </meiHead>
  <music/>
</mei>
```

Figure 3-30: Code snippet of the metadata (<meiHead> element) of an MEI file exported by the MP Editor. It shows that the reference to the IIIF manifest is stored in the <source> element, with the @targettype=IIIF and the @target attribute pointing to the IIIF URI. Based on a figure from Desmond et al. (2021)

Therefore, to provide a compatible MEI file to the MP Editor, MuRET uses IIIF for storing and serving images. When importing a set of images to process in MuRET, the images are uploaded into a IIIF-compliant system called Cantaloupe.[108] After finishing the OMR process and exporting the MEI file encoding the recognized symbols, MuRET generates and stores a IIIF manifest. It encodes the IIIF URI in the @target attribute of the <source> element and uses the attribute @targettype=IIIF to identify it as IIIF, according to the MP Editor's conventions presented in Figure 3-30.

[108] Cantaloupe is an open-source dynamic image server, see https://cantaloupe-project.github.io (accessed 7 January 2022).

Also following the MP Editor's conventions, MuRET links the individual images using a IIIF URI in the @target attribute of the <graphic> element, which is part of the facsimile module that includes information about images and regions in those images (see Figure 3-31). The syntax followed by MuRET to encode the IIIF URI is "[IIIF-server]/[work-name]/canvas/f[folio-number]" (Desmond et al. 2021).

```
<mei>
   <meiHead/>
   <music>
      <facsimile>
         <surface xml:id="image_3058" ulx="0" uly="0" lrx="1335" lry="2000">
            <graphic target="https://muret.dlsi.ua.es/iiif/6:223/canvas/f29 " xml:id="graphic_155981"/>
            <zone xml:id="page_3023" ulx="0" uly="0" lrx="1335" lry="2000" type="page"
label="Page #1"/>
            <zone xml:id=" region_26180" ulx="320" uly="138" lrx="1216" lry="309" type="region"
label="staff"/>
            <zone xml:id="symbol_132619" ulx="320" uly="138" lrx="394" lry="309" type="symbol"
label="clef.G:L2"/>
         </surface>
      </facsimile>
   </music>
</mei>
```

Figure 3-31: Facsimile element with the <graphic> elements including the reference to the IIIF URI for an image, as exported by MuRET to work as input for the MP Editor.

In addition to including a reference to a manifest (in the metadata <source> element) and the images of the manifest related to an encoded piece (in the facsimile <graphic> elements), the MP-Editor-compatible file exported by MuRET also encodes the music as a Parts-based MEI file following the structure shown in Figure 3-28. All <part> elements follow certain rules regarding the encoding of structural elements, such as page and system beginnings, and the encoding of clef, key signature, and mensuration symbols of the part as illustrated at the end in Figure 3-34. The next three bullet points present these rules:

- **Page-begin and system-begin elements (<pb> and <sb>).** Page-begin elements are included at the beginning of each part and at every page turn. They do not indicate a physical page begin—otherwise, they would be included only in the top voice (see Figure 3-32)—but rather the beginning of the part in that page—therefore, they are included in both voices of the page (see Figure 3-32).

The <sb> element indicates a system begin (see the systems shown in Figure 3-32). In the MEI file, these elements precede the musical content of that system, and the first system begin in a page for a given voice is encoded by a <sb> following the <pb> element (see the gray highlighted elements in Figure 3-34).

Figure 3-32: One page with two parts, shown in red rectangles. Even though only the part at the top starts in the physical beginning of the page, both parts include a page-begin <pb> element since this element is encoding the start of the part in this new page. The two parts have a different number of systems, the top voice has five systems and bottom one has four; the beginning of each of these systems is represented by a <sb> element. This figure can be found in Desmond et al. (2021).

- **Clef and key signature elements (<clef> and <keySig>).** Clefs and key signatures are physically written at the beginning of each system (see Figure 3-33); therefore, the elements of <clef> and <keySig> (if a key is provided) are placed after each <sb> element in the <layer> element, which contains the stream of notes in the piece (see the light-blue highlighted elements in Figure 3-34).

- **Mensuration elements (<mensur>)**. Opposite to clefs (and key signatures), the mensuration sign appears only once at the beginning of a voice (see Figure 3-33). Therefore, instead of encoding the mensuration in the stream of notes inside <layer>, the <mensur> element is encoded inside <staffDef> (see the pink highlighted element in Figure 3-34), which contains the staff meta-information (e.g., the number of staff lines and the notation type). While the main mensuration is encoded by <mensur> within the <staffDef> element, any change in mensuration is encoded by a <mensur> element within the <layer>.

Figure 3-33: Top part (voice) in Figure 3-32. Each system starts with a clef (light blue circles), while the mensuration sign is only included at the beginning of the voice (pink circle). This figure can be found in Desmond et al. (2021).

```xml
<part>
  <scoreDef>
    <staffGrp>
      <staffDef n="1">
        <mensur/>
      </staffDef>
    </staffGrp>
  </scoreDef>
  <section>
    <staff n="1">
      <layer n="1">
        <pb/>
        <sb/>
        <clef/>
        <!-- notes in the 1st system of the voice -->
        <sb/>
        <clef/>
        <!-- notes in the 2nd system of the voice -->
      </layer>
    </staff>
  </section>
</part>
```

Figure 3-34: Encoding within one <part>, with its two children <scoreDef> and <section> for meta-information and music content, respectively. The <pb> indicates the beginning of a page for that part (dark gray), and <sb> indicates the beginning of a system for the given part within that page (light gray). Each <sb> element is followed by the <clef> (light blue) that is written at the beginning of the system, and each clef contains a @facs attribute pointing to the id of the zone where it appears in the image (see Figure 3-33). The <mensur> element is encoded in the <staffDef> (pink), and it also includes a @facs attribute pointing to the region where the sign is in the image (see Figure 3-33).

3.5 Integration of Humlib's Dissonance Filter in the MP Editor

This section presents the humlib Dissonance Filter (DF) in more detail and the work of integrating this tool within the MP Editor. It starts by restating the goal behind the integration of the DF tool in the MP Editor and the reason why I chose this particular tool over the other computational music analysis options. Then, Section 3.5.1 provides the list of Renaissance dissonance labels identifiable by the DF and presents an example that includes a couple of the most common dissonance types. Section 3.5.2 provides the details of the work done to include the DF into the MP Editor, the definition of legal and illegal dissonances labels, and how these are marked in different colors in the MP Editor. Finally, Section 3.5.3 summarizes the implications (or benefits) of this work on software development and encoding of mensural notation.

Similar to how the user of MuRET verifies and corrects the output of the OMR stage, in the MP Editor the user can also verify and correct the results of the automatic voice alignment. The voice alignment might be incorrect due to different factors: an OMR error that was not caught before, a scribal error (i.e., an error in the original source), or an error in the automatic voice alignment algorithm itself. The MP Editor facilitates the detection of voice-alignment errors already by (1) presenting the piece lined-up in score format, permitting the user to visualize the vertical sonorities that were not visible in the original separate-parts format; (2) allowing to bar the piece by different note values; and (3) allowing to switch the piece to modern clefs to ease reading. All these actions facilitate looking at the counterpoint of the piece and finding errors.

In this dissertation, I am testing whether explicitly showing counterpoint errors (in particular, illegal dissonances) to the user facilitates the correction task even further, measuring the time invested on error correction with and without these counterpoint errors markers. For this purpose, from the computational music analysis tools presented in Section 2.4.3, I decided on using Humdrum, specifically the *humlib* library. Only Humdrum and the VIS Framework have out-of-the-box counterpoint tools. Both contain a "Renaissance dissonance labelling" tool developed by Alex Morgan. The tool is known as the *dissonance indexer* in the VIS Framework, which was developed during Morgan's PhD with the help of Christopher Antila (VIS Framework developer); and its humlib's counterpart is known as the *dissonance filter,* which was developed by Morgan during his postdoc with the help of Craig Sapp (humlib's developer) as part of the Josquin Research Project analysis tools. There were two reasons why I chose the Humdrum implementation of the dissonance labeler than its VIS Framework implementation. First, humlib's dissonance filter includes more dissonance types than the VIS Framework dissonance indexer, therefore, being more precise in its dissonance classification.[109] Second, humlib's dissonance filter is available in Verovio Humdrum Viewer (VHV), allowing to render the dissonance labels with the notes they relate to by using Verovio as the render engine, as will be shown in Section 3.5.1 that presents the details about the filter. Nevertheless, the integration of humlib's dissonance filter into the MP Editor was not a straight-forward process, as will be detailed in Section 3.5.2.

3.5.1 The Dissonance Filter (DF)

Table 3-5 shows the dissonance labels available in humlib's dissonance filter and their definitions. The dissonance filter (DF) works on **kern, the Humdrum format for Common Western Music Notation (CWMN). As indicated before, these dissonance labels can be rendered in Verovio close to the dissonant note, as shown in Figure 3-35. This figure shows a screenshot of the Verovio Humdrum Viewer (VHV) for a fifteenth-century piece of the Josquin Research Project (Ockeghem's chanson Ma maistresse). The VHV webapp shows the **kern file encoding the piece in CWMN on the left panel, and the score rendered by Verovio in the right panel. The Humdrum file (on the left) contains three sets of columns corresponding to the three voices of the piece. The columns (called spines in Humdrum) that contain the music events (notes and rests) in the voices are indicated by the header **kern. The **cdata spine next to each **kern column indicates the dissonance labels for that particular note. Figure 3-35 shows examples of three dissonance labels: the passing tone (p), and the suspension (s) and agent (g). I will use this example to explain some of these common dissonance types.

[109] See https://vis-framework.readthedocs.io/en/v3.0.5/_modules/vis/analyzers/indexers/dissonance.html for the dissonance types available in the VIS Framework dissonance indexer and see https://doc.verovio.humdrum.org/filter/dissonant/#dissonant-function-labels for the dissonance types in humlib's dissonance filter.

Table 3-5: Dissonance function labels as defined in humlib (https://doc.verovio.humdrum.org/filter/dissonant/).

Label	Meaning	Label	Meaning
P	rising passing tone	p	downward passing tone
N	upper neighbor	n	lower neighbor
D	double neighbor upper then lower	d	double neighbor lower then upper
E	upper échappée	e	lower échappée
C	ascending short nota cambiata	c	descending short nota cambiata
K	ascending long nota cambiata	k	descending long nota cambiata
A	rising anticipation	a	descending anticipation
I	reverse ascending nota cambiata	i	reverse descending nota cambiata
J	reverse upper échappée	j	reverse lower échappée
S	ternary suspension	s	binary suspension
G	ternary suspension agent	g	binary suspension agent
F	fake susp. Approached by step up	f	fake susp. Approached by step down
x	note of resolution of a suspension against the suspension's dissonance	r	suspension repeated note
M	suspension missing agent approached by step up	m	suspension missing agent approached by step down
o	purely ornamental suspension	h	chanson idiom
Q	dissonant 3rd quarter rising passing tone	q	dissonant 3rd quarter falling passing tone
B	dissonant 3rd quarter upper neighbor	b	dissonant 3rd quarter lower neighbor
T	appoggiatura approached from below	t	appoggiatura approached from above
V	ascending accented passing tone	v	descending accented passing tone
W	accented upper neighbor	w	accented lower neighbor
Y	only dissonant against known dissonance asc.	Y	only dissonant against known dissonance desc.
Z	unclassified dissonance, 2nd or 7th interval	z	unclassified dissonance, 4th interval
L	unclassified dissonance in parallel accompaniment to an identifiable dissonance, moving by step up	l	unclassified dissonance in parallel accompaniment to an identifiable dissonance, moving by step down

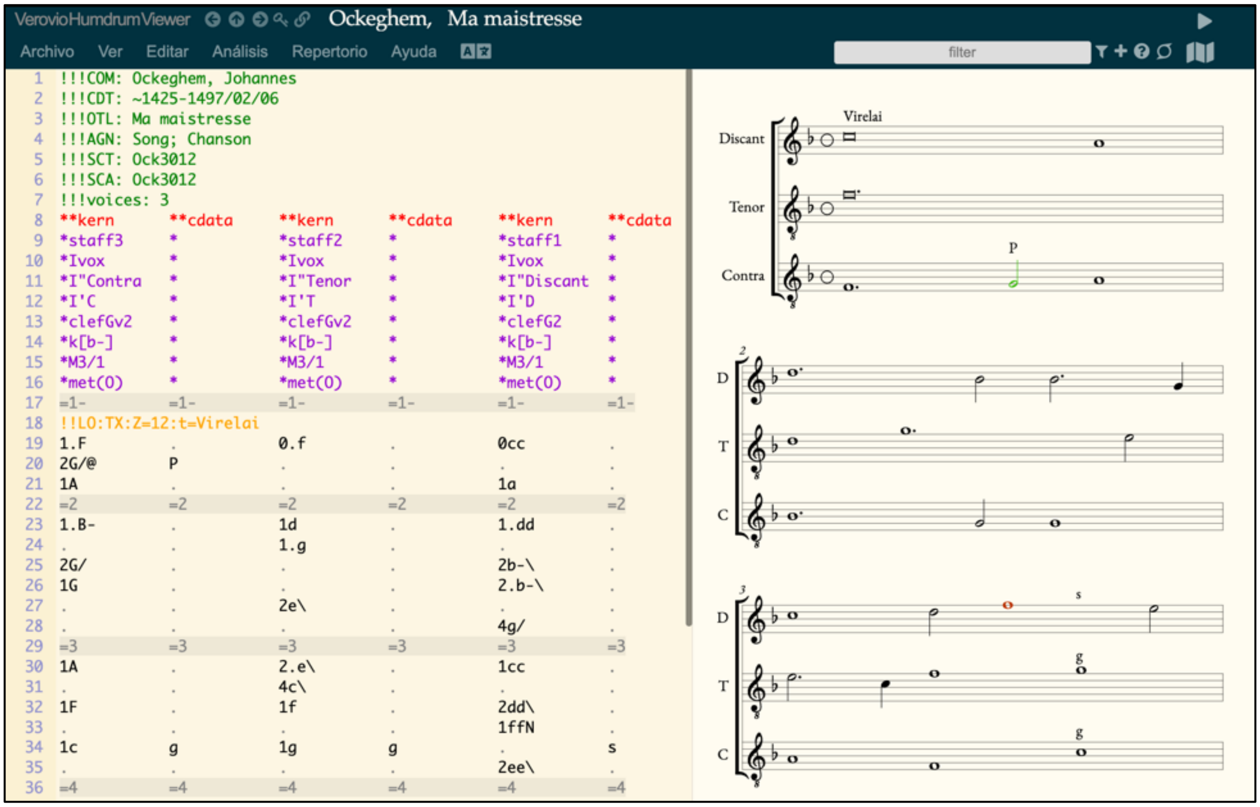


*Figure 3-35: Use of the dissonance filter in Verovio Humdrum Viewer. The left panel shows the encoding of the piece in **kern (Humdrum format for CWMN), and the right side renders the piece using Verovio. The dissonance labels are shown in both panels, in the left panel through the **cdata column, and on the right panel they are shown as labels (e.g., 'P' for ascending passing tone, 's' for suspension, and 'g' for the agent of the suspension) colored according to their metric position.*

Figure 3-35 illustrates two types of dissonances: passing tones and suspensions, both of which are common in Renaissance and later styles. Therefore, I will use the example presented on this Figure 3-35 to define and illustrate these two dissonances. A **passing tone** is a dissonance that is approached by a step and left by step, with both steps happening in the same direction, as can be seen by the note marked in green with a 'P' label on top in measure one—the 'P' is a capital letter to indicate that this is an ascending passing tone. Here, the contratenor's G is dissonant against the tenor's F as it forms a second, and it is also dissonant against the discant's C as the C forms a fourth with the lowest voice—a sonority that is considered dissonant in the Renaissance, even though it might be deemed as consonant on later periods. This dissonant G is an (ascending) passing tone, moving by ascending steps from F to A.

The other example of a dissonant note is found in measure three. There are two types of labels at this point, an ‘s’ and a ‘g,’ that are marking a suspension. “A **suspension** consists of a voice that becomes dissonant either by sustaining or restriking a note before resolving the dissonance down by step. This sustained voice was referred to as the patient by Artusi, and it gets an S or s label at the moment of the dissonance. Another voice strikes a note that is dissonant against the suspended note, and this voice gets a G or g label for agent.”[110] Looking at the example in measure three of Figure 3-35, the F in the Discant is consonant with the latter half of the tenor’s F. However, when the tenor changes to a G, the “suspended” F in the discant becomes dissonant. This dissonance is resolved by a step down to the half-note E at the end of the measure. The discant’s F is then classified as a suspension (with the ‘s’ label), and the tenor’s E that causes the suspended note to form a dissonance is classified as the agent of the suspension (with a ‘g’ label). In this case, there are two agents for the suspension, as there is a second ‘g’ label on the contra’s C.

3.5.2 The DF in the MP Editor: Integration Process and Distinction of Legal vs. Illegal Dissonances

As was indicated before, the integration of humlib’s dissonance filter into the MP Editor was not a straight-forward process. The dissonance filter works only for **kern, Humdrum format for CWMN, while the output of the MP Editor is a Mensural MEI file. Craig Sapp (humlib’s developer) and I designed and implemented a process to transform the Mensural MEI file output of the MP Editor into **kern to be able to use the dissonance filter and retrieve the dissonance labels of the mensural score; and, with the help of Juliette Regimbal (MP Editor’s lead developer),

[110] https://doc.verovio.humdrum.org/filter/dissonant/ - suspensions-s-s-g-g

these labels were displayed in the MP Editor. Figure 3-36 presents the process of converting from Mensural MEI to **kern and including the dissonance labels into the MP Editor.

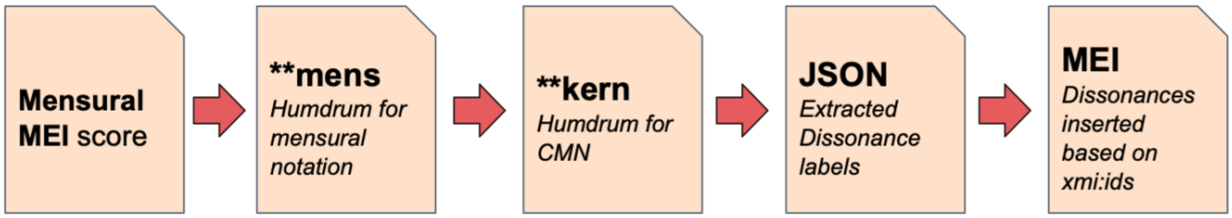


*Figure 3-36: Conversion process starting with the output of the MP Editor, a Mensural MEI file. This file was converted into **mens (humdrum for mensural notation) and then **kern (humdrum for CWMN) to retrieve the dissonance labels. This figure can be found in Desmond et al. (2021).*

The Score-based Mensural MEI output of the MP Editor was converted into **mens, which was then converted into **kern, with the IDs from the MEI notes (which are encoded in the attribute @xml:id of the <note> element) being preserved along the conversion process. Once in **kern, humlib's dissonance filter was used. See Figure 3-37 for an example of the output of this step for a three-voice piece. Each group of three columns represents a voice, with the **kern spine representing the music events (notes and rests), the **xmlid spine representing the @xml:id value of the event in the MEI file, and the **cdata column representing the dissonance label for that sonority according to the dissonance filter. In the next step, the dissonances (in the **cdata column) and the xml:ids of the notes related to those dissonances (in the **xmlid column) were extracted and encoded in a JSON list, as the one shown in Figure 3-38. The JSON list of notes' ids and related dissonances was then used to display the dissonance labels in the MP Editor by introducing them as lyrics of the corresponding notes in the rendered MEI file as will be seen later in Figure 3-39.

```
!!!OTL: 6 Missa sobre las voces
!!!system-decoration: [(s1,s2,s3)]
**kern      **xmlid     **cdata   **kern      **xmlid     **cdata   **kern      **xmlid     **cdata
*part3      *part3      *         *part2      *part2      *         *part1      *part1      *
*staff3     *staff3     *         *staff2     *staff2     *         *staff1     *staff1     *
*I"tenor    *           *         *I"altus    *           *         *I"superius *           *
*clefC3     *           *         *clefC2     *           *         *clefG2     *           *
*met()_0222 *           *         *met()_0222 *           *         *met()_0222 *           *
*MM600      *           *         *MM600      *           *         *MM600      *           *
1G          PART3_A12   .         00r         .           .         00r         .           .
2G          PART3_A15   .         .           .           .         .           .           .
2A          PART3_A16   .         .           .           .         .           .           .
=           .           =         .           .           =         .           .           =
2B          PART3_A17   .         .           .           .         .           .           .
2c          PART3_A18   .         .           .           .         .           .           .
1d          PART3_A19   .         .           .           .         .           .           .
=           =           =         =           =           =         =           =           =
1.e         PART3_A21   .         00r         .           .         1g          PART0_A20   .
.           .           .         .           .           .         2g          PART0_A23   .
4d          PART3_A24   .         .           .           .         2a          PART0_A25   .
4c          PART3_A26   .         .           .           .         .           .           .
=           =           =         =           =           =         =           =           =
2d          PART3_A28   .         .           .           .         2b          PART0_A27   .
2e          PART3_A30   .         .           .           .         2cc         PART0_A29   .
2f          PART3_A31   .         .           .           .         1dd         PART0_A32   .
2.g         PART3_A33   .         .           .           .         .           .           .
=           =           =         =           =           =         =           =           =
.           .           .         1c          PART6_A35   .         2ee         PART0_A34   .
4f          PART3_A36   p         .           .           .         .           .           .
=           .           =         .           .           =         .           .           =
4e          PART3_A37   .         .           .           .         1gg         PART0_A38   .
4d          PART3_A39   n         .           .           .         .           .           .
2e          PART3_A40   .         2c          PART6_A41   .         .           .           .
2d          PART3_A43   .         2d          PART6_A44   .         2ff         PART0_A42   .
=           =           =         =           =           =         =           =           =
2c          PART3_A46   .         2e          PART6_A47   .         2ee         PART0_A45   .
2d          PART3_A48   .         2f          PART6_A49   .         1dd         PART0_A50   .
1B          PART3_A51   .         1g          PART6_A52   .         .           .           .
.           .           .         .           .           .         4cc         PART0_A53   v
.           .           .         .           .           .         4b          PART0_A54   .
=           =           =         =           =           =         =           =           =
0A          PART3_A56   .         0a          PART6_A57   .         2cc         PART0_A55   .
=           =           =         =           =           =         =           =           =
.           .           .         .           .           .         2ee         PART0_A58   .
.           .           .         .           .           .         2ee         PART0_A59   .
.           .           .         .           .           .         2ee         PART0_A60   .
=           =           =         =           =           =         =           =           =
1r          .           .         2r          .           .         2a          PART0_A61   .
.           .           .         2c          PART6_A64   .         1ee         PART0_A65   .
2r          .           .         2c          PART6_A67   .         .           .           .
2e          PART3_A69   .         2d          PART6_A70   V         2dd         PART0_A68   v
=           =           =         =           =           =         =           =           =
2e          PART3_A72   .         2.e         PART6_A73   .         2cc         PART0_A71   .
2e          PART3_A75   .         .           .           .         4b          PART0_A74   .
.           .           .         4f          PART6_A77   P         4a          PART0_A76   n
2A          PART3_A79   Z         1g          PART6_A80   .         2b          PART0_A78   Z
1d          PART3_A82   z         .           .           .         2ee         PART0_A81   Z
=           =           =         =           =           =         =           =           =
```

*Figure 3-37: Result of the conversion from **mens to **kern and the use of the dissonance filter for a three-voice piece. Each group of three adjacent columns provide information for a different voice, orange for tenor, green for altus, and blue for superius. The **kern column contains the notes (and rests) in CMN values of the voice, the **xml:id contains the @xml:id values corresponding to each of these notes in the original MEI file, and the **cdata column contains the dissonance labels corresponding to those notes.*

```
{
        "PART3_A36": "p",
        "PART3_A39": "n",
        "PART0_A53": "v",
        "PART6_A70": "V",
        "PART0_A68": "v",
        "PART6_A77": "P",
        "PART0_A76": "n",
        "PART3_A79": "Z",
        "PART0_A78": "Z",
        "PART3_A82": "z",
        "PART0_A81": "Z",
        "PART0_A83": "T",
        "PART3_A85": "v",
        "PART3_A87": "n",
        "PART3_A91": "V",
        "PART3_A94": "P",
        "PART6_A101": "n",
        "PART0_A100": "p",
        "PART0_A105": "Z",
        "PART0_A106": "V",
        …
}
```

*Figure 3-38: JSON list showing the mapping of xml:ids and corresponding dissonance labels retrieved from the **kern file shown in Figure 3-37.*

The JSON list of note's ids (corresponding to the @xml:id values) and dissonance labels allows to show the labels under the appropriate notes in the MP Editor, as can be seen in Figure 3-39. The labels are shown in two colors: blue for legal dissonances (no matter how common or uncommon these might be) and orange for illegal dissonances. From the list of dissonances in Table 3-5, the ones I considered as illegal dissonances are the dissonances that do not fall in a "known" category (e.g., passing tones, lower/upper neighbors, suspensions, appoggiaturas, cambiatas, and échappées). These "illegal dissonances" are:

- Z or z: dissonances not assignable to any of the other categories and, therefore, its function is considered "unknown"
- L or l: an unidentifiable dissonance like Z, but that is in parallel motion with an identifiable dissonance

- Y or y: a note that is only dissonant against known dissonance types
- x: resolution against suspension dissonance; this is, when a voice sounds the note of resolution of a suspension in a descending line against the dissonant portion of the suspension in another voice

For examples on any of these categories, please consult the VHV page on the dissonant filter.[111]

Figure 3-39: Dissonance labels displayed in the MP Editor, with blue indicating legal dissonances and orange indicating illegal dissonances.

3.5.3 Contributions of the Integration of Humlib's DF into the MP Editor

The implementation of the conversion process introduced in the previous section (and outlined in Figure 3-36) had beneficial implications:

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

1. Improvement of **mens. I proposed a few ways to encode certain mensural notation features that were missing in **mens (and are already supported in MEI).[112] These features included adding a token to represent alterations, adding a way to encode the semantic aspects of a mensuration sign, and adding a token for showing coloration (this last one was suggested by Rizo). Humlib's developer, Craig Sapp, implemented these features in **mens. The improvement of **mens, in turn, improves MuRET's and the MP Editor's support for mensural notation, as **mens is used to encode the semantic transcription in MuRET and the music entered by the user for each system in the MP Editor's input interface.
2. Sapp and I corrected a few issues in the conversion process from **mens to MEI: the Humdrum tokens representing perfect and imperfect notes were not converted correctly into MEI, and the encoding of dotted notes in MEI was invalid according to the mensural module.
3. The amplification of the mei2hum conversion tool to allow for conversion of Mensural MEI to **mens (and not only from CMN MEI to **kern).[113] This was all implemented by Sapp.
4. The creation of the mens2kern conversion tool.[114] I implemented the core of this tool that converted mensural note values into CMN, while Sapp implemented the structure of the code.

The rest of the process, extracting a JSON list with the information and then using this to add the labels as lyrics into the rendered MEI file in the MP Editor, was designed by Sapp and he gave the explicit instructions to the MP Editor's developer, Regimbal, to implement this process.

[112] I did this always in consultancy with David Rizo (MuRET's developer who is using **mens for MuRET's semantic transcription).
[113] https://github.com/craigsapp/humlib/blob/master/src/tool-mei2hum.cpp
[114] https://github.com/craigsapp/humlib/blob/master/src/tool-mens2kern.cpp

Chapter 4 Methodology – Part 2: The Dissonance Filter Experiment in the MP Editor

The previous chapter presented the tools chosen for the digitization and encoding workflow and the work done for their interoperability. A summary of this workflow and interaction of these tools to obtain the symbolic scores with editorial corrections is shown in Figure 4-1. After presenting this encoding methodology in the previous chapter, this current chapter focuses on the experiment's methodology.

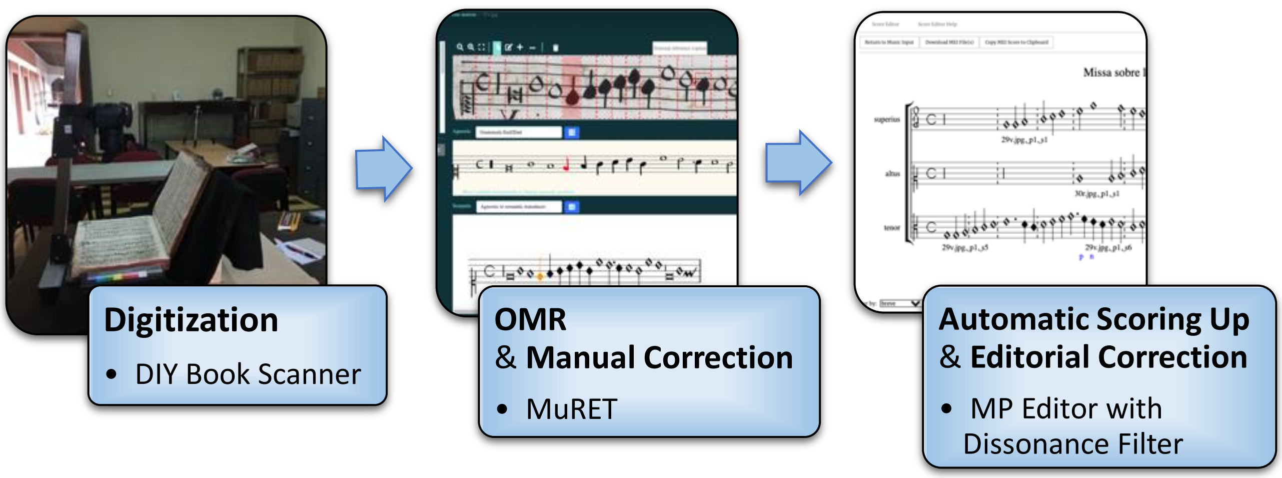


Figure 4-1: Digitization and encoding workflow with the tools involved in each stage.

The experiment consists of evaluating if the use of humlib's dissonance filter (DF) reduces the time spent in the last part of the encoding workflow. In other words, I am evaluating whether the use of counterpoint error markers showing illegal dissonances facilitates the correction of the automatic voice-alignment process in the MP Editor's score interface, where the piece is already shown lined-up as a score that can be barred by different note values and rendered in modern clefs. At this point in the workflow, I expect the score to have very few OMR errors, as they have been manually corrected in the previous stage using MuRET's specialized interface.[115] Therefore, I

[115] OMR frameworks tend to have specialized interfaces to correct the results of each of the OMR steps. As an example, Aruspix and Neon have interfaces that allow to render the recognized music symbols on top of the image to facilitate their correction. MuRET allows for similar ease in the correction process through its document analysis

expect for most of the corrections in this stage to be editorial corrections (i.e., corrections for scribal errors). Scribal errors are hard to identify and, while presenting the piece as a score facilitates this task, I want to evaluate if this highlighting of illegal dissonances can facilitate it further.

To evaluate this, I chose a random set of pieces from GuatC 1, divided it into two datasets, one using the dissonance filter (D) and one without it (ND), and determined the average time spent in correcting the voice alignment in both D and ND datasets. For the correction time to be comparable among the D and ND datasets, I distributed the pieces so that the two datasets had approximately the same average number of measures, voices, and illegal dissonances. The experiment was conducted by Geneviève Gates-Panneton, a Bachelor of Music (B.Mus.) with a major in voice and a minor in early music. Gates-Panneton's background includes a paleography course—where she studied how to read early music notation, including mensural notation—and a Renaissance musicianship course. She is also an early music singer, being part of the Ensemble Scholastica—a vocal ensemble that specializes in the performance of medieval plainchant and polyphony.

The following sections contain details about the experiment, with Section 4.1 focusing on data preparation and Section 4.2 focusing on the experimenter's tasks and method for analyzing the results. The results of the experiment (these are, the correction time per piece and resulting files) and the discussion of these results will be presented in the third section of Chapter Chapter 5 (Section 5.3).

---

interface (see Figure 3-17), agnostic transcription panel (see Figure 3-19), and semantic transcription panel (see Figure 3-20) as presented in Section 3.2.

4.1 Data Preparation

Before starting the experiment, I used one mass and one short polyphonic piece to train the experimenter (Geneviève Gates-Panneton) on using the MP Editor interface and the dissonance filter integrated into it. These pieces were items 5 and 6 from Table 2-2, corresponding to an *Asperges me* by Alonso de Trujillo and the *Missa sobre las vozes* by Cristóbal de Morales. For the actual experiment, I randomly selected sixteen pieces from the manuscript, where each piece was either a mass movement or a short polyphonic piece. The sixteen pieces, which represent around 20% of the total corpus, are shown in Table 4-1. To ease the editorial work in the MP Editor, I subdivided the long mass movements into smaller units. Here, I am considering as a "long movement" anything that has more than 2 folios. I divided these long movements into what I call self-contained units. A "self-contained unit" is a section or group of consecutive sections whose ending coincides with a page turn. For example, Figure 4-2 shows a single self-contained unit distributed over three folios (97v–100r). This self-contained unit starts with the section "Et in terra pax" (shown in red) on f.97v–98r and ends in the following pair of folios 98v–99r. However, its ending does not coincide with the page turn; instead, the new section "Qui tollis peccata mundi" (shown in blue) starts right in that pair of folios and continues into the next pair f.99v–110r, where its ending coincides with the end of the page. Therefore, those two sections together constitute what I called a self-contained unit. The reason why I work with these self-contained units rather than single sections is because when the end of a section does not coincide with the end of a page, as shown in Figure 4-2, MuRET cannot export the section into an MEI file.

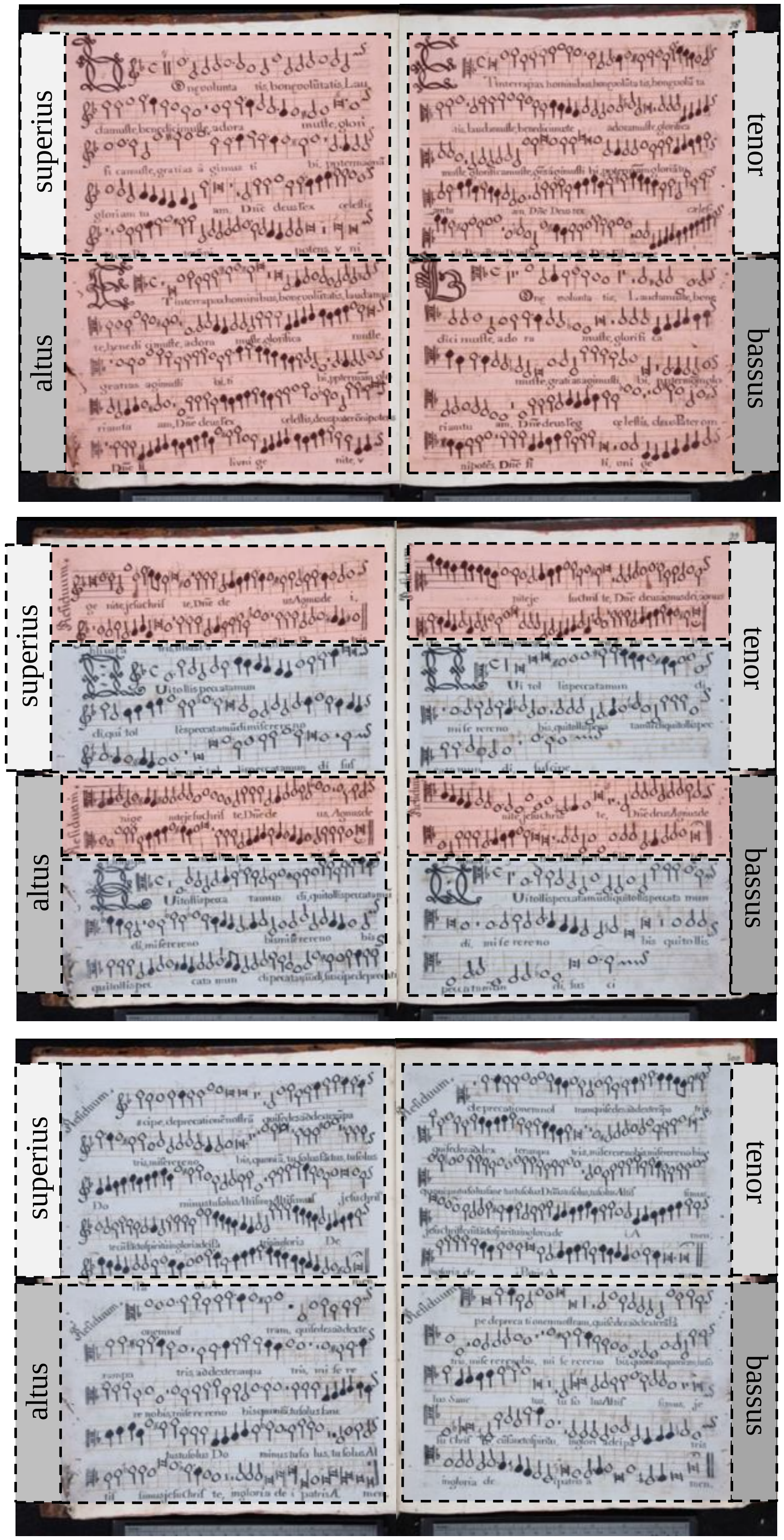


Figure 4-2: Example of a self-contained unit in folios 97v–100r (from top to bottom). This self-contained unit consists of two consecutive sections. The first section is "Et in terra pax" in folios 97v–99r (red box) and the second section is "Qui tollis peccata mundi" in f.98v–100r (blue boxes).

Table 4-1: Pieces randomly selected, their subdivision in self-contained units (i.e., section or group of sections whose ending coincide with a page turn), and data of these subdivisions.

Mass Movement or Short Piece	Folios	No. Folios	Self-contained Units	Folios	Number of: Folios	Measures	Voices	Illegal Disson.
Sanctus, Missa 7	38v-41r	3	Sanctus	38v-39r	1	29	4	0
			Benedictus & Osanna	39v-41r	2	50	4	52
Sanctus, Missa 8	52v-53r	1	Full movement	52v-53r	1	29	4	9
Gloria, Missa 9	56v-60r	4	Et in terra pax	56v-58r	2	44	4	46
			Qui tollis peccata	58v-60r	2	37	4	8
Agnus II, Missa 9	70v-71r	1	Full movement	70v-71r	1	35	5	14
Kyrie, Missa 10	71v-73r	2	Full movement	71v-73r	2	71	4	4
Credo, Missa 10	76v-81r	5	Patrem omnipotentem & Et incarnatus est	76v-78r	2	67	4	64
			Crucifixus	78v-79r	1	40	4	3
			Et in spiritum sanctum	79v-81r	2	64	4	17
Agnus I, Missa 10	83v-84r	1	Full movement	83v-84r	1	23	4	2
Gloria, Missa 12	97v-100r	3	Full movement	97v-100r	3	136	4	-
Agnus, Missa 13	118v-119r	1	Full movement	118v-119r	1	30	4	6
Kyrie, Missa 14	119v-121r	2	Full movement	119v-121r	2	56	4	23
Kyrie, Missa 15	133v-135r	2	Full movement	133v-135r	2	44	4	0
Credo, Missa 15	139v-146r	7	Patrem omnipotentem	139v-141r	2	64	4	61
			Et incarnatus est	141v-143r	2	56	4	43
			Et ascendit in celum	143v-144r	1	48	3	0
			Et in spiritum sanctum	144v-146r	2	-	4	-
Sanctus, Missa 16	165v-168r	3	Sanctus	165v-166r	1	22	4	2
			Osanna	166v-167r	1	13	4	17
			Benedictus	167v-168r	1	27	3	16
Agnus, Missa 17	179v-180r	1	Full movement	179v-180r	1	26	4	1
Surrexit, Piece 21	184v-186r	2	Full piece	184v-186r	2	17	5	21
Surrexit, Piece 25	189v-190r	1	Full piece	189v-190r	1	15	4	13

This division of mass movements resulted on twenty-five self-contained units as shown in Table 4-1. I had to remove two pieces from the dataset (shown in pink): the "Missa 12 - *Gloria* (full movement)" because it was an outlier in terms of number of measures, and the "Et in spiritum sanctum" section from "Missa 15 - *Credo*" due to an issue in Verovio regarding changes in mensuration that does not allow the MP Editor to render this piece correctly.[116] This left a total of twenty-three pieces for the experiment. I divided this set of twenty-three pieces into two datasets, one where the dissonance filter (DF) will be used to correct the voice alignment in the MP Editor (dataset D) and one where the dissonance filter would not be used (dataset ND). For a fair comparison of the results between the two datasets, the division of the pieces into these two datasets was done in a way that both D and ND would have approximately the same number of measures (in average), voice distribution, and "illegal" dissonances. I also tried to distribute the pieces so that I had around the same number of pieces in each dataset with a CPDL transcription. This division is illustrated in Table 4-2 and Table 4-3.

[116] This issue was solved after the experiment in the release version-3.10.0, specifically by commit *20a61b5* on 14 March 2022: https://github.com/rism-digital/verovio/commit/20a61b5343ed091348390b13c2a101b825f0bf31.

Table 4-2: Dataset ND, with no dissonance filter used in the correction of the voice alignment in the MP Editor. The "self-contained units" column shows the filename of the self-contained unit and the sections contained in that unit.

Self-contained Units	**Folios**	**Number of**			**CPDL Transcription**
		Measures	**Voices**	**Illegal Disson.**	
Missa8.4_Sanctus0 Missa 8, Sanctus (Full)	52v-53r	29	4	9	✓
Missa9.2_Gloria1 Missa 9, Gloria ("Et in terra pax")	56v-58r	44	4	46	✓
Missa9.2_Gloria2 Missa 9, Gloria ("Qui tollis peccata")	58v-60r	37	4	8	✓
Missa10.1_Kyrie0 Missa 10, Kyrie (Full)	71v-73r	71	4	4	✓ Full Kyrie movement (71v-72r). ✗ Extra Kyrie at the end (72v-73r)
Missa10.3_Credo1 Missa 10, Credo ("Patrem omnipotentem" & "Et incarnatus est")	76v-78r	67	4	64	✓
Missa10.5_AgnusI0 Missa 10, Agnus I (Full)	83v-84r	23	4	2	✓
Missa15.3_Credo2 Missa 15, Credo ("Et incarnatus est")	141v-143r	56	4	43	✓
Missa15.3_Credo3 Missa 15, Credo (Et ascendit in celum)	143v-144r	48	3	0	✓
Missa16.4_Sanctus2 Missa 16, Sanctus (Osanna)	166v-167r	13	4	17	✓
Missa17.5_Agnus0 Missa 17, Agnus (Full)	179v-180r	26	4	1	✓
Piece21_Surrexit0 Piece 21, Surrexit (Full)	184v-186r	17	5	21	✗
	Average	39.2	4	19.5	

Table 4-3: Dataset D, with the dissonance filter used in the correction of the voice alignment in the MP Editor. The "self-contained units" column shows the filename of the self-contained unit and the sections contained in that unit.

Self-contained Units	**Folios**	**Number of**			**CPDL Transcription**
		Measures	**Voices**	**Illegal Disson.**	
Missa7.4_Sanctus1 Missa 7, Sanctus (Sanctus)	38v-39r	29	4	0	✘
Missa7.4_Sanctus2 Missa 7, Sanctus (Benedictus & Osanna)	39v-41r	50	4	52	✘
Missa9.6_AgnusII0 Missa 9, Agnus II (Full)	70v-71r	35	5	14	✓
Missa10.3_Credo2 Missa 10, Credo (Crucifixus)	78v-79r	40	4	3	✓
Missa10.3_Credo3 Missa 10, Credo (Et in spiritum sanctum)	79v-81r	64	4	17	✓
Missa13.5_Agnus0 Missa 13, Agnus (Full)	118v-119r	30	4	6	✓
Missa14.1_Kyrie0 Missa 14, Kyrie (Full)	119v-121r	56	4	23	✓
Missa15.1_Kyrie0 Missa 15, Kyrie (Full)	133v-135r	44	4	0	✓
Missa15.3_Credo1 Missa 15, Credo (Patrem omnipotentem)	139v-141r	64	4	61	✓
Missa16.4_Sanctus1 Missa 16, Sanctus (Sanctus)	165v-166r	22	4	2	✓
Missa16.4_Sanctus3 Missa 16, Sanctus (Benedictus)	167v-168r	27	3	16	✓
Piece25_Surrexit0 Piece 25, Surrexit (Full)	189v-190r	15	4	13	✓
Average		39.7	4	17.3	

As can be seen in Table 4-2 and Table 4-3, both ND and D have a similar average number of measures (39.2 vs 39.7); both are composed of 4-voice pieces, with only one 3-voice and one 5-voice piece in each dataset; and the average number of illegal dissonances appearing in both datasets is similar, with 19.5 for ND and 17.3 for D, respectively. Finally, all pieces in the experiment have a modern transcription in CPDL except for two in each dataset.

4.2 Methodology and Evaluation Procedure

This section expands on the procedure to be followed by Gates-Panneton during the experiment—how she looked at the pieces to find the errors to correct—and the method for analyzing the results afterwards. In summary, this section provides details of the "methodology" of the experiment, which is different from the "methodology" of the digitization and music-encoding process that was presented in Chapter Chapter 3.

The experimenter's tasks consist of: (a) reporting the time invested in the correction of each piece, (b) taking notes of her general impressions after processing each piece, if she considers this to be relevant, and (c) providing the files downloaded from the MP Editor at the end of the correction process. Gates-Panneton was instructed to perform the following steps when correcting each piece: (1) look at the end of sections to see if the voices line up at the cadence, in order to determine if there is a note value missing; (2) look at cadences for the same purpose; and (3) look at the rest of the piece, which implied looking at all of the notes for the ND dataset and looking only at the notes with orange labels for the D dataset. Although most corrections at this point in the encoding workflow should be editorial corrections, I instructed the experimenter to use the regular mode of the MP Score Editor rather than the editorial mode. This was to allow her more freedom to play with changing notes, without each correction getting recorded into 'sic' and 'corr' readings within the MEI file.

To analyze the results, I need to determine where the experimenter's corrections are, and whether these are correct corrections. To facilitate this task, I wrote two auxiliary scripts: MEI2tokens and MusicXML2tokens.[117] I will use these scripts with the following MEI or

[117] The scripts can be found at https://github.com/martha-thomae/GuatC1/tree/Experiment/token_files under the names *meimensuralscore2eventlist.py* and *musicxml2eventlist.py*.

MusicXML scores as input to represent them as a list of events or tokens representing the notes in that score:

- OMR Score files. I will generate these files by uploading the OMR Parts-based MEI file into the MP Editor and exporting the score with no corrections. I will use the MEI2tokens script over these OMR Score files.
- I will also use the MEI2tokens script over the Experimenter's Score files.
- And I will use the MusicXML2tokens script over the CPDL transcriptions, which are encoded in MusicXML format.

After using these scripts over the previously mentioned files, I will be able to compare them using a diff tool.[118] Comparing the tokens of the OMR Score against the ones of the Experimenter's Score will allow me to identify the experimenter's corrections. And comparing the tokens of the Experimenter's Score against the ones of the CPDL transcription will allow me to find out whether the experimenter's corrections match those found in CPDL. While the CPDL scores cannot be considered ground truth, they do record another musician's ideas about mistakes in the sources and help in the process of checking the experimenter's corrections. Whenever I find a discrepancy between the experimenter's and the CPDL versions, I will evaluate the two scores in more detail to determine what would be the best way to handle that passage in the piece, always favoring the version that reduces the number of "true" illegal dissonances and that results in readings that incorporate repeated melodies found in other voices (imitation). I plan on consulting Professors Julie Cumming and Peter Schubert when this is difficult to judge.

[118] Diff tools allow for comparing files by displaying the differences between their contents. I used a free, online diff tool called *Diffchecker*: https://www.diffchecker.com.

To judge the experimenter's version, in addition to comparing it to the CPDL score—which, again, cannot be considered ground truth—I will also look for any orange labels left by the experimenter by re-uploading the experimenter's file into the MP Editor and activating the DF. If there are any orange labels left in the experimenter's file, I will determine whether these point to counterpoint errors missed by the experimenter, or if they are the result of "false" illegal dissonances (false positives provided by the DF). Here I will also consult Professors Cumming and Schubert when this is hard to judge.

Additionally, for the pieces in the ND dataset—where the experimenter would not use the dissonance filter (DF) during correction—I will re-upload the original OMR file into the MP Editor with the DF activated and replicate the experimenter's corrections to see if these corrections result on the elimination of illegal dissonances (i.e., reducing the number of orange labels). The two actions of checking for the orange labels left by the experimenter's corrections and replicating the experimenter's corrections to determine if they result on the elimination of orange labels will help me have a better picture of whether the use of the DF in the pieces of the ND dataset would be helpful in correcting these pieces or, on the other hand, would confuse the experimenter by highlighting spots in the piece that do not need correction.

Chapter 5 Experiment

This chapter provides details on each of the three stages of the workflow, starting from the capture of the digital images. Section 5.1 provides the details of the digitization process, carried out during my second visit to Guatemala in January 2019. Section 5.2 provides the details of the OMR process of the GuatC 1 within MuRET. Finally, Section 5.3 focuses on the experiment of evaluating whether the use of the dissonance filter within the MP Editor makes the process of editorial correction more efficient. It includes subsections on results and discussion.

5.1 The Digitization Process (Second Visit)

The digitization process of GuatC 1 took place in the Archivo Histórico Arquidiocesano de Guatemala (AHAG, the archive holding the sources) between the 07 and 15 of January 2019. Table 5-1 shows the activities performed for each of these dates. Section 5.1.1 presents the results of this digitization process; this is, information regarding the format and number of image files obtained. Section 5.1.2 presents the details of what happened during the digitization process. This includes the procedure followed for the digitization of individual pages, the issues encountered during the process, the solutions, and future recommendations.

Table 5-1: Calendar of digitization process.

Monday	Tuesday	Wednesday	Thursday	Friday
JAN 07 - Conservation treatment - Test two book cradle configurations with GuatC 1	**08** - Set up book scanner & camera parameters - Test digitization procedure to find future issues	**09** - Photographing manuscript	**10** - Photographing manuscript	**11** - Photographing manuscript
14 - Photographing manuscript	**15** - Reshooting folios that had issues	**16**	**17**	**18**

5.1.1 Results

The result of this digitization stage consisted of images of the folios, spine, and covers of the book. The folio images have a size of 4016 x 6016 pixels, they are stored with a 300 dpi resolution, with color model RGB, and 24 bit depth (i.e., 8 bits per color channel); the color profile used at the moment of capture was sRGB. Ignoring the test shots of the second date on Table 5-1, I took a total of approximately 580 photographs to digitize the manuscript folios and discarded around 180 of these. Upon each capture, two image files were generated, one in JPEG and one in NEFF format, the only formats available in the Nikon D750 camera, with NEFF being Nikon's equivalent of a raw image. I converted the NEFF files into TIFF using the official Nikon software *Capture NX-D*, which I downloaded from their website since the one included with the camera was outdated.[119] Both NEFF and TIFF images were stored as master files (without any edits) in three different hard drives (one located at McGill University, and two located at home). These hard drives also contain the JPEG files obtained upon capture. The average size of the TIFF files is of 72.8 MB, while the average size of the JPEG files is of 17.4 MB. I have a 395 TIFF files and the same number of JPEG files—including all folios (blank pages, music pages, index page, and the table of contents page)

[119] The official Nikon software for the Nikon D750 camera can be downloaded in this webpage, in the tab 'software:' https://downloadcenter.nikonimglib.com/en/products/175/D750.html.

the images for the front and back covers, and the book's spine—resulting on a total size of approximately 35.6 GB of stored images.

5.1.2 **Discussion**

As indicated before, the digitization of GuatC 1 took place between the 07 and 15 of January (2019) in the AHAG. This avoided the transportation of the book outside of its current location. As indicated in the methodology, since the AHAG does not have its own conservation and digitization departments, the conservation and digitization tasks had to be outsourced. The Centro de Rescate, Estudio y Análisis Científico del Arte (CREA)[120] took on the conservation task. CREA's conservators worked on the assessment of the material for digitization (presenting a conservation report on November 2018), and its preparation for it. Regarding the digitization task, unfortunately, I was not able to find a digitization technician with training on special collections. However, the Centro de Investigaciones Regionales de Mesoamérica (CIRMA, the institution holding the microfilms of some of the AHAG's music documents)[121] provided me with the contact of Daniel Hernández Salazar,[122] an internationally known Guatemalan professional photographer, who once collaborated with CIRMA on the digitization of a large map.

On 07 January, GuatC 1 received basic conservation treatment to prepare it for digitization. The DIY book scanner was set up and tested the next day, with Hernández Salazar setting up the camera parameters to be used during the rest of the project. Imaging started on 09 January and continued for the next four business days. I conducted the digitization of the folios, with Hernandez Salazar occasionally visiting (on 09 and 10 January) to answer any questions and to photograph

[120] https://www.crea.com.gt/
[121] http://cirma.org.gt/glifos/index.php/Página_principal
[122] Salazar is famous for his historical memory work, including his photo exhibition of "Genocide Dismissed, Guatemala a Silenced Tragedy." Personal website of the photographer: https://danielhernandezsalazar.com/.

the book's covers and spine. I used the last day for reshooting pages that presented issues—e.g., a tilted color patch or a small, detached piece of paper covering some of the text or music content of the page. The whole process took seven business days or, equivalently, a week and a half.

On 07 January, while CREA's conservators cleaned the GuatC 1 choirbook and prepared it for digitization, I mounted the book scanner in its facing-up V-shaped configuration to test it with the original source. I confirmed that, just like in the test I ran at home (Section 3.1.3.1), this configuration was not suitable given the equipment I had available—a tripod rather than an adjustable camera mount and with no platen allowed for flattening the pages—and I started preparing the open-at-one-side configuration for the next day. As indicated in the methodology (Section 3.1.3.1), in the open-at-one-side configuration it was easier to keep the camera perpendicular to the photographed folio and at a constant distance from it by using a copystand (as was shown in Figure 3-2), and it also allows for an even distribution of the lighting by locating each light source at one side of the page (as previously shown in Figure 3-3).

The second day (08 January) was mostly used for testing and making decisions about the equipment and digitization procedure to start the actual imaging process the next day. The photographer set up the camera parameters on this day, and they remained fixed for the complete digitization of the folios. The first thing he did was to position the camera in the copystand arm and determine the zoom based on the distance between the camera and the manuscript. The sliding measure along the copystand arm reaches up to 59 cm in increments of 0.5 cm. The photographer positioned the camera at 50 cm and based on this distance he adjusted the zoom of the lens, setting it to a focal length of 44 mm. The lens that came with the camera was an AF-S Micro NIKKOR 24-85 mm f/3.5-4.5G ED VR. According to the photographer, the ideal would have been to keep the focal length closer to the median value of the lens—this is, the median value between 24 mm and 85 mm, which is approximately 55—and then place the camera at the appropriate distance. However, we were limited by the length of the copystand arm and that is why he decided to fix the distance first and determine the zoom second.[123]

After placing the camera and considering that I was going to conduct the digitization on my own, the photographer set the camera to one of the automatic modes, setting it to shoot in Mode A (Aperture-Priority Auto), where one sets the aperture and the camera automatically adjust the shutter speed to obtain optimal exposure.[124] He set the aperture by testing different f-numbers, photographing a folio with each of these values and comparing the results, especially in the corners of the page (the areas that are the farthest from the camera, which is looking directly at the center of the page). He evaluated the photos to determine which f-number made the corners look clearer, better defined, and chose that as the aperture value to use, ending up with an f/16 aperture. He also set the camera to perform automatic white balancing.

I conducted two digitization sessions per day on the following four business days (09–11 and 14 January), one in the morning and one in the afternoon. While the photographer was not able to set the parameters prior to every session—as it was recommended by the FADGI guidelines and digitization experts at BAnQ—the parameters set by him on 08 January and the imaging

[123] When using a copystand, one is always restricted by the length of its arm. Therefore, a combination of a sturdy tripod with a horizontal tube to place the camera (as shown in Figure 2-9 of Section 2.2.2.1) can be used instead to extend the maximum reach.

[124] "Aperture controls the brightness of the image that passes through the lens and falls on the image sensor. It is expressed as an f-number (written as "f/" followed by a number), such as f/1.4, f/2, f/2.8, /f4, f/5.6, f/8, f/11, f/16, f/22, or f/32. Changing the f-number changes the size of the aperture, changing the amount of light that passes through the lens. The higher the f-number, the smaller the aperture and the less light that passes through the lens; the lower the f-number, the larger the aperture and the more light that passes through the lens." https://imaging.nikon.com/lineup/dslr/basics/04/04.htm

"Shutter speed is a measurement of the time the shutter is open, shown in seconds or fractions of a second… The faster the shutter speed, the shorter the time the image sensor is exposed to light; the slower the shutter speed, the longer the time the image sensor is exposed to light." https://imaging.nikon.com/lineup/dslr/basics/04/03.htm conditions remained constant during the whole digitization period. I used the same room to photograph the whole manuscript. The only source of light were the two natural-light LED panels at fixed positions (I marked the spot in the floor for each of the three legs of the light stands). The lights in the room were turned off during digitization and the door and window's wooden doors remained closed to avoid external light sources.

Also on 08 January, the photographer pointed out a few issues before starting the digitization process:

- He indicated that the lens included with the borrowed camera was not the ideal for this kind of job, and that a macro lens would be preferrable. Since no other lens was available, I used the one provided with the camera.
- He pointed out the need to clean the lens, bringing a blower and a lens cloth the next day to remove dust and fingerprints.
- He proposed an easy way to keep track of the versos (which, contrary to the rectos, are not numbered). His way was to include the number in the photograph itself by including a small piece of paper with the printed number, which I placed next to the color patch (see Figure 5-1 and Figure 5-2).
- He advised me to buy a longer data cable than the one provided with the camera, so that it would be easier to connect the camera—which was way up high in the copystand—to the computer to transfer the photos.

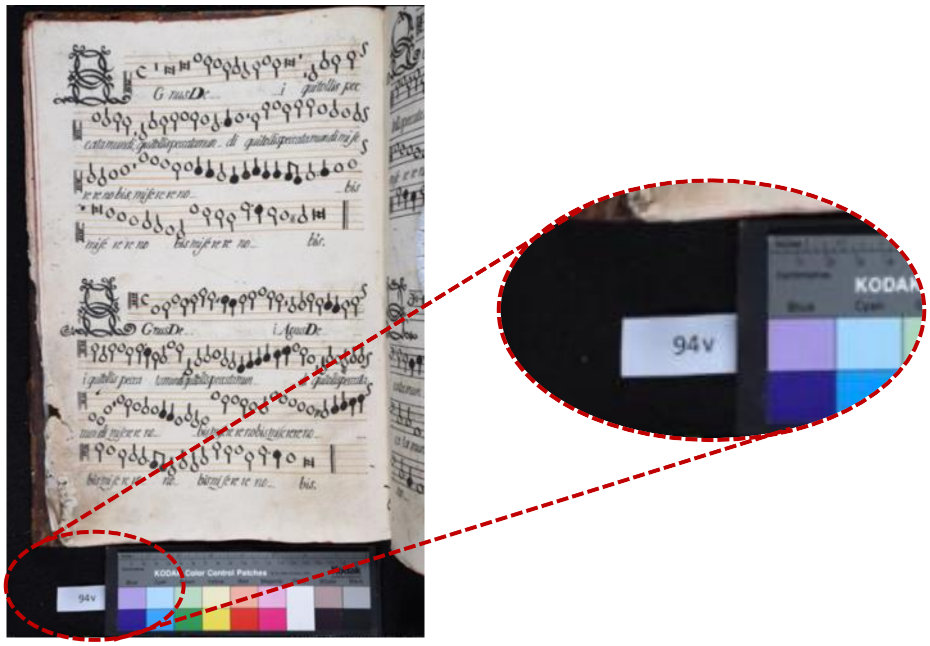


Figure 5-1: Photograph of a verso folio, with the number indicated by a piece of paper at one side of the color patch.

Figure 5-2: Setup for the photograph shown in Figure 5-1, with a black-felt sheet covering the cradle, the color patch standing over a stack of cardboards, a printed number placed next to the color patch to keep track of the verso numbers, and a book serpent holding the pages still on the cradle's door.

I decided on shooting tethered from my laptop using the *DslrDashboard* application and connecting the camera wirelessly to my computer.[125] Shooting tethered allowed me to review the

[125] This application was free and allow me to preview the photos on my computer. The application can be downloaded at www.dslrdashboard.info. For information on how to use it with a wireless connection to a Nikon

photos on my laptop; after a few seconds of firing the shutter button, the image appears on the computer screen. Reviewing the pictures on the camera would have been very inconvenient given that it was placed up high in the copystand. Besides, evaluating the photos in the camera monitor implies manipulating the camera, which can result in unwanted changes to the camera's position. In addition to reviewing the photos on my laptop, shooting tethered allowed me to shoot from my laptop. Shooting remotely allows to avoid pressing the camera's shutter button, which is important because (1) this action can introduce unwanted vibrations on the photograph, (2) it is impractical to press the button when the camera is set up high, and (3) if the camera is far away from the manuscript, you might need one person to handle the manuscript and another one to handle the camera.[126] At the end of each day, I would connect my laptop to the camera with the (long) data cable and download all the photos, which I then copied to an external hard drive.

I used most of the equipment listed in the methodology in Table 3-3, except for the control remotes which were no longer needed since I was shooting from the computer. The whole setup was placed close to a power outlet so that I could keep both camera and laptop connected to avoid running out of battery. As mentioned in page 138, I also used some extra material that was not mentioned in the table. I borrowed a color and gray-scale reference patches, both including a dimensional scale, from the DDMAL Lab. At the end, I did not have enough space to locate both color and gray-scale patches and decided on using the color patch since I was taking color photographs. I could have cut the color and gray-scale squares into smaller squares to fit them (this was part of DIAMM's and FADGI's advice, see G19), but I preferred not to do it since these patches were borrowed. I also bought two book serpents to help holding the pages that were in the

D750 camera, I consulted this video: https://www.youtube.com/watch?v=5WsS-e8wUSI&list=PLC08PttZdHNBVP6qU9jOqXl0qlhqRGJRY&index=1.

[126] Prof. Ichiro Fujinaga suggested a possible solution to this third point: using a pedal to activate the shutter button, having more hands free to manipulate the manuscript.

door of the cradle so that they did not fall on top of the page to be photographed (as shown in Figure 5-2).

Because I only had a single camera with the open-at-one-side configuration, I photographed all the rectos first, turned the book around, and photographed all the versos. I came up with a small protocol/procedure for digitizing each folio to make the process as uniform as possible. The digitization procedure consisted of the following steps:[127]

1. Turn the page.
2. Place the book serpent on top of the page that is on the door side—so that it does not fall over the page to be photographed.
3. Change the piece of paper including the printed number, if photographing a verso.
4. Quickly review the picture taken before (previous folio), which should be uploaded in the computer by then. This inspection consists of quickly evaluating if the picture is in focus and to look for shadows. If there is a shadow in the low part of the page, this comes from the cardboard stack where the color patch is placed, and it is necessary to remove a cardboard to lower the color patch and take the photo again.
5. Press focus on the *DslrDashboard* app, which uses the auto-focus function of the camera.
6. Shoot from the app.
7. In the case of a verso, check once again that I used the right printed number.
8. Go to step 1.

[127] The steps can be seen in the "Digitization of Music Manuscript – Part 2" found at https://www.youtube.com/watch?v=4KFCom8Vndo. This video is part of a playlist about the "Digitization of Guatemalan Music Manuscript," which can be seen at https://www.youtube.com/watch?v=VDbJQSMjB0s&list=PLC08PttZdHNBpl8z1MV2U3UZFNBL3jQ5g.

I omitted steps 3 and 7 when photographing rectos, since the number is included at the top of the folio for rectos.

As I moved through the pages of the book, eventually the stack of cardboards holding the color patch projected a shadow over the photographed page. When a shadow started to appear at the bottom of the page, I removed a cardboard from the stack. Similarly, as I moved through the pages, the photographed folios start getting farther apart from the camera—which starts showing in the photographs by the black felt area getting bigger. To try to keep the distance between the camera and the folio constant, I lowered the camera 0.5 cm when I have gone through approximately the same distance in terms of folios.

As I started getting closer to the other end of the book, most of the pages would lie on the door-side of the cradle (rather than in its horizontal plane) and they would tend to fall as the book serpent was no longer heavy enough to hold them in place. Moreover, the page to be photographed tended to curve more because the many pages lying on the cradle door pulled the few pages that were lying on the cradle's horizontal surface. To solve the first issue and avoid the book from closing abruptly, I increased the angle of the opened book by moving the cradle's door a bit behind. To solve the second issue (the increase in curvature of the page to be photographed), I lifted the book spine by cutting two pieces of the black felt, rolling each of them into a conic shape, and inserting each cone below one of the ends of the spine (always under the black-felt sheet). This helped both giving support to the spine and flattening the page that lied on the horizontal surface of the cradle by rising the spine and gutter.[128]

[128] The solution to these two issues can be seen in the "Digitization of Music Manuscript – Part 3" found at https://www.youtube.com/watch?v=cDn3b8RsQBM. This video is part of a playlist about the "Digitization of Guatemalan Music Manuscript," which can be seen at https://www.youtube.com/watch?v=VDbJQSMjB0s&list=PLC08PttZdHNBpl8z1MV2U3UZFNBL3jQ5g.

Figure 5-3: Two pieces of the black felt sheet (used for a neutral background), cut and rolled into conic shapes to lift each of the ends of the book spine.

Another issue that emerged during the digitization process was that the computer application I was using for shooting tethered would get stuck from time to time. My solution at that moment to avoid delaying work was to switch to *Nikon's Wireless Mobile Utility* (*WMU*, version 1.6.2),[129] a mobile application I already had downloaded on my phone. After all, the photos were stored in the camera with the same parameters that have been used so far regardless of the app used since I was not using the software for setting or changing the camera settings. After further inspection, I discovered that tethering software is known for being finicky and that sometimes it stops working for no apparent reason, so it is always good to have a plan B just in case. Another possible reason for the sudden termination of the software could have been related to memory-issues with my computer. The JPEG images were saved internally on my laptop, and I erased them every couple of folios—since I was only using the laptop to review the pictures. In the future, I plan to take measures to reduce the chances of having this kind of issues. First, I will use the official software for the camera. I have found a few options for tethering software in the Nikon website including NX Tether and Camera Control Pro 2.[130] Using the official software can

[129] I found this app and information on how to use it in this video: https://www.youtube.com/watch?v=GuXGiqhC2f0&list=PLC08PttZdHNBVP6qU9jOqXl0qlhqRGJRY&index=2.
[130] Check https://www.nikonusa.com/en/nikon-products/imaging-software/nx-tether.page for Nikon NX Tether, and https://www.nikonusa.com/en/nikon-products/product/imaging-software/camera-control-pro-2---full-version-%28boxed%29.html for Nikon Camera Control Pro 2.

help avoid compatibility issues. Second, I will connect the camera to the computer by (a long) cable rather than wirelessly to avoid connectivity issues. And, finally, I will explore the option to directly store the images in the hard drive to avoid memory issues. This was my first time shooting tethered and while I took advantage of its preview and remote shooting functionalities, I think that further exploration of the tethering software would have been beneficial since there are options in this kind of software to change the filepath where images are be saved (useful to save the images to the hard drive directly), add comments (useful to keep track of the verso numbers), and add stars to images among other things.

5.2 OMR of GuatC 1 in MuRET

Once I obtained the images, I proceeded to the OMR stage using MuRET. In the MuRET framework, I performed the automatic steps of document analysis, agnostic transcription, and semantic transcription, and manually corrected the results of these steps when needed.

Section 5.2.1 presents the results of the full OMR process; these are, the Parts-based MEI files retrieved by MuRET. Section 5.2.2 presents the details of what happened during the OMR process. This includes the status of MuRET at the beginning of the OMR stage; the dataset used for training the various machine-learning models; the common sources of error in the MuRET steps for document analysis, agnostic transcription, and semantic transcription; and the time to process each page through each of the steps in MuRET's interface (i.e., document analysis, part assignation, agnostic transcription, and semantic transcription).

5.2.1 Results

The results of the OMR process in MuRET are Part-based MEI files encoding the music symbols of each part/voice in a piece. For the experiment, I generated twenty-three files, corresponding to the twenty-three self-contained units mentioned in Table 4-2 and Table 4-3, having a total size of 3.6 MB. Therefore, the average individual file size is of 157 KB. All the OMR files for the experiment can be found in the "Experiment" branch of the "GuatC1" GitHub repository.[131]

5.2.2 Discussion

MuRET was still in development and when I got to the OMR stage (June 2019), MuRET did not include the models for document analysis and holistic staff-recognition of mensural symbols yet; the only available model was the one working at symbol level.[132] Therefore, I used this time to get familiar with MuRET, prepare training data for the document analysis by drawing bounding boxes for the staff regions, and use the symbol-level classifier to start annotating the symbols in these staff regions to prepare training data for the holistic staff-level recognition. During this process, I became one of the first testers of MuRET for handwritten mensural notation and raised several issues that were fixed by David Rizo. By August 2019, the holistic staff-recognition model for mensural notation that was trained on Spanish manuscripts—mentioned in Section 2.3.2—was implemented into MuRET under the name "Zaragoza End2End."[133] This Zaragoza model made the classification of symbols in the staff more efficient than using the symbol-level classifier, and I prepared more folios ready to be used as training data for a Guatemalan holistic staff-level classifier (still the document analysis part had to be done manually). The training data consisted of folios that were classified with the symbol-level and the Zaragoza End2End models; these randomly selected folios are:

- The complete folios from the first piece, *Asperges me* by Bermúdez (1v–4r). These were processed with the symbol-level classifier.

[131] https://github.com/martha-thomae/GuatC1/tree/Experiment
[132] For details about these models, please consult Section 2.3.3.3.
[133] This model has had different names within MuRET including "Python A end2end with muret_model_v0_40."

- Two folios from the twelfth piece, the mass *Ave maris stella* by Victoria (106v–107r). These were processed with the symbol-level classifier.
- The complete folios from the sixteenth piece, the *Missa de Bomba* by Bermúdez (152v–169r). These were processed with the holistic staff-level classifier, except for f. 158v–152r (first two sections of the *Credo—Pater omnipotente* and *Crucifixus*), which were processed with the symbol-level classifier.
- The complete folios from the twenty-second piece, *Lumen ad revelationem* by Franco (186v–187r). These were processed with the holistic staff-level classifier.

With this training data, Jorge Calvo-Zaragoza and Antonio Ríos-Vila trained a new holistic staff-recognition model for the Guatemalan manuscript. I used this "Guatemala End2End" holistic staff-recognition model to classify the rest of the symbols in the manuscript.[134] While I kept on doing the document analysis by hand for a while, eventually the "Guatemala SAE" document-analysis model was uploaded into MuRET, and I used it to automatically retrieve the staff regions for the remaining folios (around half of the manuscript). The Guatemala SAE was trained with the folios that were previously classified, so, around half of the manuscript.[135]

These two models, the Guatemala SAE and the Guatemala End2End model, facilitated the tasks of staff recognition and symbol classification within those staff regions for the GuatC 1 book. The Guatemala SAE's document-analysis model provided the bounding boxes for each folio, and I just expanded the area of some of them to cover any note that was above or below the usual staff

[134] This model has had different names within MuRET including: "Guatemala End2End 2019" and "Python A end2end with Guatemala_model_190."
[135] This model has had different names within MuRET including "Python D.A. with Guatemala."

region (Figure 5-4a). On a few occasions I had to remove extra bounding boxes—there was one bounding box that was consistently drawn at the edge of the versos, in the black felt area (Figure 5-4b).

(a) Some staff bounding boxes needed to be resized for notes on top or below the staff lines (see selected staff region).

(b) Extra staff region in the black background area on versos (see left bottom corner).

Figure 5-4: Types of errors in the document analysis stage.

On the other hand, the Guatemala End2End model provided me with the symbols in the staff region, which I then corrected (if needed) or added any missing symbols with the symbol-level classifier. The biggest sources of error for the Guatemala End2End model were:

- It did not recognize F clefs, and all C-clefs on the first line were recognized as C-clefs on the second line. The issues with these two types of clefs are due to the lack of them in the training data.
- Beamed notes were recognized as semiminimas (or umbeamed fusas). This is also because there are very few beamed notes in the training data compared to the amount of semiminimas.
- It recognized very few custos. I had to use the symbol-level classifier to add them and most of the time this model will recognize them as semiminimas, maximas, or flats. The symbol-level classifier was trained on a different corpus, where the custos look different.
- It did not perform well when symbols matched vertical positions (see Figure 5-5). This was the case for fermatas and accidentals that are placed above the notes, and for key signatures with more than one accidental since these accidentals would share the same vertical position. The issue here is due to the fact that the holistic staff-level model assumes that is working with a sequence of symbols (and that there are no stacked symbols).

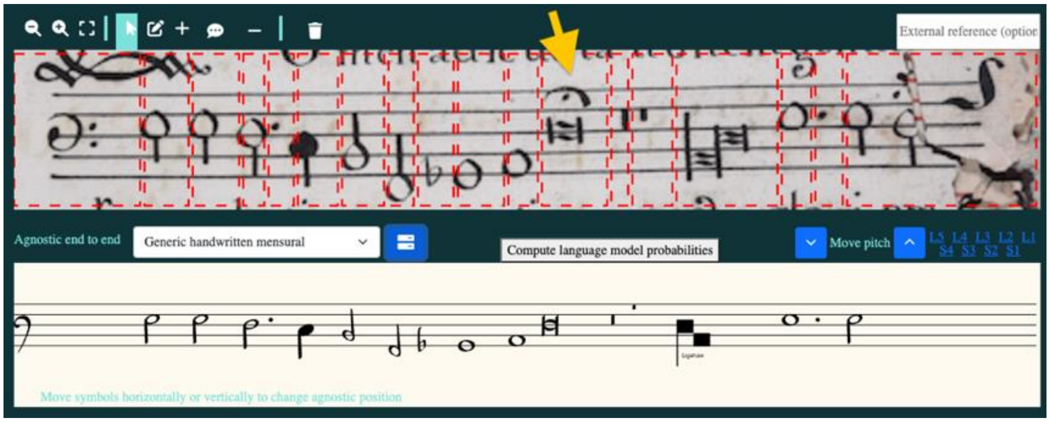


Figure 5-5: The fermata and the F breve note share the same vertical position, but only the breve is present in the agnostic transcription while the fermata is missing. This is because the holistic staff-level symbol recognition model is designed for processing a single sequence of symbols per staff (as in monophonic music) and cannot process symbols that share the same vertical position.

- There were other kinds of mistakes, but they were not as common as the first three and were easier to fix (e.g., minim rests, accidentals, and dots were confused sometimes; some rests were recognized with the wrong value every now and then; and some ligatures were recognized as breves).

The errors mentioned are related to the document-analysis and agnostic-transcription steps of MuRET, for which I used the models trained on the GuatC 1 data. For the semantic-transcription step, no training was needed since it uses the pre-existing "Agnostic-to-Semantic Transducer" presented in Section 2.4.2.[136] The semantic step also resulted in a common set of corrections, which are listed below:

- Providing the correct note values and pitches for ligatures in the **mens table. This must be done for all ligatures because they are not supported in MuRET yet.

[136] This transducer has had different names in MuRET including "Java Agnostic 2 Graph Transducer."

- Correcting accidentals that are mistakenly recognized as part of the key signature because they are located at the start of the system (since they belong to the first note in that system).
- Correcting the stem direction of a few notes.[137]
- Relate accidentals (or, less likely, dots) to the right note. This is needed when the accidental (or dot) is closer in proximity to a note other than the one it is supposed to affect.
- Add gestural information for accidentals. This needs to be done for accidentals that are cancelling the preceding accidental or a key signature's accidental. During the experiment, I had to enter this correction after exporting the MEI file from MuRET.[138]

Some of these issues have been addressed already, that is the case with the last and antepenultimate issues as explained in their footnotes. The time invested in each part of the OMR in MuRET varied during the processing of the complete manuscript as the issues in MuRET were being found and solved. By the end of the OMR process, and considering future improvements, the time invested in each part of the process is:

- Document analysis: less than a minute when using the SAE to retrieve the staff regions, resize a few, and delete any extra region. This is an improvement compared to the time invested on manually drawing these bounding boxes, which took between two and three and a half minutes.

[137] This was a bug in the transducer that should be fixed by now.
[138] However, recent improvements on MuRET allow for this correction to be entered in the **mens table directly.

- Part assignation: seconds, considering that the user already knows the parts involved in that section of the piece.
- Agnostic transcription: 7–8 minutes in average using the Guatemala End2End model. Pieces that took a bit longer involved a combination of the symbols that were not handled well by the model: presence of F clefs and C1 clefs (which were not present in the training data), pieces with key signatures and many accidentals (which implied lots of vertical position matching), and beamed notes. If the user wanted to spend time correcting the bounding boxes coordinates for the symbols, the time spent in a page could increase up to 20 minutes.[139]
- Semantic transcription: less than 5 minutes using the "Agnostic-to-Semantic Transducer." Most of the time in this semantic transcription step consists of quickly verifying that the notes rendered in the semantic transcription panel are ok (which provides a kind of double-check since these notes were verified once in the agnostic panel) and making quick corrections for the type of errors enumerated before. I guess that with the more recent fixes to the transducer and MuRET's semantic-encoding support, the amount of time invested in the semantic transcription could be reduced to 3 minutes.

5.3 The Dissonance Filter Experiment in the MP Editor

The following sections contain details about the experiment, which consists of evaluating whether highlighting illegal dissonances according to counterpoint rules of Renaissance style (using

[139] Since the bounding boxes for symbols are not preserved in the MP Editor, which only supports bounding boxes for the systems, this is not important unless one retrieves the symbols' bounding boxes encoded in the <zone> elements and adds them back into the MP Editor's Score-based MEI file using a post-processing script.

humlib's dissonance filter, DF) makes the process of correcting the scored-up OMR files more efficient. At this point, since the user has already manually corrected the OMR errors using MuRET's interface (specially designed to support this action as is the case of various OMR frameworks), I expect most errors in voice alignment to be due to scribal errors. Since scribal errors are hard to find and tend to require visualizing the piece in score layout, I am including the illegal dissonances' highlighting functionality (through the DF) in the MP Editor's score interface and evaluating whether this functionality facilitates the identification of the scribal errors and their editorial correction. Chapter Chapter 4 presented the two datasets prepared for this experiment—the D dataset where the experimenter can use the DF, and the ND dataset where the experimenter cannot use this filter—and the evaluation procedure. The current chapter focuses on the results and discussion of the experiment, with Section 5.3.1 presenting the results (these are, the correction time per piece and resulting files) and Section 5.3.2 providing the discussion of these results.

5.3.1 **Results**

The experimenter corrected the voice-alignment using the MP Editor score interface, activating the dissonance filter to correct the pieces from dataset D and keeping the filter deactivated when correcting the pieces from dataset ND, and she annotated the correction time for each piece. Two types of results were obtained from this process: correction time and output files. The time invested in the correction of each piece in these two datasets is shown in Table 5-2 and Table 5-3, with the mean and standard deviation of the corresponding dataset.

Table 5-2: Voice-alignment correction time for each piece in the ND dataset.

Self-contained Units	Folios	Number of			Time (min)
		Measures	Voices	Illegal Disson.	
Missa8.4_Sanctus0	52v-53r	29	4	9	8
Missa9.2_Gloria1	56v-58r	44	4	46	13
Missa9.2_Gloria2	58v-60r	37	4	8	17
Missa10.1_Kyrie0	71v-73r	71	4	4	18
Missa10.3_Credo1	76v-78r	67	4	64	16
Missa10.5_AgnusI0	83v-84r	23	4	2	7
Missa15.3_Credo2	141v-143r	56	4	43	15
Missa15.3_Credo3	143v-144r	48	3	0	15
Missa16.4_Sanctus2	166v-167r	13	4	17	11
Missa17.5_Agnus0	179v-180r	26	4	1	6
Piece21_Surrexit0	184v-186r	17	5	21	10
				Average	12.4
				Standard Deviation	4.2

Table 5-3: Voice-alignment correction time for each piece in the D dataset.

Self-contained Units	Folios	Number of			Time (min)
		Measures	Voices	Illegal Disson.	
Missa7.4_Sanctus1	38v-39r	29	4	0	1
Missa7.4_Sanctus2	39v-41r	50	4	52	10
Missa9.6_AgnusII0	70v-71r	35	5	14	7
Missa10.3_Credo2	78v-79r	40	4	3	2
Missa10.3_Credo3	79v-81r	64	4	17	30
Missa13.5_Agnus0	118v-119r	30	4	6	6
Missa14.1_Kyrie0	119v-121r	56	4	23	7
Missa15.1_Kyrie0	133v-135r	44	4	0	1
Missa15.3_Credo1	139v-141r	64	4	61	2
Missa16.4_Sanctus1	165v-166r	22	4	2	1
Missa16.4_Sanctus3	167v-168r	27	3	16	9
Piece25_Surrexit0	189v-190r	15	4	13	6
				Average	6.8
				Standard Deviation	8.0

The resulting files can be consulted in the "Experiment" branch of the "GuatC1" GitHub repository,[140] which contains:

- The Mensural MEI Parts-based files that resulted from the OMR process, and that were used by the experimenter as the starting point. These are the files that Gates-Panneton uploaded into the MP Editor.
- The Mensural MEI Parts- and Score-based files that came out of this experiment. These are the output files from the MP Editor that contain Gates-Panneton's corrections.
- The text files that encode the OMR score, the experiment resulting scores, and the CPDL transcription scores as a set of tokens (see description of these in Section 4.2). One can use a diff tool to compare these files to quickly determine the corrections made by the experimenter (by comparing the OMR tokens vs. the experiment results tokens) and find the differences between the experimenter's results and the CPDL transcription (by comparing the experimenter results tokens vs. the CPDL transcription tokens).[141]

The following section provides a summary of the analysis of the results (time and files), discussing the "usefulness" of the DF in the MP Editor.

5.3.2 Discussion

In this section, I discuss the results given in Section 5.3.1 by presenting the lessons learned from the use of the dissonance filter (DF) in the editorial correction stage of the encoding workflow and providing examples for each of these lessons. All the examples presented below use modern clefs and have a "note about clefs" statement in the caption clarifying the clefs of each voice.

[140] https://github.com/martha-thomae/GuatC1/tree/Experiment
[141] An example of an online diff tool can be found at https://www.diffchecker.com. There are many others.

By analyzing the corrections of the individual pieces and looking at their correction time (see Table 5-2 and Table 5-3),[142] the following points become apparent: the DF detects some "false" illegal dissonances; it saves time in editorial corrections and improves the accuracy of these; it is also useful for detecting OMR errors; and it is sensitive to changes in style. All these points are discussed in detail below.

5.3.2.1 *False Illegal Dissonances*

False illegal dissonances are notes that are wrongly identified as illegal dissonances by the DF; in other words, they are false positives, being in fact legal dissonances or even consonances. The main types of false illegal dissonances found in the dataset are rearticulations of the same pitch, short note values doubling the agent, and notes of extremely small rhythmic level (e.g., fusas). Examples of each of these cases are presented below.

Rearticulations. When a pitch is repeated (or rearticulated) within a single beat the DF sometimes identifies one of those pitches as an illegal dissonance, although if the two notes were combined into one note value, they would not be flagged as illegal dissonances. Figure 5-6 shows an example of this. Its left panel shows the original reading of a mass containing a rearticulated D minim in the tenor. The DF marks the first of the two D minims as an illegal dissonance with an orange label. If the repeated D minims are replaced by a single D semibreve, as shown in the right panel of Figure 5-6, the orange label is replaced by the correct dissonance label, this is a blue 'g' representing the agent of the suspension that is now marked with an 's' in the superius. As indicated in Schubert (2008, 146), "you may now [in two voices with mixed values] have a simultaneously attacked

[142] The full details of the corrections of each piece and their comparison to corrections done by the CPDL transcriber can be found in the Experiment branch of the GuatC1 GitHub repository at https://github.com/martha-thomae/GuatC1/blob/Experiment/Appendix%20C.pdf.

dissonance if it results from the rearticulation of a pitch… the agent or the patient in a suspension may be broken in two and rearticulated… This usually happens when the text has a new syllable.…”

Figure 5-6: Original reading of Missa9.6_AgnusII0 in MP Editor (left), showing an illegal ‘Z’ dissonance detected by the dissonance filter in the tenor. This is not a true illegal dissonance, but the result of a rearticulated note given the syllables “na” and “no” below each minim (from “dona nobis”). This is evident when we replace the two minims with a single semibreve (right) to show the correct dissonance labels in that passage. ***Note about clefs:*** *From the bottom up, the voices are in F clef, suboctave G clef, and the three upper voices are in G clef.*

Short notes doubling the agent. Here it is important to make a distinction between a “note doubling the agent” or “doubled agent,” and a “double agent.” When using the terms “note doubling the agent” or “doubled agent,” I am referring to an agent that is repeated in two voices (same pitch class). When using the term “double agent,” I am referring to two agents with *different* pitches for a suspended note—both of which behave as agents.

Examples of double agents are given in Figure 5-7, where both the G breve in the upper voice I and the C semibreve in the bass (yellow boxes) act as agents of the tenor's suspended F (blue box). As can be seen in the figure, "double agents" are handled correctly by the DF as both notes are identified with a blue 'g' label. These are not false illegal dissonances; they are legal dissonances.

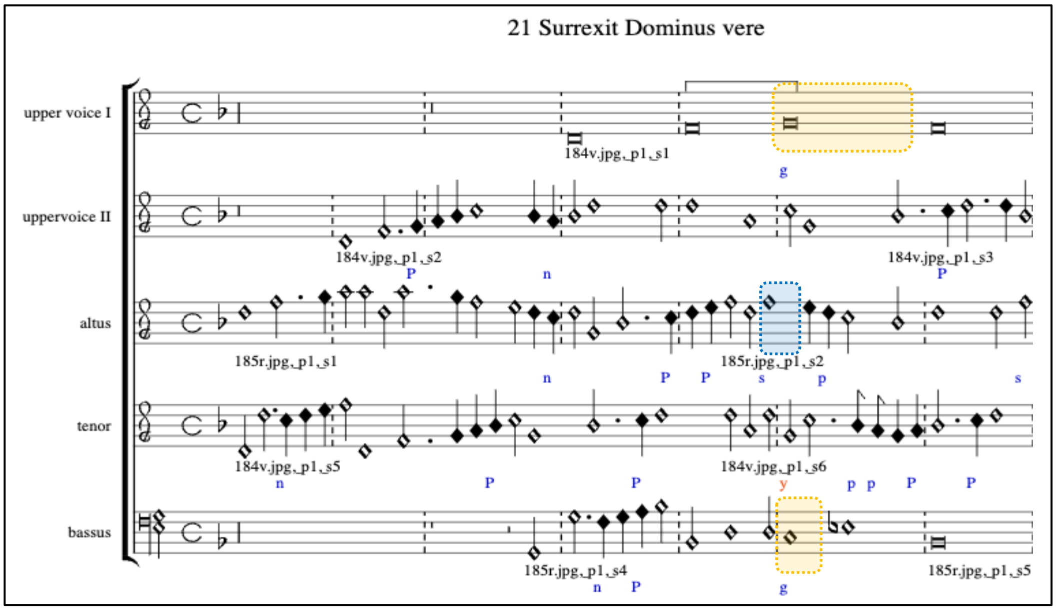
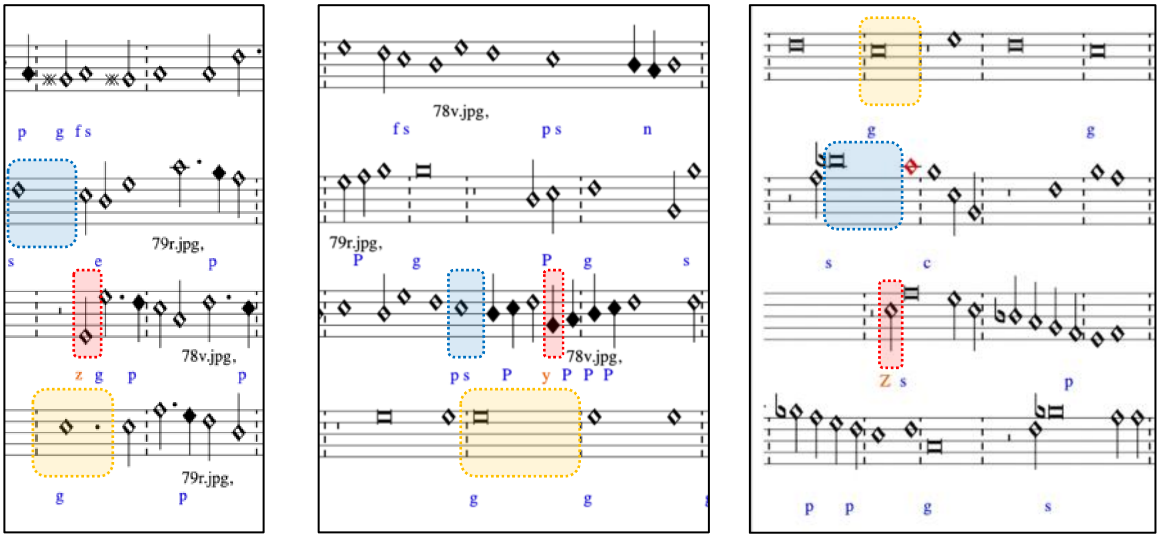


Figure 5-7: Example of double agents (yellow boxes) for a single suspension (blue box) in Piece21_Surrexit0. ***Note about clefs:*** *The upper voices I and II are in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

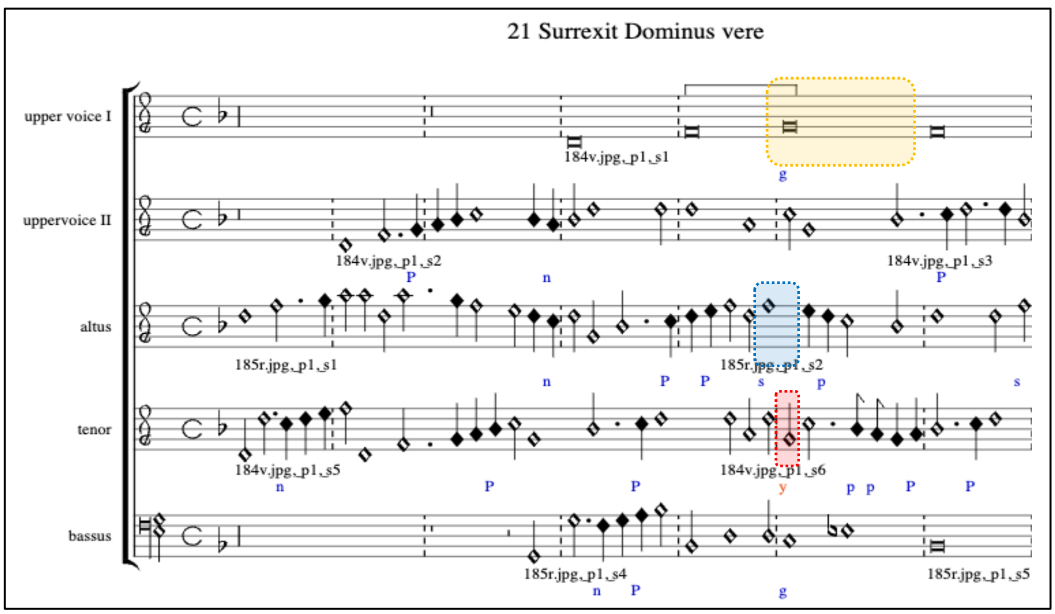
The false illegal dissonance issue happens in the case of double**d** agents—where the pitch class of the agent is already sounding on another voice—when the note doubling the agent has a value shorter than a semibreve. Even though these short notes form a unison or an octave with the agent, the DF assigns them an illegal dissonance (orange) label. There are several examples of this, including those shown in Figure 5-8. Figure 5-8a shows an orange 'z' label under the tenor's E minim (pink box), despite this minim having the same pitch as the E dotted semibreve still sounding in the bass (yellow box), which has been labelled as an agent with a blue 'g' label. Similarly, in Figure 5-8b, the tenor's G semiminim labelled with an orange 'y' (pink box) is doubling the G breve in the bass (yellow box), which behaves as an agent as indicated by the blue 'g' label. While these examples show doubling of the agent's exact pitch, there are examples of doubling of the agent's pitch class, as can be seen in Figure 5-8c, where the tenor's C (pink box) duplicates the pitch class sounding on the superius, a breve C an octave higher that behaves as an agent (yellow box).

The examples of illegal dissonances in Figure 5-8a–c are all short notes (minims or semiminims) doubling the agent with a later onset than the agent; some of them are acting as agent themselves (Figure 5-8a and Figure 5-8c) and should be marked with a blue 'g' label, while others should be considered consonant (Figure 5-8b) and have no dissonance label. However, even when the short note is sharing the same onset and pitch class as the longer note identified with a blue agent 'g' label in another voice and, therefore, it is clearly acting as an agent itself, it is still misidentified as an illegal dissonance. This is shown in Figure 5-8d by the 'y' orange label under the G minim of the tenor (pink box), which doubles the breve G in the upper voice I (yellow box and marked with a blue 'g' label). Both are agents of the suspended F in the alto, and both should have a blue 'g' label rather than an illegal-dissonance orange label.[143]

[143] According to Schubert (2008, 144), an agent can be a short note (like a minim) and one can skip or step away from the agent: "The note that creates the dissonance in a suspension (the agent) may skip or step to a consonance on the second half-note beat, while the patient is resolving. This does not apply to diminished fourth species, where the agent must stay put."



(a) *Missa10.3_Credo2* (b) *Missa10.3_Credo2* (c) *Missa10.3_Credo1*



(d) *Piece21_Surrexit0*

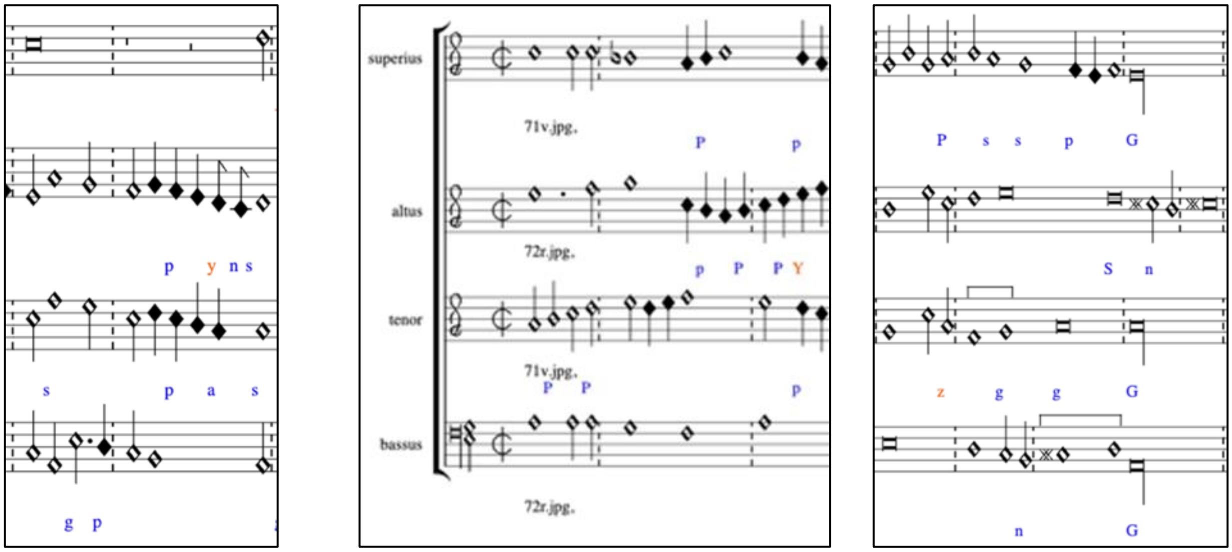
Figure 5-8: Examples of short notes (pink boxes) that are doubling the agent (yellow boxes) of a suspended note (blue boxes) and that are wrongly identified as illegal dissonances.
Note about clefs: *The superius in a–c and the upper voices I and II in d are in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

To solve this issue, short note values labelled as illegal dissonances could be doubled-checked to verify whether they are doubling an agent or not. If they are, the orange label should be removed as we are not dealing with an illegal dissonance and then analyzed to see whether it should be assigned a blue ‘g’ label (when acting as an agent itself) or no label (when consonant).

Notes of small rhythmic value and other false illegal dissonances. Some notes in stepwise passages of small note values, such as semiminims or fusas, and on weak beats are marked with orange labels even though they are acceptable dissonances—and they were kept in both the experimenter and the CPDL version of the pieces.[144] A few examples of these false illegal dissonances are shown in Figure 5-9. Figure 5-9a shows a stepwise passage of semiminims and fusas in the alto. The alto’s D fusa is a fourth above the tenor’s A, which is already dissonant with the bass; therefore, the alto’s D fusa is marked with a ‘y’ orange label. However, this is a false illegal dissonance as the alto’s D fusa is of a very small rhythmic level.[145] Similarly, Figure 5-9b shows a sequence of stepwise semiminims in the alto, where an E semiminim is a second above the tenor’s passing D and, therefore, is marked with a ‘Y’ orange label. However, its status as an illegal dissonance is debatable. The late Renaissance theorist Galilei would have categorized it as an acceptable (legal) dissonance while other Renaissance scholars would have not, as a quarter note against a quarter note was considered as first species counterpoint and all notes attacking together must be consonant with each other (Peter Schubert, personal communication, 27 June 2022). A setting on the MP Editor could be implemented to decide whether to follow Galilei’s theory to allow for this type of dissonance to be legal in a more seventeenth-century style, or to mark them as illegal—as in the current implementation—in a more Renaissance style.

[144] This happens in Missa9.6_AgnusII0, Missa10.1_Kyrie0, and Missa15.3_Credo2.

[145] “Eighth notes may be used in pairs on weak quarter beats in stepwise passing or lower neighbor motion” (Schubert 2008, 88). Examples like this one, with a sequence of semiminims and a pair of fusas with one of the fusas incorrectly labelled with as an illegal dissonance occur in different instances over the corpus.



(a) Fusa in a stepwise passage of semiminims and fusas. *(b) Seminiminim in a sequence of stepwise semiminims.* *(c) Minim by skip on a weak beat.*

Figure 5-9: Examples of some "false" illegal dissonances in Missa10.1_Kyrie0 (b and c) and Missa15.3_Credo2 (a). ***Note about clefs for Missa15.3_Credo2 (a):*** *The superius and altus are in G clefs, and the tenor and bass are in suboctave G clefs.* ***And for Missa10.1_Kyrie0 (b, c):*** *The superius is in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

Finally, Figure 5-9c shows an interesting example of “doubling a passing tone” (rather than doubling the agent). Here, the tenor’s A minim is dissonant as it is a fourth above the E breve in the bass; and because it is a skipped dissonance, it is illegal and, thus, it is marked with an orange ‘z’ label. However, that tenor’s A minim is doubling the A minim in the superius, which is acting as a passing tone. This is another example of a dissonance that is not legal in the Renaissance, but it is acceptable later. The DF is not wrong in its classification of this sonority as an illegal dissonance, it is just necessary to be aware of this case (the doubling of a passing tone); and in the future, as indicated at the end of the previous paragraph, one can implement a setting that deactivates the “illegal” classification of these sonorities for later pieces. Figure 5-9b and Figure 5-9c belong to Mass 10, which is a later mass as will be presented in more detail in Section 5.3.2.2.

In future work, revisions to the DF could avoid showing false illegal dissonances labels for rearticulations, short notes doubling the agent, and notes of extremely small rhythmic level (e.g., fusas and semifusas) as these orange labels distract the user in their correction process.

5.3.2.2 *The DF is Sensitive to Changes in Style – Eighteenth-Century Works*

The DF was designed for Renaissance polyphonic music, especially music of the Josquin generation; it does not work for eighteenth-century musical works. For most pieces in the D dataset, the correction took between 1–10 minutes. But one piece in the D dataset was an outlier compared to the others in terms of time taken for corrections: Missa10.3_Credo3, that took nearly 30 minutes. The experimenter spent considerable time trying to eliminate the orange labels in one passage, and finally decided to leave it as it was (see Figure 5-13), which was the correct decision since this passage combines Baroque and Renaissance style, as pointed out by Professor Peter Schubert.[146] This piece is a section from a movement of the only known eighteenth-century mass of the manuscript.

Mass 10 is one of the six masses added to the contents of the manuscript in the eighteenth century—the other six masses of the manuscript were recopied from a *Libro de Kyries* originally copied in 1602 by Gaspar Fernández. From the six masses copied in the eighteenth century, two of them were composed by Victoria (*Missa11* and *Missa12*), a sixteenth-century composer; one by Matías de Rivera (*Missa17*), a seventeenth-century composer; one by Serra (*Missa13*) and one by Alegre (*Missa14*), for both of whom we do not know their region and period of activity; and one by Torres y Martínez Bravo (*Misssa10*), who lived ca. 1670–1738.[147] Mass 10 is, therefore, the only mass whose composer is known to have lived in the eighteenth century. Moreover, it has

[146] Personal communication, 09 February 2022.
[147] https://hispanicpolyphony.eu/es/person/13479

musical and notational characteristics of later pieces, such as the presence of seventh chords and ties;[148] in particular, ties make it possible to notate durations that are impossible to represent in mensural notation, as shown in Figure 5-10. In Figure 5-10, the note in the orange circle has a length equivalent to five minims. This value is impossible to represent in mensural notation, but it is possible to express in its successor system (common Western music notation) by using ties, which were commonly used with barlines—introduced for solo parts and partbooks in the seventeenth century—to connect notes separated by them and to facilitate the notation of values that cannot be written with a single note (Rastall 2001; Hiley 2001). This mass was composed later than the other masses of the experiment.

In mass 10, the DF did not work as well as it did in the earlier works, because the stylistic changes in this piece no longer conform to Renaissance dissonance rules as shown in Figure 5-11 and Figure 5-12. Figure 5-11 shows two dominant seventh chords (F6/5 and G7) pointed out by the experimenter in Missa10.3_Credo1 (measures 11 and 12). Figure 5-12 shows another example of seventh chords, now in Missa10.5_AgnusI0 (measure 16). The 'y' label under the superius G# represents a note that is only dissonant against a known dissonance, which the case of the superius G# as it is dissonant against the D in the tenor that forms a suspension with the E in the alto. However, these sonorities could also be considered as an E4/2 dominant seventh chord (lacking the B). Finally, Figure 5-13 shows the passage in Missa10.3_Credo3 (measure 20) that took the experimenter 30 minutes to process. As indicated before, this passage combines Baroque and Renaissance style by including seventh chords and fake suspensions, respectively. The third inversion seventh chord can be found over the second half of the red D semibreve in the bass

[148] 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.

(shown by the orange ‘y’). The fake suspension is shown in the blue box, with the alto’s G semibreve over the D in the tenor and bass; the dissonant preparation of the suspension happens at the end of measure 20, the suspended G in the alto forms a diminished fifth with the C# at the beginning of measure 21, and then it resolves down by step. Fake suspensions, which were used in Renaissance music, are among the dissonances included in the Dissonance Filter; they are marked with a blue ‘f’ label (see Table 3-5). However, in the example shown in Figure 5-13, the fake suspension is not marked with the appropriate label. Its simultaneous attack with the tenor’s D minim does not allow the DF to recognize the alto’s G as a fake suspension and, instead, it marks it as an unknown dissonance with an orange ‘z’ label (blue box). The tenor’s D minim itself is also marked with a ‘z’ label (red box). The ‘z’ in the tenor’s D minim is another case of rearticulation issue (red box), but happening in another voice and an octave above—as the D is also sounding in the bass (yellow box).

The experimenter made the right decisions regarding whether or not to correct a note/rest in the sections of this mass when she approached them from a “chord-oriented” point of view. This is what she did for most of the Missa10 pieces that were part of the ND dataset (where the DF was not activated) and she had no orange labels to mislead her in the correction process. For the Missa10.3_Credo2 of the D dataset the experimenter eventually decided to ignore the orange labels, treating the notes as either part of a chord or as doubling the agent.

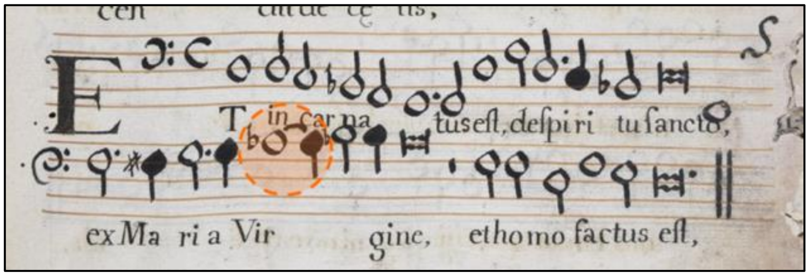


Figure 5-10: Bass voice of the Et incarnatus section of Missa10.3_Credo1 with a tied note expressing a note value impossible to represent in mensural notation (orange circle).

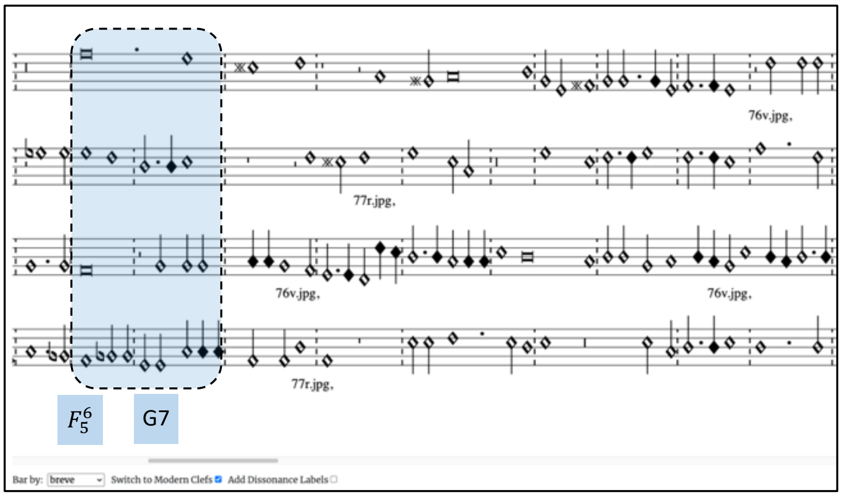


Figure 5-11: Dominant seventh chords pointed out by the experimenter in the Patrem omnipotentem section of Missa10.3_Credo1.
Note about clefs: *The superius is in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

Figure 5-12: Only illegal dissonance in Missa10.5_AgnusI0 that is not due to any of the false illegal dissonances categories (mentioned in Section 5.3.2.1), but to the presence of a dominant seventh chord, E4/2 (lacking the B).

Note about clefs: *The superius is in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

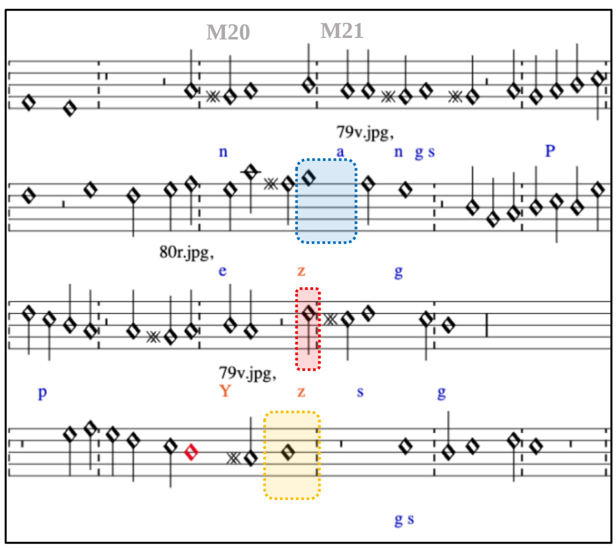


Figure 5-13: Passage in Missa10.3_Credo3 where the experimenter invested a lot of time. It is a combination of Baroque and Renaissance style, with seventh chords (measure 20) and fake suspensions (blue box).

Note about clefs: *The superius is in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

5.3.2.3 ***The DF Saves Time***

When comparing the average correction time reported for the *No Dissonance filter dataset* (ND dataset in Table 5-2) and the *Dissonance filter dataset* (D dataset in Table 5-3), it becomes apparent that the use of the dissonance filter (DF) reduces the time invested in correcting the results of the automatic voice-alignment. The reduction in time goes from 12.4 to 6.8 minutes (see Table 5-2 and Table 5-3), almost halving the correction time. However, the standard deviation in the D dataset is high (8.0) compared to the one in the ND dataset (4.2). This is due to the outlier piece in the D dataset, the eighteenth-century Missa10.3_Credo3, whose correction time was 30 minutes while the other pieces in the dataset had a correction time between 1–10 minutes. Details about mass 10, the only eighteenth-century mass in the corpus, have been discussed already in Section 5.3.2.2. Removing the outlier from the D dataset reduces the average correction time to 4.7 minutes (and the standard deviation to 3.4). Therefore, when removing the outlier, the DF reduces the correction time almost to a third. Moreover, now that we know to ignore the cases with rearticulations and doubled agents, the correction process using the DF could be even faster.[149]

For pieces with a similar amount of information (number of measures, voices, and illegal dissonances) that did not require any type of corrections—excluding mass 10 mentioned in the previous point—the reduction in time when using the DF is considerable (see Table 5-4).[150] The pieces shaded in blue in Table 5-4 have between 26–29 measures, but it took the experimenter one minute to check for counterpoint problems in the piece in the D dataset and 6 minutes to process

[149] Rearticulations confused the experimenter in Missa9.6_AgnusII0 and, most notably, for Missa10.3_Credo3. Doubled agents, while handled correctly by the experimenter, happened in Missa10.3_Credo1 and Missa10.3_Credo2, and in Piece21_Surrexit0 and Missa13.5_Agnus0 where the notes echoing the agent were also agents themselves.

[150] From the detailed analysis of the experiment results, one can see which pieces did not require any corrections (neither for scribal errors, OMR errors missed during the OMR stage, nor shortening/enlarging of ending notes). The analysis of the experiment results can be found in the Experiment branch of the GuatC1 GitHub repository at https://github.com/martha-thomae/GuatC1/blob/Experiment/Appendix%20C.pdf.

the one in the ND dataset. The pieces shaded in green have between 44–48 measures, but it took the experimenter one minute to process the one in the D dataset and 15 minutes to process the one in the ND dataset, despite the latter (the ND piece) having one voice less.

Table 5-4: Comparison of correction time between the D and ND datasets for pieces that do not need correction.

Dataset	Self-contained Units	Folios	Number of			Time (min)
			Measures	Voices	Illegal Disson	
D	*Missa7.4_Sanctus1*	38v-39r	29	4	0	1
ND	*Missa17.5_Agnus0*	179v-180r	26	4	1	6
D	*Missa15.1_Kyrie0*	133v-135r	44	4	0	1
ND	*Missa15.3_Credo3*	143v-144r	48	3	0	15

5.3.2.4 ***The DF Improves Accuracy***

By marking the illegal dissonances in orange, the DF allows users to increase the chances of making the right correction in the right place by allowing them (1) to focus their search for the error in the passage before the first orange label and (2) to experiment with different changes in that passage and observing how these edits affect the following dissonance labels until they find a satisfactory solution. An example of this is presented below.

There were numerous cases where the experimenter made a rhythmic change too late.[151] The DF would have shown her to look just before the first orange label to change the note value. This is shown in Figure 5-14 to Figure 5-16. Figure 5-14 shows the original reading of Missa15.3_Credo2 in the MP Editor with dissonance labels, which were not available to the experimenter for this piece since it belonged to the ND dataset. Looking at this score, the experimenter noticed that the cadence to G in measure 17 was not correct (yellow box), and she knew she had to cut a minim in the alto towards the beginning. Her correction, halving a semibreve

[151] This happens in Missa7.4_Sanctus2, Missa15.3_Credo2, and Missa16.4_Sanctus3.

in measure 14, is shown in Figure 5-15. However, the rhythmic change was supposed to happen three notes before (on measure 12) as shown in Figure 5-16. While the experimenter's correction in Figure 5-15 removes all following true illegal dissonances, it still leaves some illegal dissonance labels that precede it. On the other hand, the CPDL correction three notes before in Figure 5-16, removes all true illegal dissonances. The use of the DF would have facilitated identifying the best place to correct the rhythm.

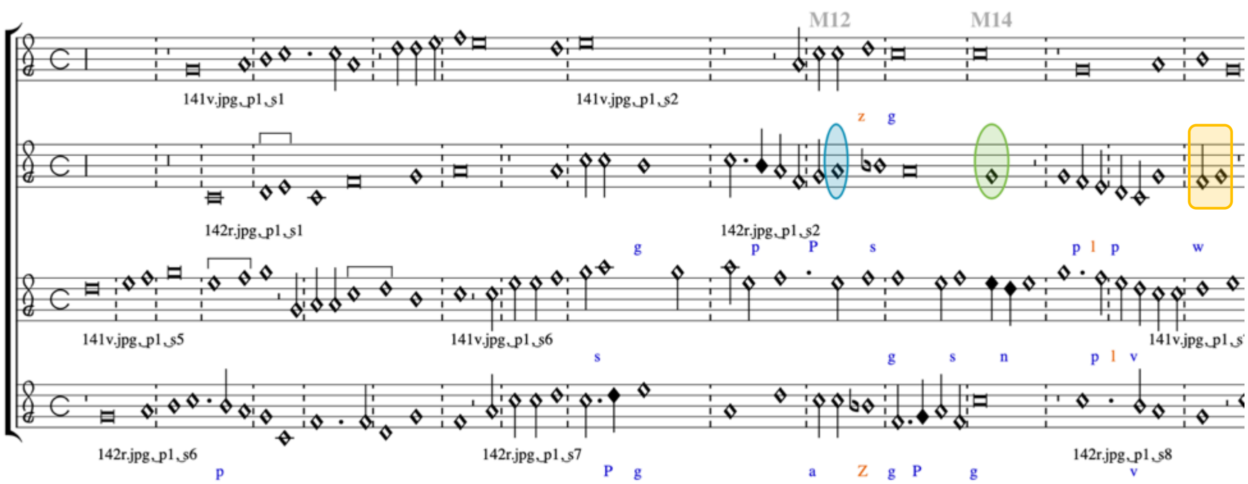


Figure 5-14: Original reading of the beginning of Missa15.3_Credo2. ***Note about clefs:*** *The superius and alto are in treble clef, while the tenor and bass are in suboctave treble clef.*

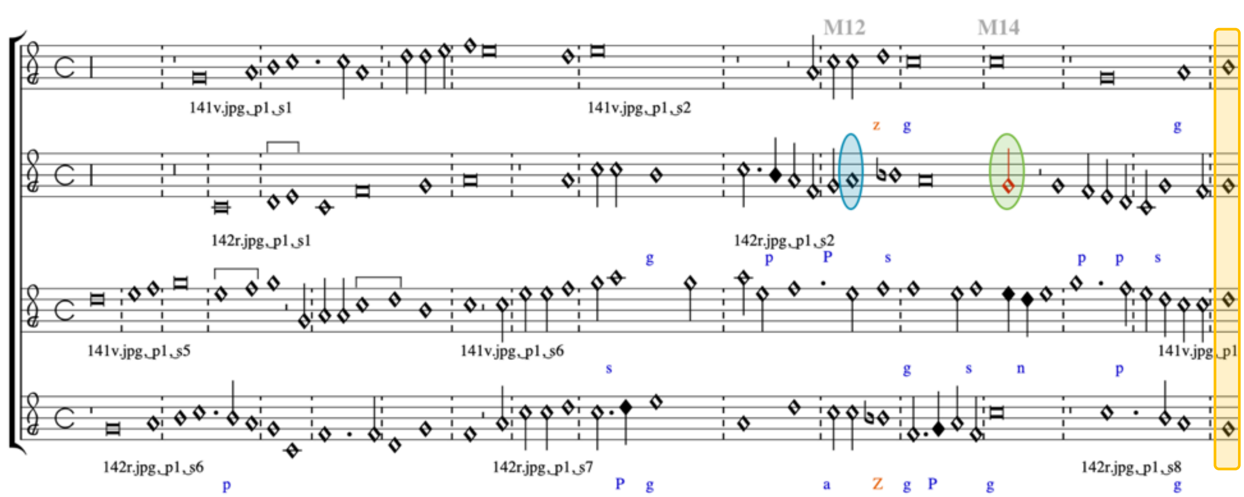


Figure 5-15: Experimenter's version of Missa15.3_Credo2, with her correction shown by the note marked in red. ***Note about clefs:*** *The superius and alto are in treble clef, while the tenor and bass are in suboctave treble clef.*

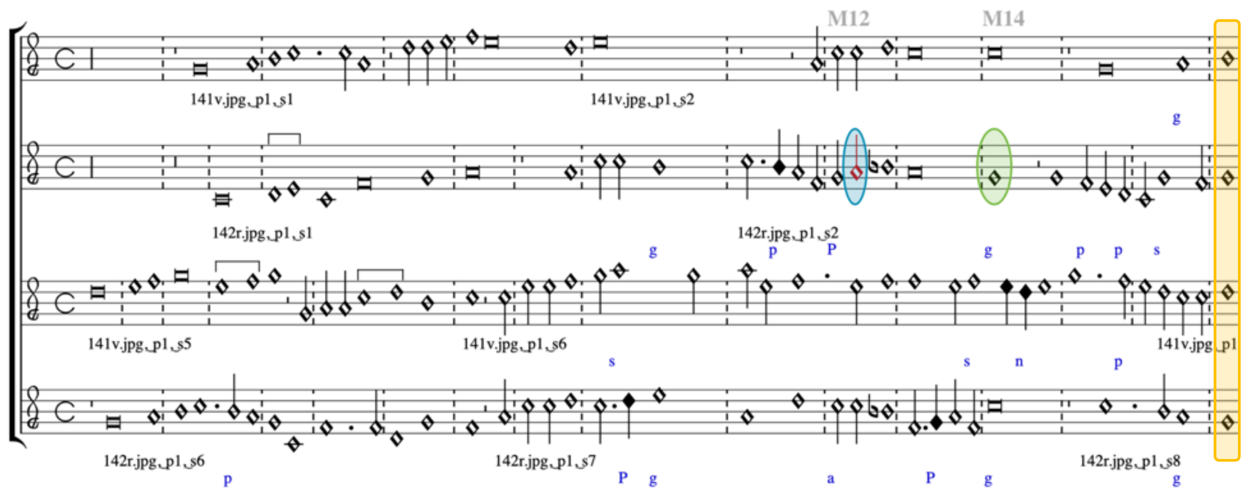
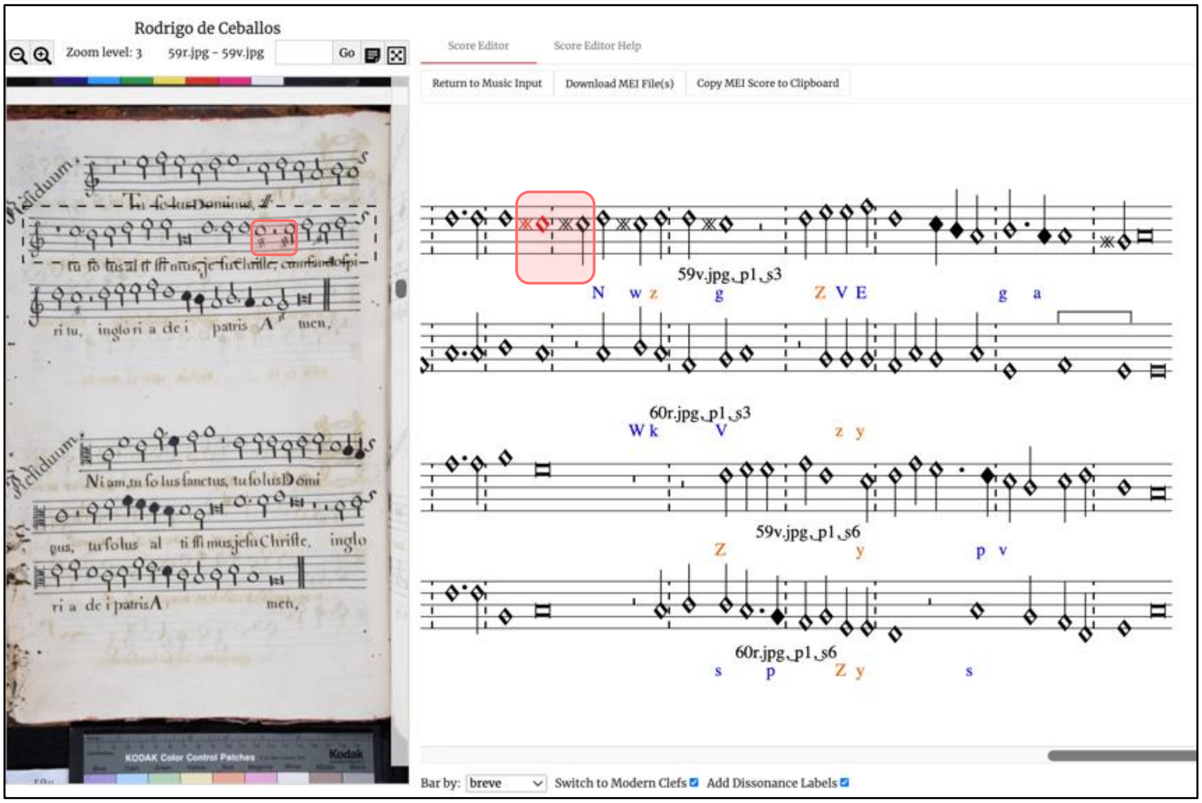


Figure 5-16: Correct version of Missa15.3_Credo2. The note marked in red shows the proper correction. ***Note about clefs:*** *The superius and alto are in treble clef, while the tenor and bass are in suboctave treble clef.*

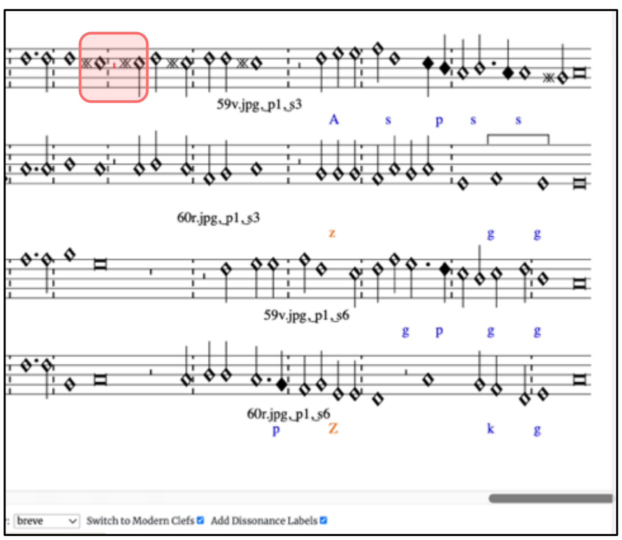
5.3.2.5 ***The DF Detects OMR Errors***

The DF was meant to help correct the results in the automatic voice-alignment stage of the encoding workflow, specifically looking at editorial corrections. However, it proved to be useful not only for detecting scribal errors, but also OMR errors, especially those that shifted the voice alignment. In a few pieces, the OMR misidentified a note (or rest) value in the manuscript or omitted a note (or rest) altogether.[152] Although these pieces were part of the ND dataset, processing them with the dissonance filter proved useful in spotting these errors quickly with the help of the orange labels. Figure 5-17 provides an example for Missa9.2_Gloria2. By finding the first orange label and checking the manuscript around that spot, the experimenter would have realized that a minim rest in the manuscript was omitted in the rendered file produced by the OMR process (see pink box in Figure 5-17a and b). The correction, adding the missing minim rest, removed all other “true” illegal dissonance labels as the only two remaining orange labels are the result of rearticulations.

[152] This happens in Missa9.2_Gloria1, Missa9.2_Gloria2, and Missa10.3_Credo1.



(a) Scored-up OMR result, before any experimenter's correction. The missing minim rest is supposed to lie between the two notes in the pink box (after the semibreve marked in red).



(b) Experimenter corrected file. The missing minim rest has been inserted; it is marked in red within the pink box.

Figure 5-17: Orange labels in Missa9.2_Gloria2 before and after the experimenter's correction of the omitted minim rest in the OMR stage. ***Note about clefs:*** *The superius and alto are in treble clef, while the tenor and bass are in suboctave treble clef.*

In addition to detecting OMR missed or misidentified notes/rests, the DF was able to detect some missing ties, an unforeseen consequence of its handling of rearticulations.[153] The OMR software (MuRET) does not handle ties in mensural notation, as they are not found before the seventeenth century in vocal music. However, it can be improved to include support for ties for later mensural sources.

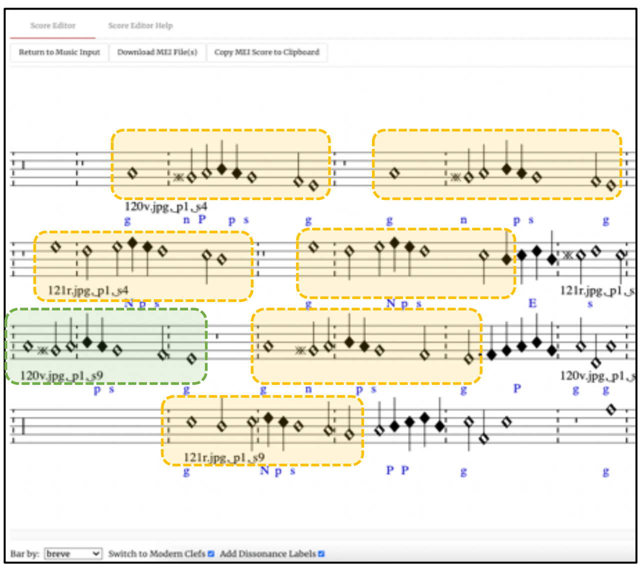
5.3.2.6 *Discussion of the Evaluation Method*

As indicated in Section 4.2 about the evaluation procedure, in addition to the elimination of orange labels, a good argument for an editorial change is if that change results in imitation—that is, recurring melodies (melodic intervals and rhythms) in different voices typical of sacred polyphony. There were various instances where the versions by the experimenter and by the CPDL transcriber did not coincide, and a good way to identify which solution was the best was by looking for imitation (see Figure 5-18) or for motivic consistency (see Figure 5-19). Figure 5-18 shows the experimenter version (Figure 5-18a) and the CPDL transcription (Figure 5-18b) of a passage from Missa14.1_Kyrie0. The experimenter kept the original notes of the manuscript (see green box). Notice that this sequence appears as a pattern in other voices as well (yellow boxes). The last notes in the green box are a third lower in the CPDL transcription (red box), which breaks the pattern seen in other voices (yellow boxes). Therefore, the CPDL correction was discarded as it contradicts the imitative texture of the passage. Figure 5-19 shows the experimenter version (top) and the CPDL transcription (bottom) of another passage of Missa14.1_Kyrie0. The experimenter kept the pitch of the note marked in red as it is in the manuscript, while the CPDL transcriber moved it a third down (see red box), contradicting the motivic pattern in the superius (shown by the yellow

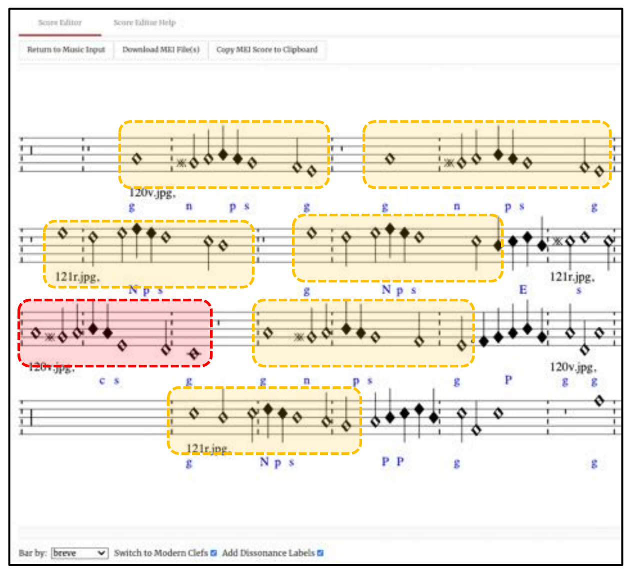
[153] This happens in Missa10.3_Credo1 and Missa10.3_Credo3.

boxes). Therefore, once again, the CPDL correction was discarded because it does not maintain motivic consistency.

Moreover, the process of re-uploading the experimenter's file and activating the DF to look for orange labels left by the experimenter allowed me to catch a scribal error that was missed by both the experimenter and the CPDL transcriber. This happened in Missa10.1_Kyrie0 (see Figure 5-20), where an orange 'Z' label below the alto's high A marks the undetected error. Favoring imitative textures for editorial corrections, allowed me to easily detect and correct the error, as moving this note down to a G results in imitation of the passages in the other voices marked in green boxes.



(a) Experimenter and OMR version.



(b) CPDL transcription.

Figure 5-18: Example of imitation. Comparison of the experimenter and OMR version (a) against the CPDL transcription (b) of part of the second Kyrie from Missa14.1_Kyrie0. The difference is in the last three notes of the green and red boxes. The CPDL edit (red box) breaks the imitative texture. ***Note about clefs:*** *The superius is in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

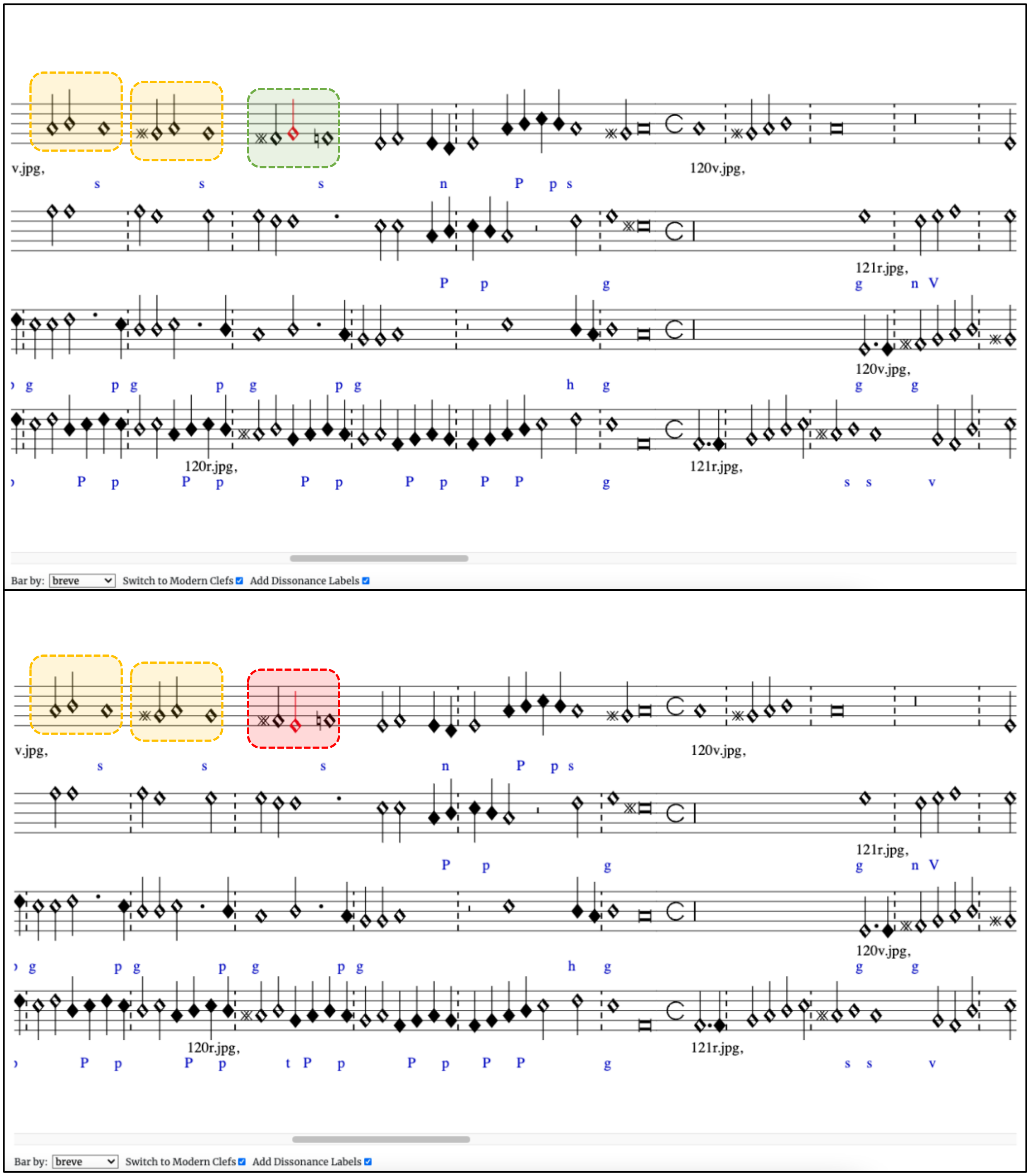


Figure 5-19: Example of motivic consistency. Comparison of the experimenter and OMR version (top) against the CPDL transcription (bottom) of the ending of the first Kyrie of Missa14.1_Kyrie0. The CPDL edit, shown by the red note, breaks motivic consistency. ***Note about clefs:*** *The superius is in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

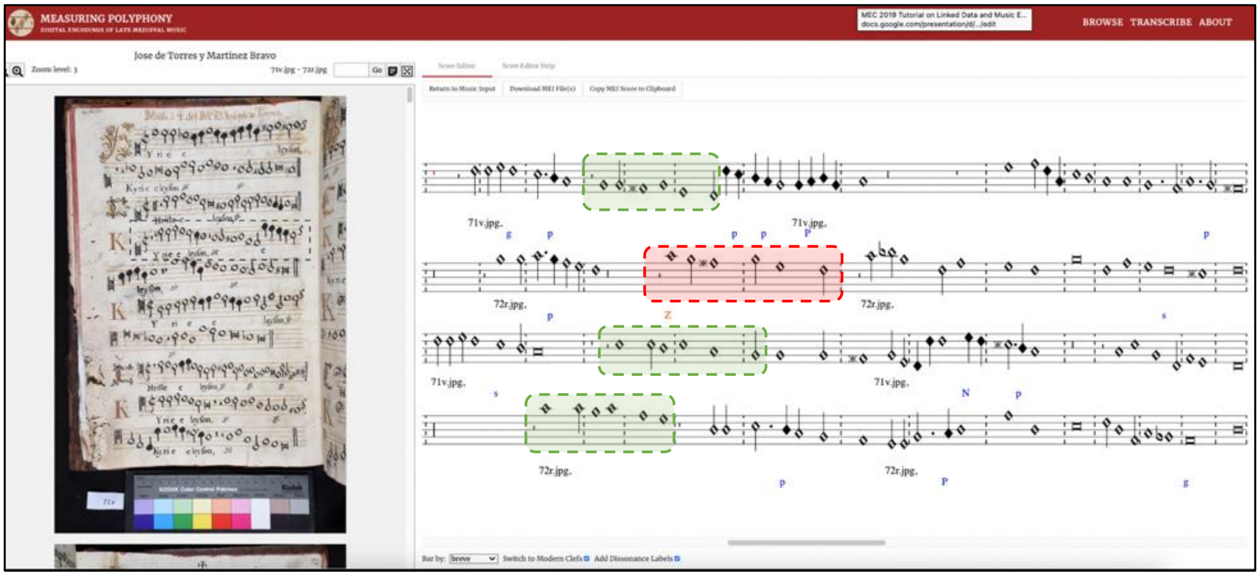


Figure 5-20: Example of an undetected error by the experimenter and CPDL transcriber in the Christe section (f.71v–72r) of Missa10.1_Kyrie0 with the DF turned on. ***Note about clefs:*** *The superius is in G clef, the alto and tenor are in suboctave G clefs, and the bass is in F clef.*

Chapter 6 Conclusions, Contributions, and Future Work

In this dissertation, I presented a digitization and music encoding workflow for obtaining high-resolution, color images and symbolic files encoding the corrected scores from books containing mensural notation. I designed this workflow to eventually apply it to the entire GuatC choirbooks, a collection of six choirbooks held at the Archivo Histórico Arquidiocesano de Guatemala (Archdiocesan Historical Archive of Guatemala, AHAG). In this dissertation, I focussed on the first choirbook of the collection, the book of masses GuatC 1. I digitized and encoded this choirbook as a proof of concept, testing the designed workflow and each of its technologies (see Figure 3-1 and Figure 4-1). This pilot project resulted in two outcomes: (1) ideas on planning for each of the stages of the workflow—we now know the approximate time consumed by each stage—and further improvements, and (2) digital image files and edited scores for all pieces of the GuatC 1. The GuatC 1 images are accessible through DIAMM,[154] and the MEI files encoding each of the shorter pieces and each of the sections of the individual mass movements can be found in the GuatC1 GitHub repository.[155] These MEI files can be rendered and played back by Verovio.[156]

The digitization and music encoding workflow has three stages that are performed by different tools (as shown in Figure 3-1 and Figure 4-1):

1. **Digitization stage:** Obtaining high-resolution digital images with a do-it-yourself (DIY) book scanner following the best practices outlined by various institutions and guidelines. The setup of the book scanner, the camera and lights that were used, and the details about

[154] https://www.diamm.ac.uk/sources/985/#/
[155] https://github.com/martha-thomae/GuatC1
[156] https://www.verovio.org/mei-viewer.xhtml

how to build the book cradle (with diagrams and dimensions) are provided in Section 3.1.3. The procedure followed and issues found during digitization are detailed in Section 5.1.

2. **Optical music recognition (OMR) stage:** Retrieving a Parts-based MEI file that encodes the music in separate-parts layout with information about the recognized symbols (shapes and pitches for notes) and the regions where they are in the image. This step was completed with the OMR framework MuRET.

3. **Automatic voice alignment (or scoring up of the voices) and editorial correction stage:** Retrieving the Score-based MEI file that encodes the music in score layout with information about the durational value of mensural notes (perfect / imperfect / altered encoded in the @dur.quality attribute of the <note> elements) and correction of scribal errors (through original and corrected readings encoded in <sic> and <corr> elements, respectively), both necessary for the correct alignment of the piece into a score.

The dissertation's final stage of the experiment focused on the last stage of the workflow, evaluating whether counterpoint error markers (specifically, illegal dissonances) make the correction of the voice alignment process more efficient. At this last stage, I expected very few OMR corrections to be needed as most of the OMR errors would have been manually corrected in the previous stage (stage 2) of the workflow using the MuRET OMR framework. Therefore, in this last stage, I mostly expected editorial corrections, in other words, corrections to scribal errors. Scribal errors are harder to spot when compared to OMR errors. Identifying scribal errors becomes easier when the piece is presented as a score, as one can use knowledge of modal counterpoint to detect places in the piece that do not conform to usual sonorities in Renaissance style, which raises the question: Can this be facilitated even further by marking some counterpoint errors in the scored-up piece? This is what I evaluated in this dissertation's final stage of the experiment. I marked the dissonances in the piece with special emphasis on illegal dissonances (uncommon for Renaissance style) and evaluated whether the highlighting of illegal dissonances helps the editor in the correction of the scored-up OMR file. This was done through the integration of humlib's Renaissance Dissonance Filter (DF) into the MP Editor. The DF labels the types of dissonances in the score and, in its integration into the MP Editor, it shows the legal dissonances in blue (e.g., passing tones, upper/lower, neighbours, and suspensions) and the illegal ones in orange. Other types of music-theory errors (e.g., parallel perfect intervals) could be explored in the future.

The dataset for the experiment consisted of files encoding sections of the masses and short pieces of the corpus, which were retrieved by performing OMR on the corresponding images. The files amounted to a 20% of the GuatC 1 corpus. To find out whether highlighting illegal dissonances (with the DF) was useful to correct the results of the scored-up OMR file in the MP Editor, I divided the dataset into two subsets with similar number of voices, measures, and illegal dissonances (for a fair comparison) and allowed the experimenter to use the DF for one of these subsets but not for the other one. The results prove that the DF made the correction process of Renaissance music more efficient by reducing the time invested in the correction of the voice alignment and increasing the accuracy of such corrections. Although the idea of using the DF in the scored-up version of the piece was to facilitate finding scribal errors, the DF also proved useful to identify OMR errors that were not corrected in the MuRET interface at the previous OMR stage of the workflow. A suggestion to use the DF to check for OMR errors in general—rather than correcting them manually on the previous stage of the workflow with the MuRET software—is tempting. However, OMR frameworks like MuRET, are optimized to correct OMR errors, as their interface allows to easily compare the recognized symbols against the original image. Moreover, there is a higher number of OMR errors in a piece than scribal errors. Therefore, I prefer handling the OMR correction in a dedicated OMR tool (like MuRET) than in the MP Editor, and to use the latter to render the piece as a score where I could spot the harder to identify (although less common) scribal errors.

A source of error during the experiment was the presence of false illegal dissonances—in other words, false positives—either due to **(1)** errors in the DF tool, or **(2)** the use of the DF in pieces that already show Baroque characteristics. The first case includes rearticulations, short notes doubling the agent, and fusas that were erroneously classified as illegal dissonances. This is something to correct in the DF implementation itself; once corrected, I expect the correction time and accuracy to increase as these false illegal dissonance labels would not distract the editor. Regarding the second case, the DF—which was developed for Renaissance music—proved to be sensitive to changes in style. Relaxation of some of the rules behind the DF could allow for handling pieces in the transition of Renaissance to Baroque style, and another type of tool could be used for later pieces.

The following three sections (6.1 to 6.3) contain conclusions regarding each of the workflow stages, summarizing what worked and what did not work on each stage and providing recommendations and updates that could improve the time invested in certain parts of the workflow.

6.1 Conclusions regarding the Digitization Stage

In this dissertation, I provided a realistic scenario for archives that do not have the resources for a digitization project—they neither have its own digitization and conservation departments, nor the equipment related to these activities. I kept the expenses to a minimum by hiring a professional photographer for only a couple sessions, having access to conservators from a non-profit institution, and designing a book scanner with built or borrowed parts. Thanks to the recommendations of many institutions, the digitization stage was completed successfully. However, after conducting this project, I would like to share some insights regarding what worked and what can be improved.

I would suggest providing the institution holding the physical item with an estimate of the time the digitization process will take as the room to use would not be available to other visitors during this process as the lighting conditions should be kept constant. Therefore, it is important to schedule ahead of time the room to be used, allowing some extra time for unexpected issues but being as prepared as possible for any issues. This preparation consists of testing the equipment and software to be used prior to digitization visit. For me this preparation took the form of testing the cradle configuration beforehand, which helped me to have an alternative configuration ready on site (without delaying the digitization process); and to have alternatives for the tethering app on my phone. However, further testing of the tethering app would have been useful. While I took advantage of the tethering software's remote shooting and preview functionalities, further exploration of the app would have allowed me to use some of its other functionalities: changing the file path where images are saved to store them directly on the external hard drive; adding starts to distinguish the best images; and adding comments to annotate the images (e.g., for verso pages and other folios missing the page number). During digitization, I ran into an unexpected issue with the tethering software as the app would stop working from time to time. While my alternative solution to switch into the phone tethering app worked, a few suggestions to avoid this issue in the future are to: (i) use the camera's official tethering software to avoid compatibility issues, (ii) connect the camera using a (long) data cable rather than wirelessly to avoid connectivity issues, and (iii) save the images on an external hard drive to avoid any memory issues generated by storing the images in the computer.

In addition to the recommendations about the scheduling of a designated room for digitization and the testing of the equipment and software beforehand to be prepared for issues—always having a plan B on hand—for optimal use of the time on site, I have a few recommendations regarding the digitization equipment and process itself. The use of the open-at-one-side configuration to keep the pages as flat as possible without the use of a platen worked very well for this manuscript, the same as the use of a stack of pieces of cardboard for keeping the color patch at the same height as the photographed page. The solutions for the issues presented when reaching the end of the book also worked well for the manuscript: (1) increasing the angle so that the book did not close abruptly and (2) lifting the book spine with conic pieces placed below it to counteract the effect of the gutter pulling the pages lying on the flat surface, which incremented the curvature of the photographed folio.

Other actions that resulted useful were to use book serpents (or snake weights) to hold the pages in place, use small pieces of paper with the printed verso numbers and include them at one side of the color patch so that the image itself included the information of the folio number for the versos, use a long data cable for connecting the camera to the computer or other device as the data cable provided with the camera tends to be small, and bring cleaning equipment for the lens since it is not usually provided with the camera. Some improvements could be to obtain more reference targets as they get scratched during the process and to cut them in smaller pieces to include both color and grey-scale targets on the image. As a personal suggestion, and to avoid the daily headaches of directly seeing the light panels after each day of work, it would be better to use lights with barndoors.

A final remark regarding the digitization expert. Although various digitization guidelines recommend that the digitization technician is trained in special collections and that they oversee the full process, this was not possible in the current scenario. Such an expert—with training on digitization and special collections—was not available; and, even if they were, hiring someone for the entire process would require a budget that neither the archive nor I had. The best solution was to hire a photographer with experience with oversize items (as he was involved in the digitization of a large map for the CIRMA archive), and only hire him for a few sessions, one session to set up the camera parameters and the other ones for questions or issues, and for me to handle the manuscript during digitization following experts' recommendations. Therefore, the professional photographer, considering that I was going to conduct the digitization process by myself, set the camera to perform certain actions automatically: white balancing, auto-focus, and shoot in Aperture-Priority Auto, where the camera automatically calculates the shutter speed based on the aperture value set by the photographer and the room conditions. To ensure that the values automatically calculated by the camera remained as constant as possible, I kept the conditions in the room constant with the room lights off, windows blocked, and the natural-light LED panels as the only source of light. While the ideal scenario would be for the digitization technician (with special collection training) to be the one to conduct the process and set the values of white balancing and other manual operations, projects with limited resources cannot adjust to this ideal scenario. Some recommendations based on the photographer's feedback include to use a macro lens (if available), as it is more suitable for this kind of job. And, to avoid being limited by the copystand arm on how far one can set the camera from the photographed item (since this distance determines the focal length value), one could use a combination of a sturdy tripod with a horizontal arm (shown in Figure 2-9).

6.2 Conclusions regarding the OMR Stage

There have been several improvements in MuRET since the writing of the Experiment chapter. Section 5.2.2 indicated the data used to train the document analysis and the holistic staff-level symbol recognition models for this experiment. It also indicated the common errors in the document analysis, agnostic transcription, and semantic transcription steps. According to a recent workshop on MuRET presented at the Music Encoding Conference (MEC) in May 2022, new models have been trained for both document analysis and holistic staff-level symbol recognition using all labelled examples of the GuatC 1 and a Zaragozan manuscript (both are very similar in their symbols' drawings). The performance of the new document analysis model is nearly perfect for the GuatC 1, which is expected since it has been trained with the full choirbook. Similarly, the performance of the new holistic staff-level model has improved significantly when compared to the one used in this dissertation, see outlined issues in page 211. Custos and F clefs, which were undetected by the holistic staff-level model used in this dissertation, as they were underrepresented in the training data (see page 208), are now successfully detected by the new holistic model, as the training data (the entire manuscript) counts with more examples of these symbols. The issue of fermatas—and other symbols that are stacked on top of another—being undetected by the holistic staff-level model is still present since this model was designed for monophonic staves and cannot handle stacked symbols (e.g., chords).

At present, the semantic transcription is nearly perfect as most of the issues pointed out in page 212 have been addressed. The user only needs to check the semantic sequence in a staff region when there are ligatures or accidentals present. Ligatures still need to be corrected all the time, but they are easier to correct with recent changes to the interface (see Figure 6-1). In the presence of accidentals, one needs to quickly check if they are being applied to the correct note—

since sometimes they are closer to the previous or following note and the agnostic-to-semantic transducer assigns them to the closest note. There are other two cases where one needs to be mindful of accidentals: (1) when the accident of a note is at the beginning of the system as it can be confused with a key signature in the semantic encoding, and (2) when an accidental is cancelling a preceding accidental or a key signature accidental as one needs to add the gestural information in the semantic encoding to show the real meaning of the written accidental. I believe these two cases could be handled by MuRET by implementing extra rules in the agnostic-to-semantic transducer, reducing the amount of work (i.e., corrections) on the user's side.

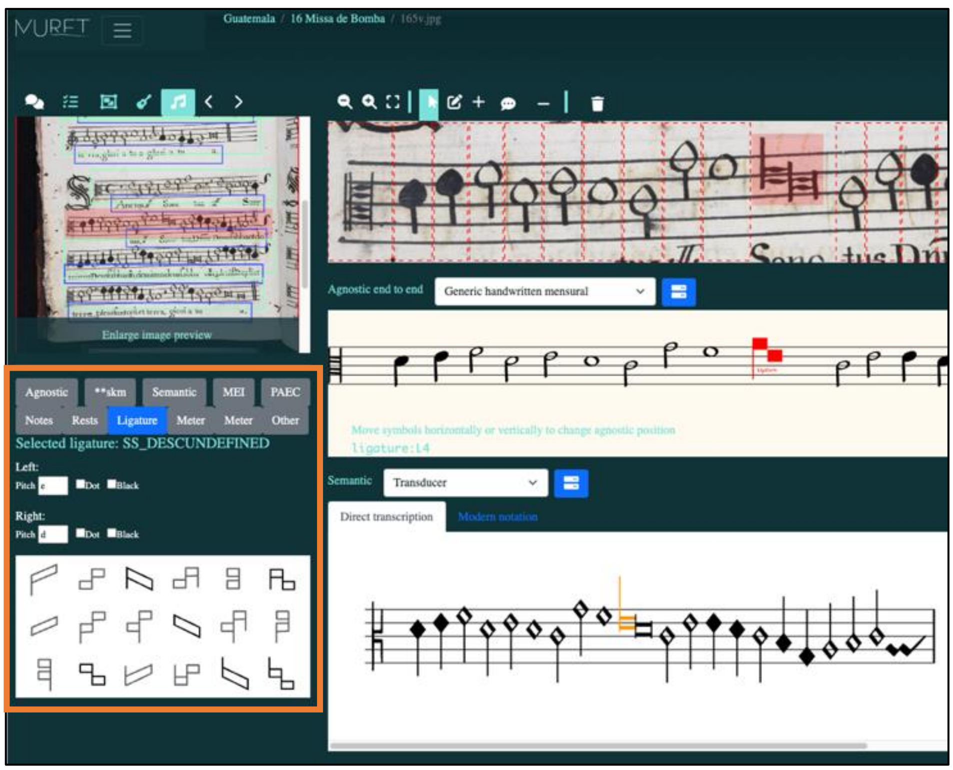


Figure 6-1: Semantic interface to correct two-notes ligatures (orange box). It provides support for correcting the shape of the ligature and the pitches of the two notes.

With all the recent improvements in MuRET and the new models for document analysis and agnostic transcription (trained with the full manuscript), the processing of one page through the full OMR process takes around seven minutes, less time than what was reported in Section 5.2

of the Experiment chapter. Therefore, processing manuscripts with a similar writing style to the one of GuatC 1 will be easier and take less time than the GuatC 1 did given that the models are ready and that their classification results require minimal correction.

6.3 Conclusions regarding the Voice-Alignment and Editorial Correction Stage

Insights of this last stage have already been summarized at the beginning of this chapter, indicating how the highlighting of illegal dissonances in the scored-up OMR file rendered in the MP Editor allows for a more efficient correction of the score by incrementing accuracy and reducing the correction time. The use of the DF in the MP Editor facilitated the correction of the score as it allowed the editor to spot the first orange label (the first illegal dissonance) in the score and focus their search for an (OMR or scribal) error at or before this point. The editor could quickly check that the notes in the passage preceding the first orange label match the notes in the manuscript to discard any OMR errors, which are highly improbable at this point since the MuRET user would have already corrected the OMR results; however, a few OMR errors went unnoticed in the experiment, but most of the errors corrected in the MP Editor were indeed scribal errors.

In addition to help reducing the error search area, the use of the DF in the MP Editor can assist the user on identifying the actual error and its solution by allowing them to experiment with the notes in the passage preceding the first orange label. The user can change the duration of the notes/rests in this passage or insert notes/rests and see the effect of these changes on the following dissonance labels, favoring the solution that results in the reduction of the orange (illegal dissonances) labels. In addition to reducing the number of illegal dissonances (orange labels) in the score, other paradigms for providing the best editorial correction include the reinforcement of motivic repetition and imitation.

Although I did not have ground truth for these scores, I had CPDL transcriptions for some of them that provided me with another musician's interpretation of the piece. An interesting result that confirms the usefulness of the DF was that the marking of the illegal dissonances allowed me to identify an error in one of the pieces of the corpus that was not identified by neither the experimenter nor the CPDL transcriber. The correction of this error resulted on reinforcing the imitative texture of the Renaissance piece.

All in all, the use of the DF helps focusing the search for errors—to the notes before or at the first orange label—and finding the right solution by allowing to change, add, remove note values and to see the effect this has on the following dissonance labels. This increases the accuracy and reduces the time of the correction process, making it more efficient. Further improvements in time and accuracy of the corrections could be obtained by improving how the DF handles the "false" illegal dissonances mentioned in Section 5.3.2.1.

6.4 Contributions

6.4.1 Original Contributions

There are two main contributions from this project: (1) the availability of a corpus as symbolic scores, and (2) the workflow used to retrieve these symbolic scores. The availability of the GuatC 1 as symbolic scores has several advantages because symbolic files allow for searchability, computational music analysis, preservation of a corpus on relatively small files, and access through machine-readable files that can be rendered in original notation, played back, and automatically transcribed into modern values. This corpus is particularly interesting for studying transmission of music from Europe to Latin America. It is possible to compare the encoded pieces to concordant sources and automatically find variant readings—as done for the Kyrie of Palestrina's mass in

Thomae (2021). In summary, the encoding of this corpus in symbolic files provides several advantages in terms of scholar study, preservation, and access.

Moreover, the workflow used to encode these symbolic scores is a contribution itself. This workflow allows for the semi-automatic transcription of digital images of mensural sources into scores, a layout more familiar for modern scholars than the original separate-parts arrangement of most mensural sources. The manual transcription of mensural music into scores is a time-consuming task; by making this process semi-automatic, this workflow allows the scholar to focus on the analysis task rather than the transcription (and encoding) task. The results of the workflow, the symbolic scores, are also beneficial for scholars given the advantages mentioned before, including searchability and computational musicological analysis. The presented digitization and encoding workflow can be applied to various Renaissance music sources, contributing to early music research.

I designed this workflow with the idea of using pre-existing tools developed for specific tasks (e.g., OMR, scoring up, or music analysis) and combine them together, trying to avoid re-inventing the wheel. For the OMR stage, I used the pre-existing OMR framework of MuRET. One could replace this by Aruspix if dealing with printed mensural notation, as long as the output of Aruspix conforms to the expected input by the MP Editor, summarized in Section 3.4. For the automatic voice alignment and editorial correction stage, I used the MP Editor, an online mensural notation editor that had support for these two functionalities. The work did not rely on developing a new monolithic tool but connecting the pre-existing tools.

Finally, I designed the digitization and encoding workflow to be accessible to everyone, including archives that lack the resources and personnel for this kind of project. With this goal in mind, I presented all the details about the DIY book scanner setup—which was designed with all parts built or borrowed—and only used free, online encoding tools (e.g., MuRET and the MP Editor). Regarding human power, I looked for non-profit initiatives and grants (when available) to support certain tasks (e.g., the conservation task), and hired a professional photographer only for the initial setup of the equipment. I propose that the encoding part—the OMR and editorial correction—can be undertaken by undergraduate or graduate students as part of their curriculum or profession practice. I hope the accessibility of this digitization and encoding method can empower archives and scholars all around the globe.

6.4.2 Additional Contributions

As indicated before, the main contributions of this dissertation are the digital image files and MEI files of the GuatC 1 pieces that help preserving and improving access to this corpus and the digitization and encoding methodology/workflow. With respect to the latter, there are three types of additional contributions enumerated in the following lines. These would not have been possible without the support of David Rizo (MuRET's developer), Juliette Regimbal (MP Editor's developer), and Craig Sapp (humlib's developer):

1. Development of a new tool:
    a. Development of a humlib tool for converting from **mens to **kern; this is, for converting Humdrum scores encoding mensural notation into Humdrum scores encoding common (Western) music notation. This was implemented by Sapp and me.
2. Improvement of the individual tools involved in the workflow:
    a. Improvement of MuRET's supports for the OMR mensural notation. I raised issues that were solved by Rizo.

b. Improvement of the MP Editor, both in terms of the symbols supported (e.g., allowing for mensuration signs and coloration to be entered by the user) and the scoring up functionality (e.g., the ability to handle changes in mensuration). I implemented these changes in the MP Editor's code with guidance from Regimbal.

c. Improvement of Humdrum's support for encoding mensural notation in **mens by adding support for: interpreting correctly imperfect, perfect, and altered values in terms of the length of the notes and a character for encoding alteration ('+'), encoding colored notes, encoding both the mensuration sign and its semantics separately (e.g., '*met(O2)_2322' for O2 mensuration sign and imperfect maximodus, perfect modus, imperfect tempus, and minor prolation), distinct characters for distinguishing between dots of division and augmentation (':' for dot of divison, and '.' for dot of augmentation). I led this work and Sapp implemented it.

d. Improvement of MEI and Verovio for supporting the encoding of the durational quality of notes in the @dur.quality attribute with values "imperfecta" / "perfecta" / "altera" and their rendering.

3. Interaction between these tools:

Avoided re-inventing the wheel by using existing tools, rather than developing new ones, and allow for these to interact with one another.

a. Allowed for the interaction of both MuRET and the MP Editor with the help of Rizo and Regimbal, and with the permission of Karen Desmond (MP Editor's principal investigator). I led the work with Regimbal and Rizo to implement the changes that allowed for the output of MuRET to be used as input of the MP Editor. This allowed us to combine the strengths of both tools, using the OMR to retrieve the music symbols rather than entering them manually in the MP Editor, and using the MP Editor's scoring up and editorial correction functionalities to obtain the corrected score.

b. Allowed for the use of humlib's Renaissance dissonance labeller within the score interface of the MP Editor. I led this work, implemented by Sapp and Regimbal, and tested by myself.

I tested the entire workflow, which is now available for encoding other pieces. All the tools involved in the workflow are meant to be accessible to places with limited resources as the encoding tools are free, open-source, and online, and the DIY scanning technology uses only already-built or borrowed parts.

The previously mentioned contributions—the preservation of a body of music and making it available for study, the design of the workflow with accessible technologies, and the development and improvement of the involved tools—help advance knowledge and research by providing access to a corpus and to a workflow that, in turn, can provide access to other music corpora. In addition to these "indirect" contributions to knowledge, I explored the use of computational counterpoint tools (in this case, humlib's Renaissance dissonance labeller) in the task of editorial correction of Renaissance music scores, and the results are promising.

6.5 Future Work

There are three aspects that I would like to explore more in the future: (1) conduct further research with the data retrieved from the GuatC 1: the images and MEI files; (2) continue exploring whether other types of music-theoretical errors, besides illegal dissonances, can help in detecting errors in the scored-up OMR files of a mensural piece; and (3) continue the work of digitization and encoding of the rest of the GuatC collection.

Regarding the first point, I would like to: (a) make the GuaC 1 images' content searchable based on the MEI files and (b) compare these pieces to other Latin American and European concordant sources. The MEI scores are machine-readable files and, therefore, their content is searchable. Moreover, they contain the coordinates of the music symbols encoded in the file. Therefore, using technologies like that of the *Cantus Ultimus* project, where the images are searchable by content in the Cantus Ultimus project website.[157] The user can search for individual symbols or pitches in a digitized manuscript and, using the coordinates of the symbols in the MEI file, bounding boxes are displayed to highlight where these symbols are in the images of the manuscript pages. The other plan for the GuatC 1 MEI files is to compare them to concordant sources. Given the European and Latin-American concordances for many of the GuatC 1 masses, I would also like to facilitate encoding the variant readings of these masses. As part of my future work, I would like to implement a 'diffing' tool that would take a set of Mensural MEI files encoding one piece from various concordant sources and produce a single Mensural MEI file that encodes the piece and its variant readings.

For the second point about future work, I would like to continue exploring whether other types of counterpoint rules (besides checking for illegal dissonances) can be applied for detecting errors in the scored-up OMR files of a Renaissance mensural piece. In this regard, in addition to illegal dissonances, it would be useful if the MP Editor could also show the presence of parallel perfect intervals (parallel fifths and octaves), places where the imitative texture or the motivic repetition of a piece is broken, and places where cadences are found. A filter for parallel perfect

[157] https://cantus.simssa.ca/manuscripts/

intervals should be relatively easy to implement in humlib; and regarding imitation and cadences, some work can be found in the *Citations: The Renaissance Imitation Mass* (*CRIM*) project.

Finally, I would like to continue my work on the preservation of the GuatC collection through digital images and symbolic edited scores, starting with requesting permission to the AHAG authorities to digitize the other choirbooks in the GuatC collection, supported by the successful results of this pilot project. I would use my DIY book scanner for GuatC 2, 4, and 5 since these are also large bound volumes. I would, however, use a simpler setup for GuatC 3 and 6, since GuatC 3 lost its binding and GuatC 6 is of small dimensions. Regarding the DIY book scanner, to avoid being restricted by a small copystand arm when deciding on the focal length to be used on the lens, I plan on exploring the use of either a copystand with a larger arm or a combination of a sturdy tripod with a horizontal arm. Regarding the OMR stage, I plan on using batch processing on the other books—a recent improvement in MuRET, still in the development version—which allows for applying the processes of document analysis, agnostic transcription, and semantic transcription to multiple pages at once. The user no longer has to click on every page to perform the document analysis and then click on every staff region to obtain the agnostic and semantic transcriptions, instead MuRET will be able to perform these three steps for all pages selected and the user will be left only with the correction task, saving lots of clicking time. I would also use the new document analysis and agnostic transcription models, which were trained with the entire GuatC 1 manuscript (as presented at the Music Encoding Conference in 2022), to classify the staff regions and the symbols in the other choirbooks of the collection (GuatC 2–6). Since GuatC 2–4 are the books in worst preservation condition, with missing sets of notes due to holes in the manuscript sources, I envision the Renaissance dissonance filter (DF) to be particularly useful to determine the missing notes. Because of these holes in the manuscripts, in addition to the editorial correction capabilities of the MP Editor—which encode the original and corrected readings in the <sic> and <corr> elements of the MEI file—I think that it would be useful to allow for encoding additions and gaps—using <add> and <gap> elements, respectively—to record where notes have been added in the source material or when there is a gap that the editor was not able to fill in but they know what the approximate duration of this gap is.

I hope this work can move us forward in the research of colonial music sources in the Americas and beyond, propelling the study of the transmission of Western music traditions across former European colonies and the identification of local traditions of counterpoint. I also hope that the technologies and workflow presented here empower small archives to digitize their music document collections, contributing to the preservation of the world's historic musical heritage.

References

Books, Articles, and Websites

Adobe Developers Association. 1992. “TIFF Revision 6.0.” Final. Mountain View, CA, US. https://adobeio-prod.adobemsbasic.com/content/dam/udp/en/open/standards/tiff/TIFF6.pdf.

Adobe Systems Incorporated. 2005. “Adobe RGB (1998) Color Image Encoding.” Version 2005-05.

Anderson, Sheila, Mike Pringle, Mick Eadie, Tony Austin, Andrew Wilson, and Malcolm Polfreman. 2006. “Digital Images Archiving Study.” Final draft. Arts and Humanities Data Service (AHDS). https://web.archive.org/web/20091118120100/http://ahds.ac.uk/about/projects/archiving-studies/digital-images-archiving-study.pdf.

Antila, Christopher. 2017. “Design Principles.” Documentation. VIS API (3.0.5). https://vis-framework.readthedocs.io/en/v3.0.5/about.html.

Antila, Christopher, and Julie Cumming. 2014. “The VIS Framework: Analyzing Counterpoint in Large Datasets.” In *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR)*, 71–76. Taipei, Taiwan.

Apel, Willi. 1953. *The Notation of Polyphonic Music 900–1600*. Fifth Edition. Cambridge, MA: The Medieval Academy of America.

Bainbridge, David, and Tim Bell. 2001. “The Challenge of Optical Music Recognition.” *Computers and the Humanities* 35 (2): 95–121.

BCR’s CPD Digital Imaging Best Practices Working Group. 2008. “BCR’s CDP Digital Imaging Best Practices.” Version 2.0. Bibliographical Center for Research - Collaborative Digitization Program (BCR CDP). https://mwdl.org/docs/digital-imaging-bp_2.0.pdf.

Bent, Margaret. 2001. "Notation, III, 3: Polyphonic Mensural Notation, C1260–1500." In *Grove Music Online*. Oxford University Press. http://www.oxfordmusiconline.com/subscriber/article/grove/music/20114pg6.

Besser, Howard. 2003. *Introduction to Imaging*. Getty Publications.

Boone, Graeme M. 1999. "Facsimiles - Oxford, Bodleian Library, MS. Canon. Misc. 213." *Notes,* Second Series, 55 (3): 762–65.

Burlet, Gregory, Alastair Porter, Andrew Hankinson, and Ichiro Fujinaga. 2012. "Neon.Js: Neume Editor Online." In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR)*, 121–26. Porto, Portugal.

Busse Berger, Anna Maria. 1993. *Mensuration and Proportion Signs: Origins and Evolution*. Oxford: Clarendon Press.

"Cabildo Catedral, Inventarios." 1633. Archivo Histórico Arquidiocesano de Guatemala (AHAG), Cabildo, Caja.

Call, Jerry, Charles Hamm, Herbert Kellman, and University of Illinois at Urbana-Champaign. Musicological Archives for Renaissance Manuscript Studies. 1979–1988. *Census-Catalogue of Manuscript Sources of Polyphonic Music, 1400-1550*. Vol. 1. 5 vols. Renaissance Manuscript Studies, 1. [n.p.]: American Institute of Musicology. https://bac-lac.on.worldcat.org/oclc/1032993823.

Calvo-Zaragoza, Jorge, Francisco J. Castellanos, Gabriel Vigliensoni, and Ichiro Fujinaga. 2018. "Deep Neural Networks for Document Processing of Music Score Images." *Applied Sciences* 8 (5): 654–74.

Calvo-Zaragoza, Jorge, Jan Hajič, and Alexander Pacha. 2020. "Understanding Optical Music Recognition." *ACM Computing Surveys (CSUR)* 53 (4): 1–35.

Calvo-Zaragoza, Jorge, and David Rizo. 2018a. "Camera-PrIMuS: Neural End-to-End Optical Music Recognition on Realistic Monophonic Scores." In *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, 248–55. Paris, France.

———. 2018b. "End-to-End Neural Optical Music Recognition of Monophonic Scores." *Applied Sciences* 8 (4): 606–29. https://doi.org/10.3390/app8040606.

Calvo-Zaragoza, Jorge, David Rizo, and José M. Iñesta. 2016. "Two (Note) Heads Are Better Than One: Pen-Based Multimodal Interaction with Music Scores." In *Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR)*, 509–14.

Calvo-Zaragoza, Jorge, Alejandro H. Toselli, and Enrique Vidal. 2019. "Handwritten Music Recognition for Mensural Notation with Convolutional Recurrent Neural Networks." *Pattern Recognition Letters* 128: 115–21. https://doi.org/10.1016/j.patrec.2019.08.021.

Calvo-Zaragoza, Jorge, Gabriel Vigliensoni, and Ichiro Fujinaga. 2017. "One-Step Detection of Background, Staff Lines, and Symbols in Medieval Music Manuscripts with Convolutional Neural Networks." In *Proceedings of the 18th International Society for Music Information Retrieval Conference (ISMIR)*, 724–30. Suzhou, China.

Calvo-Zaragoza, Jorge, Ké Zhang, Zeyad Saleh, Gabriel Vigliensoni, and Ichiro Fujinaga. 2017. "Music Document Layout Analysis through Machine Learning and Human Feedback." In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 02:23–24. Tokyo, Japan. https://doi.org/10.1109/ICDAR.2017.259.

Castellanos, Francisco J., Jorge Calvo-Zaragoza, and José M. Iñesta. 2020. "A Neural Approach for Full-Page Optical Music Recognition of Mensural Documents." In *Proceedings of the 21st International Society for Music Information Retrieval Conference (ISMIR)*, 23–27. Montréal, QC, Canada.

Castellanos, Francisco J., Jorge Calvo-Zaragoza, Gabriel Vigliensoni, and Ichiro Fujinaga. 2018. "Document Analysis of Music Score Images with Selectional Auto-Encoders." In *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, 256–63. Paris, France. http://cloud.simssa.ca/index.php/s/FjJGQ6josKEIWNn.

Catalyne, Alice Ray, and Mark Brill. 2001. "Franco, Hernando." Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.10132.

"Cathedral." 2021. In *Oxford English Dictionary (OED)*. Oxford University Press. https://www.oed.com/view/Entry/28939?rskey=vdalqO&result=1#eid.

Chapman, Stephen. 2000. "Considerations for Project Management." In *Handbook for Digital Projects: A Management Tool for Preservation and Access*, edited by Maxine K. Sitts, first edition. Andover, Massachusetts: Northeast Document Conservation Center.

Chew, Geoffrey, and Richard Rastall. 2001. "Notation, III, 4: Mensural Notation from 1500." In *Grove Music Online*. Oxford University Press. http://www.oxfordmusiconline.com/subscriber/article/grove/music/20114pg7.

Choudhury, G. Sayeed, Cynthia Requardt, Ichiro Fujinaga, Tim DiLauro, Elizabeth W. Brown, James W. Warner, and Brian Harrington. 2000. "Digital Workflow Management: The Lester S. Levy Digitized Collection of Sheet Music." *First Monday* 5 (6).

CompuServe Incorporated. 1990. “Graphics Interchange Format: Version 89a.” Columbus, OH. https://www.w3.org/Graphics/GIF/spec-gif89a.txt.

Condit-Schultz, Nathaniel, and Claire Arthur. 2019. “HumdrumR: A New Take on an Old Approach to Computational Musicology.” In *Proceedings of the 20th International Society for Music Information Retrieval Conference (ISMIR)*, 715–22. Delft, Netherlands. https://archives.ismir.net/ismir2019/paper/000087.pdf.

Coros de Catedral. 1995. Compact disc. Vol. 2. Música Histórica de Guatemala. Guatemala. Fundación para la Cultura y el Desarrollo.

Craig-McFeely, Julia. 2008. “Digital Image Archive of Medieval Music: The Evolution of a Digital Resource.” *Digital Medievalist* 3. https://doi.org/10.16995/dm.16.

Cumming, Julie E. 2013. “From Two-Part Framework to Movable Module.” *Medieval Music in Practice. Studies in Honor of Richard Crocker. Münster: American Institute of Musicology*, 177–215.

Cumming, Julie E., and Cory McKay. 2021. “Using Corpus Studies to Find the Origins of the Madrigal.” In *Proceedings of the Conference Future Directions of Music Cognition*. https://doi.org/10.18061/FDMC.2021.0005.

Cuthbert, Michael Scott, and Christopher Ariza. 2010. “Music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data.” In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR)*, 637–42. Utrecht, Netherlands.

Cuthbert, Michael Scott, Christopher Ariza, and Lisa Friedland. 2011. “Feature Extraction and Machine Learning on Symbolic Music Using the Music21 Toolkit.” In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, 387–92. Miami, FL, USA. http://web.mit.edu/music21/papers/Cuthbert_Ariza_Friedland_Feature-Extraction_ISMIR_2011.pdf.

Cuthbert, Michael Scott, and cuthbertLab. 2021. “Music21 Documentation.” Documentation. https://web.mit.edu/music21/doc/index.html.

Dalitz, Christophe, and Thomas Karsten. 2005. “Using the Gamera Framework for Building a Lute Tablature Recognition System.” In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*, 478–81. London, UK.

DeFord, Ruth I. 2015. *Tactus, Mensuration and Rhythm in Renaissance Music*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781107587717.

Desmond, Karen, Andrew Hankinson, Laurent Pugin, Juliette Regimbal, Craig Sapp, and Martha E. Thomae. 2020. “Next Steps for Measuring Polyphony: A Prototype Editor for Encoding Mensural Music.” In *Music Encoding Conference Proceedings 26–29 May, 2020*, 121–24. Tufts University, Boston. http://dx.doi.org/10.17613/5k88-9z02.

Desmond, Karen, Emily Hopkins, Samuel Howes, and Julie E. Cumming. 2020. “Computer-Aided Analysis of Sonority in the French Motet Repertory, ca. 1300-1350.” *Music Theory Online* 26 (4).

Desmond, Karen, Laurent Pugin, Juliette Regimbal, David Rizo, Craig Sapp, and Martha E. Thomae. 2021. “Encoding Polyphony from Medieval Manuscripts Notated in Mensural Notation.” In *Proceedings of the Music Encoding Conference*, edited by Stefan Münnich and David Rizo, 197–219. Alicante, Spain (online): Humanities Commons. https://doi.org/10.17613/tf2j-x697.

Digital Preservation Coalition. 2015. *Digital Preservation Handbook*. 2nd ed. https://www.dpconline.org/handbook.

"Diocese." 2021. In *Oxford English Dictionary (OED)*. Oxford University Press. https://www.oed.com/view/Entry/53084?redirectedFrom=diocese#eid.

Drake, Karl Magnus, Borje Justrell, and Anna Maria Tammaro. 2003. "Good Practice Handbook." Version 1.2. Minerva. https://www.academia.edu/download/66703355/bestpracticehandbook1_2.pdf.

Droettboom, Michael, Ichiro Fujinaga, Karl MacMillan, G. Sayeed Choudhury, Tim DiLauro, Mark Patton, and Teal Anderson. 2002. "Using the Gamera Framework for the Recognition of Cultural Heritage Materials." In *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*, 11–17.

Droettboom, Michael, Karl MacMillan, and Ichiro Fujinaga. 2003. "The Gamera Framework for Building Custom Recognition Systems." In *Symposium on Document Image Understanding Technologies*, 275–86.

Eadie, Michael. 2005. "Preservation Handbook: Bitmap (Raster) Images." Version 1.3. Preservation Handbooks. Arts and Humanities Data Service (AHDS). https://web.archive.org/web/20150921145430/http://www.ahds.ac.uk/preservation/Bitmap-preservation-handbook_d6.pdf.

Federal Agencies Digital Guidelines Initiative. 2016. "Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files." http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final_rev1.pdf.

Federal Agencies Digitization Guidelines Initiative. 2022. "Glossary: JPEG 2000." https://web.archive.org/web/20220621025912/https://www.digitizationguidelines.gov/term.php?term=jpeg2000.

Fiorentino, Giuseppe. 2016. "Unwritten Music and Oral Traditions at the Time of Ferdinand and Isabel." In *Companion to Music in the Age of the Catholic Monarchs*, edited by Tess Knighton, 1:504–48. Brill's Companions to the Musical Culture of Medieval and Early Modern Europe. Boston: Brill. https://doi.org/10.1163/9789004329324.

Fujinaga, Ichiro, and Andrew Hankinson. 2015. "Single Interface for Music Score Searching and Analysis (SIMSSA)." In *Proceedings of the First International Conference on Technologies for Music Notation and Representation*. Paris, France.

Fujinaga, Ichiro, Andrew Hankinson, and Julie Cumming. 2014. "Introduction to SIMSSA (Single Interface for Music Score Searching and Analysis)." In *Proceedings of the International Conference on Digital Libraries for Musicology (DLfM)*, 100–102. London, UK.

Garfinkle, David, Peter Schubert, Claire Arthur, Julie Cumming, and Ichiro Fujinaga. 2017. "PatternFinder: Content-Based Music Retrieval with Music21." In *Proceedings of the 4th International Workshop on Digital Libraries for Musicology*. Shanghai, China.

Gembero Ustárroz, María. 2005. "El compositor español Hernando Franco (1532-85) antes de su llegada a México: Trayectoria profesional en Portugal, Santo Domingo, Cuba y Guatemala." *Latin American Music Review* 26 (2): 273–317.

Good, Michael. 2001. "MusicXML: An Internet-Friendly Format for Sheet Music." In *XML Conference Proceedings*. Orlando, FL. http://michaelgood.info/publications/music/musicxml-an-internet-friendly-format-for-sheet-music/.

Goodliffe, Jonathan. 2020. “Música Colonial Archive.” Wiki. Free Choral Music. https://www.cpdl.org/wiki/index.php/M%C3%BAsica_Colonial_Archive.

Hankinson, Andrew. 2014. “Optical Music Recognition Infrastructure for Large-Scale Music Document Analysis.” PhD diss., Montreal, Canada: Schulich School of Music, McGill University.

Hankinson, Andrew, Perry Roland, and Ichiro Fujinaga. 2011. “The Music Encoding Initiative as a Document-Encoding Framework.” In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, 293–98. Miami, FL, USA.

Helsen, Kate, and Inga Behrendt. 2016. “The Optical Neume Recognition Project.” In *Proceedings of the International Winterschool Digital Musicology – Digitalisierung in Der Musikwissenschaft*. Musikwissenschaftliches Institut, Ebernhard Karls Universität Tübingen. http://dx.doi.org/10.15496/publikation-24002.

Hiley, David. 2001. “Bar.” In *Grove Music Online*. Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.01972.

Huang, Yu-Hui, Xuanli Chen, Serafina Beck, David Burn, and Luc Van Gool. 2015. “Automatic Handwritten Mensural Notation Interpreter: From Manuscript to MIDI Performance.” In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, 79–85. Málaga, Spain.

Huron, David. 1997. “Humdrum and Kern: Selective Feature Encoding.” In *Beyond MIDI: The Handbook of Musical Codes*, edited by Eleanor Selfridge-Field, 375–401. Cambridge, MA: MIT Press. http://beyondmidi.ccarh.org/beyondmidi-600dpi.pdf.

———. 1999. “Music Research Using Humdrum: A User’s Guide.” *Center for Computer Assisted Research in the Humanities, Stanford, California*. http://ccarh.org/publications/manuals/humdrumuserguide/humdrum-users-guide-1998.pdf.

“Image Capture.” 2021. DIAMM. https://web.archive.org/web/20211006223046/https://www.diamm.ac.uk/about/technical-overview/image-capture/.

“Image Quality.” 2021. DIAMM. https://web.archive.org/web/20211006184205/https://www.diamm.ac.uk/about/technical-overview/image-quality/.

“Imaging.” 2021. DIAMM. https://web.archive.org/web/20211006182243/https://www.diamm.ac.uk/services/imaging/.

Iñesta, José M., Pedro J. Ponce de León, David Rizo, José Oncina, Luisa Micó, Juan Ramón Rico-Juan, Carlos Pérez-Sancho, and Antonio Pertusa. 2018. “Hispamus: Handwritten Spanish Music Heritage Preservation by Automatic Transcription.” In *Proceedings of the 1st International Workshop on Reading Music Systems (WoRMS)*, 17–18. Paris, France.

Iñesta, José M., David Rizo, and Jorge Calvo-Zaragoza. 2019. “MuRET as a Software for the Transcription of Historical Archives.” In *Proceedings of the 2nd International Workshop on Reading Music Systems (WoRMS)*, 12–15. Delft, Netherlands.

JPEG. 2022. “JPEG 1.” JPEG. https://web.archive.org/web/20220621090421/https://jpeg.org/jpeg/.

Kelly, Thomas Forrest. 2015. *Capturing Music: The Story of Notation*. New York, NY: W. W. Norton & Company.

Kolb, Paul. 2017. “Divisions of Dots and Dots of Division: History, Theory, and Practice.” *Tijdschrift van de Koninklijke Vereniging Voor Nederlandse Muziekgeschiedenis* 67 (1/2): 177–90.

Lehnhoff, Dieter. 1984. *Música de la época colonial en Guatemala: primera antología*. Antigua Guatemala, Guatemala: Centro de Investigaciones Regionales Mesoamericanas.

———. 1986. *Espada y pentagrama: la música polifónica en la Guatemala del siglo XVI*. Guatemala City: Centro de Reproducciones, Universidad Rafael Landívar.

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

———. 2001. *Las misas de Pedro Bermúdez*. Guatemala City: Universidad Rafael Landívar, Instituto de Musicología.

———, ed. 2008. *Choral Music from Guatemala: For SATB Choir*. Niedernhausen: Edition Kemel.

Lemmon, Alfred E. 1984. “Las obras musicales de dos compositores guatemaltecos del siglo XVIII: Rafael Antonio Castellanos y Manuel José de Quiroz.” *Mesoamérica* 5 (8): 389–401.

———. 1993. “Toward an International Inventory of Colonial Spanish American Cathedral Music Archives.” *Revista de Musicología* 16 (1): 92–98. https://doi.org/10.2307/20795878.

Library of Congress. 2021a. “JPEG 2000 Part 1, Core Coding System.” Sustainability of Digital Formats: Planning for Library of Congress Collections. https://www.loc.gov/preservation/digital/formats/fdd/fdd000138.shtml.

———. 2021b. “JPEG 2000 Part 1 (Core) Jp2 File Format.” Sustainability of Digital Formats: Planning for Library of Congress Collections. https://www.loc.gov/preservation/digital/formats/fdd/fdd000143.shtml.

———. 2021c. “TIFF, Revision 6.0.” Sustainability of Digital Formats: Planning for Library of Congress Collections. https://www.loc.gov/preservation/digital/formats/fdd/fdd000022.shtml.

“Libro de Actas Capitulares 1.” 1531. Archivo Histórico Arquidiocesano de Guatemala (AHAG).

MacMillan, Karl, Michael Droettboom, and Ichiro Fujinaga. 2001. “Gamera: A Python-Based Toolkit for Structured Document Recognition.” In *Proceedings of the Tenth International Python Conference*.

McIlwaine, John, Jean-Marc Comment, Clemens de Wolf, Dale Peters, Borje Justrell, Marie-Thérèse Varlamoff, and Sjoerd Koopman. 2002. “GUIDELINES FOR DIGITIZATION PROJECTS for Collections and Holdings in the Public Domain, Particularly Those Held by Libraries and Archives.” Guidelines. International Federation of Library Associations and Institutions (IFLA). http://repository.ifla.org/handle/123456789/697.

McKay, Cory. 2018. “Performing Statistical Musicological Research Using JSymbolic and Machine Learning.” Presented at the International Conference on the Anatomy of Polyphonic Music around 1500, Cascais, Portugal, June. http://jmir.sourceforge.net/publications/mckay18performing.pdf.

McKay, Cory, and María Elena Cuenca. 2021. "Musical Influences on the Masses and Motets of Cristóbal de Morales and Francisco Guerrero: A Statistical Approach." Presented at 49th Medieval and Renaissance Music Conference, Universidade NOVA de Lisboa, Portugal, July. http://jmir.sourceforge.net/publications/mckay21musical.pdf.

McKay, Cory, Julie Cumming, and Ichiro Fujinaga. 2017. "Characterizing Composers Using JSymbolic2 Features." In *Extended Abstracts for the Late-Breaking Demo Session of the 18th International Society for Music Information Retrieval Conference*. Suzhou, China. http://jmir.sourceforge.net/publications/mckay17characterizing.pdf.

———. 2018. "JSymbolic 2.2: Extracting Features from Symbolic Music for Use in Musicological and MIR Research." In *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, 348–54.

McKay, Cory, Tristano Tenaglia, Julie Cumming, and Ichiro Fujinaga. 2017. "Using Statistical Feature Extraction to Distinguish the Styles of Different Composers." In *Abstracts of the Annual International Medieval and Renaissance Music Conference*, 111. Prague, Czech Republic: MedRen. http://jmir.sourceforge.net/publications/mckay17using.pdf.

Morales Abril, Omar. 2003. "'Jesu nostra redemptio', himno de vísperas para la Ascensión, de Pedro Bermúdez." *Heterofonía: revista de investigación musical*, no. 129: 109–27.

———. 2007. "Características de estilo en la obra de Pedro Bermúdez (fl. 1574-1604)." *Revista de Musicología* 30 (2): 343–91. https://doi.org/10.2307/20797891.

———. 2013. "Gaspar Fernández: su vida y obras como testimonio de la cultura musical novohispana a principios del siglo XVII." *Ejercicio y enseñanza de la música en México*, 71–125.

———. 2015. “Música local y música foránea en los libros de polifonía de las Catedrales de Guatemala y México.” *Jahrbuch für Renaissancemusik: troja.* Vokalpolyphonie zwischen Alter und Neuer Welt: Musikalische Austauschprozesse zwischen Europa und Lateinamerika im 16. und 17. Jahrhundert (14): 95–124.

———. unpublished. *Catálogo de los acervos musicales del Archivo Histórico Arquidiocesano de Guatemala*.

Morales, Omar, Jorge Pellecer, Igor de Gandarias, and Arturo Duarte. 2001. *Missa de bomba a 4: Guatemala, siglo XVI*. Guatemala City: ADESCA.

Morgan, Alexander. 2017. “Renaissance Interval-Succession Theory: Treatises and Analysis.” PhD diss., Montreal, Canada: Schulich School of Music, McGill University.

Müller, Meinard. 2015. *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. Switzerland: Springer.

Murray, James D., and William Van Ryper. 2008. “GIF: Summary from the Encyclopedia of Graphics File Formats (Second Edition).” FileFormat.Info: The Digital Rosetta Stone. https://web.archive.org/web/20080312145216/http://www.cs.albany.edu/~sdc/CSI333/Fal07/Lect/L18/Summary.html.

Music Encoding Initiative. 2021. “Tools: MEI to Music21 Converter.” Resources website. Music Encoding Initiative. April 27, 2021. https://music-encoding.org/resources/tools.html#mei-to-music21-converter.

National Library of Australia. 2022. “Image Capture Standards.” Digitisation Guidelines. https://web.archive.org/web/20220619232223/https://www.nla.gov.au/about-us/corporate-documents/policy-and-planning/standards/digitisation-guidelines/image-capture.

OCLC. 2021. “Preparing Digital Surrogates for RLG Cultural Materials.” OCLC Research. https://www.oclc.org/research/archive/projects/culturalmaterials/surrogates.html.

Owens, Jessie Ann. 1997. *Composers at Work: The Craft of Musical Composition 1450–1600*. New York, NY: Oxford University Press.

Pacha, Alexander, and Jorge Calvo-Zaragoza. 2018. “Optical Music Recognition in Mensural Notation with Region-Based Convolutional Neural Networks.” In *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, 23–27. Paris, France.

Parrish, Carl. 1978. *The Notation of Medieval Music*. New York, NY: Pendragon Press.

Pérez de Antón, Francisco. 1999. *In Praise of Francisco Marroquín*. Universidad Francisco Marroquín.

Pugin, Laurent. 2006. “Optical Music Recognition of Early Typographic Prints Using Hidden Markov Models.” In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*, 53–56. Victoria, BC, Canada.

———. 2009. “Editing Renaissance Music: The Aruspix Project.” In *Digitale Edition zwischen Experiment und Standardisierung*, edited by Peter Stadler and Joachim Veit, 147–56. De Gruyter. https://doi.org/10.1515/9783110231144.147.

———. 2018. “Aruspix and the Marenzio Online Digital Edition.” *Troja. Jahrbuch Für Renaissancemusik* 17: 103–23.

Pugin, Laurent, and Tim Crawford. 2013. “Evaluating OMR on the Early Music Online Collection.” In *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*, 439–44. Curitiba, Brazil.

Pugin, Laurent, Jason Hockman, John Ashley Burgoyne, and Ichiro Fujinaga. 2008. "Gamera versus Aruspix: Two Optical Music Recognition Approaches." In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR)*, 419–24. Philadelphia, PA, USA.

Pujol, David. 1965. "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* 20: 3–10.

Rastall, Richard. 2001. "Tie." In *Grove Music Online*. Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.27937.

Rebelo, Ana, G. Capela, and Jaime S. Cardoso. 2010. "Optical Recognition of Music Symbols." *International Journal on Document Analysis and Recognition (IJDAR)* 13 (1): 19–31.

Rebelo, Ana, Ichiro Fujinaga, Filipe Paszkiewicz, Andre R. S. Marcal, Carlos Guedes, and Jaime S. Cardoso. 2012. "Optical Music Recognition: State-of-the-Art and Open Issues." *International Journal of Multimedia Information Retrieval* 1 (3): 173–90.

Recinos, Adrian. 1980. *Memorial de Sololá. Anales de Los Cakchiqueles. Título de Los Señores de Totonicapán*. Guatemala: Editorial Piedra Santa.

Regimbal, Juliette, Zoé McLennan, Gabriel Vigliensoni, Andrew Tran, and Ichiro Fujinaga. 2019. "Neon2: A Verovio-Based Square-Notation Editor." Presented at the Music Encoding Conference (MEC). Vienna, Austria. https://music-encoding.org/conference/abstracts/abstracts_mec2019/Neon2.pdf.

Regimbal, Juliette, Gabriel Vigliensoni, Caitlin Hutnyk, and Ichiro Fujinaga. 2020. “IIIF-Based Lyric and Neume Editor for Square-Notation Manuscripts.” In *Music Encoding Conference Proceedings 26-29 May, 2020,* 15–18. Tufts University, Boston. http://dx.doi.org/10.17613/d41w-n008.

Reuse, Timothy de, and Ichiro Fujinaga. 2019. “Robust Transcript Alignment on Medieval Chant Manuscripts.” In *Proceedings of the 2nd International Workshop on Reading Music Systems (WoRMS)*. Delft, Netherlands.

Riley, Jenn, and Ichiro Fujinaga. 2003. “Recommended Best Practices for Digital Image Capture of Musical Scores.” *OCLC Systems & Services: International Digital Library Perspectives* 19 (2): 62–69.

Ríos-Vila, Antonio, Miquel Esplà-Gomis, David Rizo, Pedro J. Ponce de León, and José M. Iñesta. 2021. “Applying Automatic Translation for Optical Music Recognition’s Encoding Step.” *Applied Sciences* 11 (9). https://doi.org/10.3390/app11093890.

Rizo, David, Jorge Calvo-Zaragoza, and José M. Iñesta. 2018. “MuRET: A Music Recognition, Encoding, and Transcription Tool.” In *Proceedings of the 5th International Conference on Digital Libraries for Musicology (DLfM),* 52–56. Paris, France: ACM. https://doi.org/10.1145/3273024.3273029.

———. 2020. “Preservación Del Patrimonio Musical Mediante Transcripción Asistida Por Ordenador.” In *La Música de La Corona d’Aragó: Investigació, Transferència i Educació,* 275–98. Universitat de València.

Rizo, David, Teresa Delgado, Jorge Calvo-Zaragoza, Antonio Madueño, and Patricia García-lasci. 2022. "Speeding-up the Encoding of Mensural Collections from Spanish Libraries." Presented at the International Association of Music Libraries, Archives, and Documentation Centres (IAML). Prague, Czech Republic.

Rizo, David, Nieves Pascual León, and Craig Stuart Sapp. 2018. "White Mensural Manual Encoding: From Humdrum to MEI." *Cuadernos de Investigación Musical*, no. 6: 373–93.

Rizo, David, Beatriz Pascual Sánchez, José M. Iñesta, Antonio Ezquerro Esteban, and Luis Antonio González Marín. 2017. "Towards the Digital Encoding of Hispanic White Mensural Notation." *Anuario Musical* 72: 293–304.

Roesner, Edward H. 2001. "Rhythmic Modes [Modal Rhythm]." In *Grove Music Online*. Oxford University Press. http://www.oxfordmusiconline.com/subscriber/article/grove/music/23337.

Roland, Perry. 2002. "The Music Encoding Initiative (MEI)." In *Proceedings of the First International Conference on Musical Application Using XML*, 55–59. Glasgow, Scotland, UK.

Roland, Perry, Andrew Hankinson, and Laurent Pugin. 2014. "Early Music and the Music Encoding Initiative." *Early Music* 42 (4): 605–11.

Sáenz de Santa María, Carmelo. 1964. *El licenciado don Francisco Marroquín, primer obispo de Guatemala (1499–1563)*. Madrid: Ediciones Cultura Hispánica.

Saleh, Zeyad, Ké Zhang, Jorge Calvo-Zaragoza, Gabriel Vigliensoni, and Ichiro Fujinaga. 2017. "Pixel.Js: Web-Based Pixel Classification Correction Platform for Ground Truth Creation." In *Proceedings of the 12th IAPR International Conference on Graphics Recognition*. Kyoto, Japan: Springer LNCS.

Sapp, Craig Stuart. 2017a. "Verovio Humdrum Viewer." Presented at the Music Encoding Conference (MEC). Tours, France. https://docs.google.com/presentation/d/1vUnojO3OmNt4x9wLOZfFhllyGcf8PTRebcO_57xlep0/edit#slide=id.g2194b8aaa8_1_41.

———. 2017b. "Verovio Humdrum Viewer." Documentation website. Verovio Humdrum Viewer Documentation. https://doc.verovio.humdrum.org/index.html.

———. 2021. "Welcome." Documentation website. Humdrum. https://www.humdrum.org/index.html.

Schele, Linda, and Peter Mathews. 1999. *The Code of Kings: The Language of Seven Sacred Maya Temples and Tombs*. Simon and Schuster.

Schmidt, Thomas Christian, Christian Thomas Leitmeir, and J. P. Gumbert. 2018. *The Production and Reading of Music Sources: Mise-En-Page in Manuscripts and Printed Books Containing Polyphonic Music, 1480-1530*. Centre d'études Supérieures de La Renaissance. Collection "Epitome Musical." Turnhout: Brepols.

Scholes, Percy, Judith Nagley, and James Grier. 2011. "Score." In *The Oxford Companion to Music*, edited by Alison Latham. Oxford University Press. http://www.oxfordmusiconline.com/subscriber/article/opr/t114/e6021.

Schubert, Peter. 2008. *Modal Counterpoint, Renaissance Style*. 2nd ed. New York, NY: Oxford University Press.

Schubert, Peter, and Julie Cumming. 2015. "Another Lesson from Lassus: Using Computers to Analyse Counterpoint." *Early Music* 43 (4): 577–86.

Snow, Robert J., ed. 1996. *A New-World Collection of Polyphony for Holy Week and the Salve Service: Guatemala City, Cathedral Archive, Music MS 4*. Monuments of Renaissance Music 9. Chicago, IL: University of Chicago Press.

———. 2001. "Bermúdez, Pedro." Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.02833.

Snow, Robert J., and Rodrigo de Ceballos. 1980. *The Extant Music of Rodrigo de Ceballos and Its Sources*. Detroit Studies in Music Bibliography, No. 44. Detroit, MI: Information Coordinators.

———. 1995. *Obras completas de Rodrigo de Ceballos*. 5 vols. Granada: Junta de Andalucía, Consejería de Cultura, Centro de Documentación Musical de Andalucía.

Snow, Robert J., and Gaspar Fernandes. 1990. *Gaspar Fernandes, obras sacras*. Portugaliae musica. Serie A. Lisboa: Fundaçao Calouste Gulbenkian, Serviçio de Música.

Stevenson, Robert. 1970. *Renaissance and Baroque Musical Sources in the Americas*. Washington, DC: General Secretariat, Organization of American States.

———. 1980. "Guatemala Cathedral to 1803." *Inter-American Music Review* 2 (2): 27–71.

———. 2001. "Fernandes, Gaspar." Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.09491.

Strunk, W. Oliver, and Leo Treitler. 1998. *Source Readings in Music History*. Vol. 2. New York, NY: Norton.

Thomae, Martha E. 2017. "Automatic Scoring up of Mensural Music Using Perfect Mensurations, 1300-1550." Master's Thesis, Montreal, Canada: McGill University.

———. 2021. “The Guatemalan Choirbooks: Facilitating Preservation, Performance, and Study of the Colonial Repertoire.” In *Christian Music Traditions in the Americas*, edited by Andrew Shenton and Joanna Smolko. New York: Rowman & Littlefield.

Thomae, Martha E., Julie E. Cumming, and Ichiro Fujinaga. 2019. “The Mensural Scoring-Up Tool.” In *Proceedings of the 6th International Conference on Digital Libraries for Musicology (DLfM)*, 9–19. National Library of The Netherlands, The Hague, NL: ACM. https://doi.org/10.1145/3358664.3358668.

———. 2022a. “Digitization of Choirbooks in Guatemala.” In *Proceedings of the 9th International Conference on Digital Libraries for Musicologists (DLfM)*, 19–26. Prague, Czech Republic: ACM. https://doi.org/10.1145/3543882.3543885.

———. 2022b. “Counterpoint Error-Detection Tools for Optical Music Recognition of Renaissance Polyphonic Music.” In *Proceedings of the 23rd International Society for Music Information Retrieval Conference (ISMIR)*. Bengaluru, India.

Tinctoris, Johannes. 1979. *Proportions in Music: Proportionale Musices*. Translated by Albert Seay. Translations (Colorado College. Music Press) 10. Colorado Springs, CO: Colorado College Music Press.

Vigliensoni, Gabriel, John Ashley Burgoyne, Andrew Hankinson, and Ichiro Fujinaga. 2011. “Automatic Pitch Detection in Printed Square Notation.” In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, 423–28. Miami, FL, USA.

Vigliensoni, Gabriel, Gregory Burlet, and Ichiro Fujinaga. 2013. “Optical Measure Recognition in Common Music Notation.” In *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*, 125–30. Curitiba, Brazil.

Vigliensoni, Gabriel, Jorge Calvo-Zaragoza, and Ichiro Fujinaga. 2018. “An Environment for Machine Pedagogy: Learning How to Teach Computers to Read Music.” In *Proceedings of the Intelligent User Interfaces Workshops on Music Interfaces for Listening and Creation*. Tokyo, Japan.

W3C. 2003. “Portable Network Graphics (PNG) Specification (Second Edition).” W3C Recommendation. https://www.w3.org/TR/PNG/.

Wick, Christoph, Alexander Hartelt, and Frank Puppe. 2019. “Staff, Symbol and Melody Detection of Medieval Manuscripts Written in Square Notation Using Deep Fully Convolutional Networks.” *Applied Sciences* 9 (13): 2646–73.

Wick, Christoph, and Frank Puppe. 2019a. “OMMR4all-a Semiautomatic Online Editor for Medieval Music Notations.” In *Proceedings of the 2nd International Workshop on Reading Music Systems (WoRMS)*, 31–34. Delft, Netherlands.

———. 2019b. “OMMR4all-a Semiautomatic Online Editor for Medieval Music Notations.” In *2nd International Workshop on Reading Music Systems*, 31–34.

Wick, Christoph, Christian Reul, and Frank Puppe. 2018. “Calamari: A High-Performance Tensorflow-Based Deep Learning Package for Optical Character Recognition.” *ArXiv Preprint ArXiv:1807.02004*.

Wong, Yue-Ling. 2016. “2. Fundamentals of Digital Imaging.” In *Digital Media Primer: Digital Audio, Video, Imaging, and Multimedia Programming*, Third edition. Boston, MA: Pearson.

Manuscripts

GuatC 1 (or GCA-Gc): Guatemala. Guatemala City. Archivo Histórico Arquidiocesano de Guatemala (AHAG), Archivo Catedralicio, Sección Litúrgica. Libro de Coro 1.

Prints

RISM A/I P 660: Palestrina, Giovanni Pierluigi da. Missarum liber secundus. Rome: Dorico, Valerio, eredi and Luigi Dorico, eredi. 1567. Print.

Appendix A: CPDL Transcriptions

The following are the pieces from the GuatC 1 choirbook with transcriptions in the Choral Public Domain Library (CPDL) *Musica colonial scores* website:

- Missa sobre las vozes from Cristóbal de Morales
- Missa Pere de nous from Pierre Colin
- Missa de 3º tono from Rodrigo de Ceballos
- Missa sine nomine from José de Torres y Martínez Bravo
- Missa O quam gloriosum from Tomás Luis de Victoria
- Missa de 5º tono from Luis Serra
- Missa de 4º tono from Alegre
- Missa de 8º tono from Rodrigo de Ceballos
- Missa de Bomba from Pedro Bermúdez
- Missa sine nomine from Juan Matías de Rivera
- Christus natus est, a 4, from Pedro Bermúdez (piece no. 18 in Table 2-2)
- Christus natus est, a 5, Anonymous (piece no. 20 in Table 2-2)
- Lumen ad revelationem, a 5, from Hernando Franco (piece no. 22 in Table 2-2)
- Surrexit Dominus vere, a 4, Anonymous (piece no. 25 in Table 2-2)

Since GuatC 1 is one of the best-preserved choirbooks in the collection (see Section 3.1.1), it is not surprising that most of the transcriptions in CPDL's Musica colonial scores website that are related to the GuatC choirbooks (on reel 9 of the CIRMA microfilms) come from GuatC 1.

Appendix B: External Documents

Letter of permission sent on 23 October 2018 to the father Eddy René Calvillo, chancellor of the Ecclesiastical Curia of Santiago de Guatemala and director of the AHAG, asking for permission for a pilot project involving the digitization and encoding of one of the choirbooks in the GuatC collection. It shows father Calvillo's signature approving the pilot project.

Octubre 23, 2018

Presbítero Eddy René Calvillo
Canciller de la Curia eclesiástica de Santiago de Guatemala
Director del Archivo Histórico Arquidiocesano de Guatemala

Reverendo padre Eddy René Calvillo:

Mi nombre es Martha Thomae, soy guatemalteca y actualmente estoy estudiando un doctorado en *Música y Tecnología* en la *Universidad de McGill* en Montréal, Canadá. Durante los pasados tres años he estado trabajando en proyectos de preservación de música medieval y renacentista. Mi estudio se ha enfocado en aplicar tecnologías que permiten preservar el patrimonio musical al hacerlo accesible de forma digital. Es mi anhelo poder aplicar estas tecnologías a la música de mi país para preservar los manuscritos musicales de la catedral por largo tiempo—más allá de lo que el manuscrito físico, aún con todos los cuidados requeridos, puede soportar.

En la presente me dirijo a Usted solicitando su permiso para la elaboración de un proyecto piloto que aplique esta tecnología en miras de asegurar la conservación del patrimonio musical de la Iglesia en Guatemala. Quisiera su permiso para trabajar dicho proyecto con uno de los manuscritos musicales de los que dispone el Archivo Histórico Arquidiócesano de Guatemala. Debido al repertorio musical con el cual estoy familiarizada (música del Renacimiento en notación mensural), quisiera poder trabajar dicho proyecto haciendo uso de algún libro de coro en notación mensural.

El proyecto piloto involucraría el obtener imágenes de alta resolución del manuscrito seleccionado. Estas imágenes serán utilizadas únicamente con fines investigativos—no de publicación ni lucro—y una copia de las mismas será donada a la catedral. El objetivo de hacer imágenes de alta resolución, es que éstas son necesarias para aplicar las tecnologías desarrolladas en la Universidad de McGill y que hacen que esta música sea más accesible a personas no expertas en música antigua. El resultado del proyecto piloto sería presentado a Usted y demás interesados dentro del Archivo Histórico. El resultado final permitiría presentar la música en un formato más asequible para los "no expertos" y escuchar la pieza. Todo lo generado por dicho proyecto—las imágenes de alta calidad, la nueva presentación de las piezas y el audio—será donado a la Iglesia, para que ésta disponga del material que ha sido generado a partir de sus manuscritos.

Nuevamente, solicito su permiso para llevar a cabo este proyecto piloto. Considero que los resultados del mismo facilitarían el conocimiento de las piezas del repertorio sacro de la Guatemala de los siglos XVI y XVII por parte tanto de músicos no expertos en música antigua y de personas sin formación musical alguna. De igual importancia, la labor de digitalización de este proyecto facilitaría la preservación de la música catedralicia, ya que la imagen digital no sólo reduciría el uso y consecuente desgaste del manuscrito, si no que funcionaría como un sustituto en caso de algún desastre natural.

Agradezco mucho su atención. Sin otro particular, se despide de usted respetuosamente,

Martha E. Thomae

¿Cuándo inicia?
¿duración?

Vo. Bo.

Conservation report for the GuatC 1–6 collection written by *Centro de Rescate, Estudios y Análisis Científico del Arte* (*CREA*) on 23 November 2018. A few issues with the metadata for some of the books: **(1)** the date of choirbook 1 (*Libro de Coro No. 1*) should be XVIII century, and **(2)** the metadata of choirbooks 2 (*Libro de Coro No. 2*) and 3 (*Libro de Coro No. 3*) is inverted.

FUNDACIÓN CREA ROZAS-BOTRÁN
Centro de Rescate, Estudio y Análisis Científico para el Arte

Guatemala, 23 de noviembre del 2018

Licenciada
Martha Thomae
Presente

Estimada licenciada Thomae:

Por medio de la presente me permito saludarle nuevamente y desearle éxitos en sus labores diarias. El motivo de mi comunicación primeramente es para agradecerle la invitación que nos ha realizado a la Fundación CREA para formar parte en el proceso de conservación y restauración de los libros de coro del Archivo Arquidiocesano de la Catedral Metropolitana de la Ciudad de Guatemala.

Habiendo realizado previamente las evaluaciones y estudios respectivos de seis (6) libros de coro, hacemos entrega de los documentos del estado de conservación de cada uno de ellos con una breve propuesta de intervención para realizar después del interesante proceso de digitalización que usted realizará.

Sin otro particular y esperando los documentos adjunto sean funcionales para su investigación me despido,

Atentamente,

Arq. Ana Lucía González Muñoz
Directora Fundación CREA
Tel. 5440-5581

Centro de Rescate, Estudio y Análisis Científico para el Arte

DIAGNÓSTICO

El archivo Histórico Arquidiocesano de Guatemala alberga gran cantidad de documentos eclesiásticos importantes desde la fundación de la ciudad. Dentro de la colección se encuentran libros de coro de diferentes dimensiones y épocas que se han ido deteriorando por las características de sus materiales de fabricación y factores ambientales.

Dado su valor histórico, es importante realizar acciones que permitan la conservación de su contenido por lo que acudimos a la petición de Martha Thomae y autoridades del Archivo Arquidiocesano, de realizar la evaluación de dichos libros para llevar a cabo la conservación preventiva antes de ejecutar el proyecto de digitalización.

Es importante mencionar que el fin de la conservación y restauración del patrimonio, en este caso documental, es intervenir las obras a modo de facilitar su comprensión y frenar el deterioro que las pone en riesgo de desaparecer, velando por que se respete en medida de lo posible su integridad histórica.

En los fondos documentales se pueden encontrar diferentes grados de deterioro que se van a tratar de acuerdo a la necesidad de la pieza o su valor cultural, histórico, artístico, etc. Los seis libros evaluados datan de los siglos XVI y XVII, su soporte es papel artesanal y presentan deterioros similares provocados principalmente por la oxidación de las tintas, humedad y ataque biológico lo cuales se describen a continuación.

ANÁLISIS TECNOLÓGICO

Éste análisis fue realizado mediante un examen organoléptico. Los datos pueden ser comprobados mediante estudios tales como: Solubilidad de tintas, medición de pH, análisis de lignina, análisis de fibras, entre otros.

Soporte: Los seis libros evaluados se encuentran sobre un soporte de papel que reúne las características del papel realizado a mano. Llamado también papel verjurado puesto que tienen las huellas de los puntizones y corondeles de la forma metálica utilizada desde el siglo XIII hasta el siglo XIX. También presentan sellos de agua o filigrana que eran utilizados por cada molino papelero para diferenciarse de otros fabricantes.

Figura 1. Sello de agua o filigrana encontrado en un folio del libro de coro No. 2

Elemento sustentado (Tintas): Destaca la presencia de tintas metaloácidas que eran fabricadas a partir de la combinación de un compuesto ácido y una sal de hierro. Estas tintas presentan gran inestabilidad química, debido a la transformación del sulfato ferroso en ácido sulfúrico el cual es altamente corrosivo y daña gravemente el soporte de papel. También se encuentran dos tipos de tintas diferentes, que puede ser a base de carbón por su tonalidad y anotaciones posteriores con tinta azul.

ESTADO DE CONSERVACIÓN

Libro de Coro No. 1	
Tipo de objeto:	Libro
Título:	Libro de Coro No. 1
Medidas:	44.5 cm alto x 30 cm ancho x 5 cm grosor
Época:	Siglo XVII
Soporte:	Papel Verjurado

Descripción de daños:

Empastado: Presenta suciedad superficial, ataque biológico, desprendimiento del lomo, desgaste de la fibra.

Figura 2. Pérdida del lomo en extremos.

Figura 3. Ataque biológico y desprendimiento del

Figura 4. Roturas y desgaste del forro.

Soporte: Presenta suciedad superficial, Amarillamiento, manchas de humedad, roturas en los bordes, desprendimiento de la última hoja.

Figura 5. Roturas y desgaste de la fibra en los bordes.

Figura 6. Amarillamiento y evidencia de manchas provocadas por humedad.

Figura 7. Lagunas ocasionadas por ataque biológico.

Tintas: Estables.

km. 28.5, Carretera a Bárcenas, Parcelamiento Santa Rosa, Santa Lucía Milpas Altas, Sacatepéquez
direccionadministrativa@crea.com.gt / Tel. +502 5440 5581

FUNDACIÓN CREA ROZAS-BOTRÁN
Centro de Rescate, Estudio y Análisis Científico para el Arte

Propuesta de tratamiento
1. Análisis Previos: a. Solubilidad de tintas b. Medición de pH c. Análisis de fibras 2. Desinfección 3. Limpieza Mecánica 4. Tratamientos en húmedo a. Hidratación b. Eliminación de manchas 5. Secado y aplanado 6. Consolidación 7. Reforzamiento del encuadernado **Observaciones:** Por las características de los daños, si se puede realizar la digitalización antes de su intervención.

Libro de Coro No. 2	
Tipo de objeto:	Libro
Título:	Libro de Coro No. 2
Medidas:	42.5 cm alto x 27.5 cm ancho x 5 cm grosor
Época:	Siglo XVII
Soporte:	Papel Verjurado

Descripción de daños:

Empastado: Pasta en mal estado de conservación, tapas separadas, se encuentra protegido con dos tapas de carton provisionales, se observan desgastes en el forro, roturas, faltantes y lomo atacado por agentes biológicos. Costuras perdidas por completo. se conserva unicamente un fragmento del lomo.

Figura 8. Tapas provisionales a manera de conservación preventiva.

Figura 9. Roturas, faltantes y desgaste en el forro.

Figura 10. Costuras perdidas completamente.

Soporte: Presenta amarillamiento, roturas, dobleces, lagunas, injertos y pérdida de borde de encuadernado a causa de ataque biológico.
Tintas: Se observan 3 tintas diferentes de las cuales solo la tinta de tonalidad marrón presenta desgaste.

Figura 11. Se observan dos bloques de folios diferentes en cuanto a la resistencia del papel y las tintas, pertenecientes al mismo libro.

Figura 12. Amarillamiento, desgaste de fibra en el borde de encuadernación e Intervenciones anteriores a nivel conservación.

Figura 13. Evidencia de anotaciones posteriores utilizando tinta de color azul.

FUNDACIÓN CREA ROZAS-BOTRÁN
Centro de Rescate, Estudio y Análisis Científico para el Arte

Propuesta de tratamiento:
1. Análisis Previos: a. Solubilidad de tintas b. Medición de pH c. Análisis de fibras 2. Desinfección 3. Limpieza Mecánica 4. Tratamientos en húmedo a. Hidratación b. Eliminación de manchas 5. Secado y aplanado 6. Eliminación de dobleces 7. Consolidación de bordes 8. Empastado (solo si se conoce el orden de los folios) **Observaciones:** Por las características de los daños, si se puede realizar la digitalización antes de su intervención.

FUNDACIÓN CREA ROZAS-BOTRÁN
Centro de Rescate, Estudio y Análisis Científico para el Arte

Libro de Coro No. 3	
Tipo de objeto:	Libro
Título:	Libro de Coro No. 3
Medidas:	40 cm alto x 27.5 cm ancho x 6 cm grosor
Época:	Siglo XVII-XVIII
Soporte:	Papel Verjurado
Descripción de daños:	

Empastado: Presenta suciedad superficial, desgaste de la fibra en bordes de la pasta y forro.

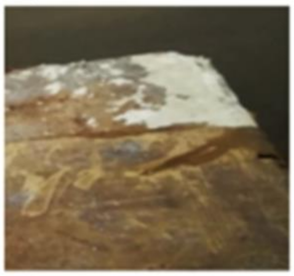


Figura 14. Desgaste de fibra.

Figura 15. Faltante de forro y desgaste fibra.

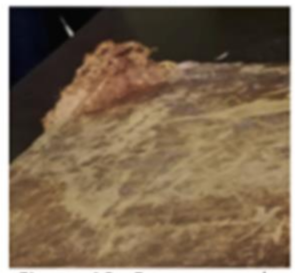


Figura 16. Forro en mal estado.

Soporte: Presenta suciedad superficial, amarillamiento, roturas por manipulación, dobleces, lagunas, injertos, folios cubiertos (pegados sobre otro).
Tintas: Oxidación leve.

Figura 17. Amarillamiento, manchas de humedad e intervenciones anteriores.

Figura 18. Deformación del plano a causa de la humedad y oxidación de las tintas.

Figura 19. Huella de puntizones de la forma metálica utilizada en la fabricación.

km. 28.5, Carretera a Bárcenas, Parcelamiento Santa Rosa, Santa Lucía Milpas Altas, Sacatepéquez
direccionadministrativa@crea.com.gt / Tel. +502 5440 5581

Centro de Rescate, Estudio y Análisis Científico para el Arte

Propuesta de tratamiento:
1. Análisis Previos: a. Solubilidad de tintas b. Medición de pH c. Análisis de fibras 2. Desinfección 3. Limpieza Mecánica 4. Tratamientos en húmedo a. Hidratación b. Eliminación de injertos c. Eliminación de manchas 5. Secado y aplanado 6. Consolidación de bordes 7. Reforzamiento del empastado **Observaciones:** Por las características de los daños, si se puede realizar la digitalización antes de su intervención.

Libro de Coro No. 4	
Tipo de objeto:	Libro
Título:	Libro de Coro No. 4
Medidas:	41.8 cm alto x 27.4 cm ancho x 7.5 cm grosor
Época:	Siglo XVII
Soporte:	Papel Verjurado

Descripción de daños:

Empastado: Pasta de madera, presenta suciedad superficial, pérdida del forro en borde inferior, y desgaste de fibra.

Figura 20. Detalles del empastado.

Figura 21. Esquinas dañadas.

Figura 22. Separación de folios mediante el uso de pellum no permite que el libro cierre de forma adecuada.

Soporte: Presenta suciedad superficial, roturas, amarillamiento, manchas provocadas por la oxidación de las tintas, friabilidad, dobleces y desprendimiento de folios.

Tintas: Tintas ferrogálicas, presentan oxidación y pérdida del contenido por desprendimientos.

Figura 23. Separación de folios utilizando pellum.

Figura 24. Daños provocados por la corrosión de las tintas.

Figura 25. Oxidación del soporte por reacción de las tintas.

km. 28.5, Carretera a Bárcenas, Parcelamiento Santa Rosa, Santa Lucía Milpas Altas, Sacatepéquez
direccionadministrativa@crea.com.gt / Tel. +502 5440 5581

FUNDACIÓN CREA ROZAS-BOTRÁN
Centro de Rescate, Estudio y Análisis Científico para el Arte

Propuesta de tratamiento:
1. Análisis Previos: a. Solubilidad de tintas b. Medición de pH c. Análisis de fibras 2. Desinfección 3. Limpieza Mecánica 4. Desmontaje 5. Tratamientos en húmedo a. Lavado i. Hidratación ii. Limpieza iii. Desacidificación 6. Secado y aplanado 7. Consolidación 8. Reforzamiento del empastado **Observaciones:** Por las características de los daños, es preferible no realizar la digitalización antes de su intervención.

Libro de Coro No. 5	
Tipo de objeto:	Libro
Título:	Libro de Coro No. 5
Medidas:	40.7 cm alto x 28.5 cm ancho x 3.3 cm grosor
Época:	Siglo XVII - XVIII
Soporte:	Papel Verjurado
Descripción de daños:	

Empastado: Presenta suciedad superficial, desgastes en los bordes y de desgaste en el lomo.

Figura 26. Detalles del empastado. Presenta golpes y desgastes en los bordes.

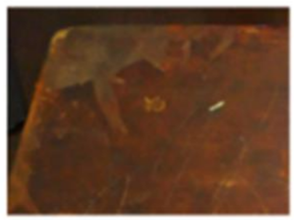


Figura 27. Daños en el forro principalmente en las esquinas.

Figura 28. Folios expuestos por pérdida del lomo.

Soporte: Presenta suciedad superficial, manchas provocadas por la oxidación de las tintas, roturas, injertos y faltantes.
Tintas: Oxidación y pérdida de material.

Figura 29. Faltante de soporte, primera hoja mutilada.

Figura 30. Amarillamiento y manchas provocadas por la oxidación de las tintas.

Figura 31. Oxidación de las tintas.

Centro de Rescate, Estudio y Análisis Científico para el Arte

Propuesta de tratamiento:
1. Análisis Previos: a. Solubilidad de tintas b. Medición de pH c. Análisis de fibras 2. Desinfección 3. Limpieza Mecánica 4. Tratamientos en húmedo a. Hidratación b. Eliminación de manchas 5. Secado y aplanado 6. Consolidación 7. Reforzamiento del empastado **Observaciones:** Por las características de los daños, es preferible no realizar la digitalización antes de su intervención.

Libro de Coro No. 6	
Tipo de objeto:	Libro
Título:	Libro de Coro No. 6
Medidas:	cm alto x cm ancho x cm grosor
Época:	Siglo XVII
Soporte:	Papel Verjurado
Descripción de daños:	

Empastado: No se encuentra empastado, está almacenado en un sobre.
Soporte: Presenta suciedad superficial, Amarillamiento, manchas leves provocadas por la oxidación de la tinta y dobleces.
Tintas: Se encuentran Oxidadas.

Figura 32. Oxidación de las tintas. Visibles en el reverso.

Figura 33. Documento doblado por la mitad.

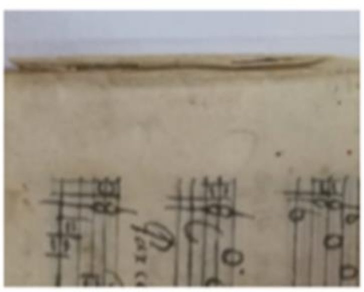


Figura 33 y 34. Dobleces en los bordes.

km. 28.5, Carretera a Bárcenas, Parcelamiento Santa Rosa, Santa Lucía Milpas Altas, Sacatepéquez
direccionadministrativa@crea.com.gt / Tel. +502 5440 5581

FUNDACIÓN CREA ROZAS-BOTRÁN
Centro de Rescate, Estudio y Análisis Científico para el Arte

Propuesta de tratamiento:
1. Análisis Previos: a. Solubilidad de tintas b. Medición de pH c. Análisis de fibras 2. Desinfección 3. Limpieza Mecánica 4. Eliminación de dobleces 5. Tratamientos en húmedo a. Hidratación b. Eliminación de manchas 6. Secado y aplanado 7. Elaboración de carpeta libre de ácido para su almacenamiento. **Observaciones:** Por las características de los daños, si se puede realizar la digitalización antes de su intervención.